HAL open science

Weiss-Weinstein bound on multiple change-points estimation

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

To cite this version:

Lucien Bacharach, Alexandre Renaux, Mohammed Nabil El Korso, Eric Chaumette. Weiss-Weinstein bound on multiple change-points estimation. IEEE Transactions on Signal Processing, 2017, 65 (10), pp.2686-2700. 10.1109/TSP.2017.2673804 . hal-01525496

HAL Id: hal-01525496

https://centralesupelec.hal.science/hal-01525496v1

Submitted on 21 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Weiss-Weinstein Bound on Multiple Change-Points Estimation

Lucien Bacharach, *Student Member, IEEE,* Alexandre Renaux, *Member, IEEE,* Mohammed Nabil El Korso, and Éric Chaumette

***Abstract*—In the context of multiple change-points estimation, performance analysis of estimators such as the maximum likelihood is often difficult to assess since the regularity assumptions are not met. Focusing on the estimators variance, one can however use lower bounds on the mean square error. In this paper, we derive the so-called Weiss-Weinstein bound (WWB) which is known to be an efficient tool in signal processing to obtain a fair overview of the estimation behavior. Contrary to several works about performance analysis in the change-point literature, our study is adapted to multiple changes. First, useful formulas are given for a general estimation problem whatever the considered distribution of the data. Second, closed-form expressions are given in the cases of i) Gaussian observations with changes in the mean and/or the variance, and ii) changes in the mean rate of a Poisson distribution. Furthermore, a semi-definite programming formulation of the minimization procedure is given in order to compute the tightest WWB. Specifically, it consists of finding the unique minimum volume covering the set constituted by hyper-ellipsoid elements which are generated using the derived candidate WWB matrices w.r.t. the so-called Loewner partial ordering. Finally, simulation results are provided to show the good behavior of the proposed bound.**

***Index Terms*—Weiss-Weinstein bound, change-point, mean square error, Bayesian lower bound**

I. Introduction

NON stationary signals are often encountered in many practical applications. The possible causes of such non stationarities are extensive, which results in numerous ways of characterizing them. One type of non stationary signals are those submitted to one or several abrupt changes. This means that, at some time instants (generally unknown), the characteristics of the signal of interest change almost instantaneously. More precisely, as the signal is often considered random, such changes can be reflected by shifts in the parameters of its distribution.

They are usually referred to as "change-points" in the literature, and arise in many signal processing applications such as instrument fault detection, system monitoring or speech processing (see e.g., [1] for an overview of the main potential applications). Depending on the application and its purpose, the problem of change-points in a signal can be formulated in different manners: i) offline detection, in which the aim is to decide whether there is a change in a given finite data set, ii) online detection, in which the aim is still to detect a change but data are assumed to be received gradually [1]. At last, the context of iii) offline estimation is adequate to locate a change as precisely as possible, once it is known to occur, see e.g., [2]. This paper deals with this latter formulation.

Among all the possible estimation schemes, the maximum likelihood estimator (MLE) is often preferred for its good statistical properties. Of course, the ideal achievement when one is interested in performance analysis is to obtain the distribution of the latter estimator. However, in the context of change- point estimation, certain regularity assumptions, usually used to prove the asymptotic normality of the MLE, are obviously not fulfilled (e.g., see Theorem 3.10 in [3, Chapter 6]), especially because of the discrete nature of the parameter to estimate. Consequently, the study of such estimator performance requires a specific analysis. For a single change-point, the asymptotic distribution of the MLE was first derived by Hinkley under certain assumptions [4]. Later on, Fotopoulos and Jandhyala derived exact computable forms of this distribution for exponentially distributed data [5], as well as Gaussian data [6].

It should be noticed that all these results have been obtained in the asymptotic context and for a single change. Therefore, one has to take interest in other ways of characterizing the estimation performance as soon as such assumptions do not hold. One of them is to work on the moments of the estimator's distribution. Even if those are not necessarily easily accessible either, lower bounds can be used to overcome these difficulties as they generally require less complex computations and can yield closed-form expressions in some specific cases.

Among the plethora of lower bounds on the Mean Square Error (MSE), the Cramér-Rao Bound (CRB) is the most famous in the signal processing community. Indeed, the attractiveness of the CRB comes from the fact that various closed-form expressions are available for a large class of observation models (see e.g., [7]–[9]. However, even if the CRB is known to be tight in the asymptotic region [3], [10], [11], it is unfortunately not the case for low SNR and/or low number of observations. An important point is that in our context of unknown discrete parameters, the regularity conditions to apply the CRB are not fulfilled since the likelihood involved in the Fisher information computation is not differentiable w.r.t. these parameters. As a first alternative to this fence, the regularization of the problem by approximating the signals by smoother one, has been studied, e.g., in [12], [13]. A second way, more natural, of overcoming the aforementioned limitations is to

L. Bacharach and A. Renaux are with Laboratory of Signals and Systems (L2S), Université Paris-Sud, 91192 Gif-sur Yvette, France.

M. N. El Korso is with Laboratory Energetics, Mechanics and Electromagnetism (LEME), Université Paris-Ouest, 91410 Ville d'Avray, France.

É. Chaumette is with Department of Electronics, Optronics and Signal, Institut Supérieur de l'Aéronautique et de l'Espace (ISAE), 31055 Toulouse, France.

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), and by the DGA MRIS.

use lower bounds with less regularity conditions, particularly lower bounds which do not involve the differentiability of the likelihood. To the best of our knowledge, only the Barankin bound [14]–[16] has been applied to change point estimation in the case of a single change point in [17], and then extended to multiple changes in [18].

In both cases, the resulting bounds appeared to be quite coarse. Lower bounds on the *global* MSE (GMSE) also exist in the Bayesian context but they received quite little attention in the change-points literature, even though early works in this area as [19], [20] were conducted using a Bayesian approach. The Weiss-Weinstein bound (WWB) is known to be one of the tightest Bayesian bounds, and for this reason, its derivation would be of great interest for change-point estimation. We recently derived it for a single change-point [21], and the aim of this paper is to generalize this study to the context of multiple changes. Note that even if the WWB is a Bayesian bound, it can still be used to assess the performance of the MLE in terms of global MSE, see e.g., [22]–[25]. Finally, the tightest WWB is computed in a fast and efficient manner using a convex equivalent formulation of the original maximization procedure.

This paper is organized as follows: in Section II, the observation model is presented and the main assumptions are stated. Section III briefly recalls the general expression of the WWB. Its derivation is achieved in Section IV for any data distribution, and this result is then applied to two particular distributions, namely the Gaussian and the Poisson distributions, in Section V. Finally, numerical results are given in Section VII to assess the behavior of the proposed bound.

II. Problem setup

We consider an independent multivariate time series $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{M\times N}$, where N denotes the number of observations and M the length of each vector $\boldsymbol{x}_i$ (possibly corresponding to the number of sensors, e.g., in the array processing context), in which a total number of change-points Q are known to occur ($Q < N$) at unknown time instants t_q, $1 \leq q \leq Q$. This can be modeled as follows:

$$\begin{cases} \boldsymbol{x}_i \sim p_{\boldsymbol{\eta}_1}(\boldsymbol{x}_i) & \text{for } i = 1, \ldots, t_1 \\ \boldsymbol{x}_i \sim p_{\boldsymbol{\eta}_2}(\boldsymbol{x}_i) & \text{for } i = t_1+1, \ldots, t_2 \\ \quad\vdots & \quad\vdots \\ \boldsymbol{x}_i \sim p_{\boldsymbol{\eta}_{Q+1}}(\boldsymbol{x}_i) & \text{for } i = t_Q+1, \ldots, N \end{cases} \tag{1}$$

where $p_{\boldsymbol{\eta}_q}$ is the probability distribution (discrete or continuous, depending on the nature of the values taken by the observations) on the q-th segment, i.e., between two consecutive change-points t_{q-1} and t_q, $q \in \{1, \ldots, Q+1\}$, t_0 and t_{Q+1} being defined as $t_0 \triangleq 0$ and $t_{Q+1} \triangleq N$. These distributions are parameterized by a vector $\boldsymbol{\eta}_q$ (e.g., in the case of a Gaussian law, $\boldsymbol{\eta}_q$ includes the mean and the covariance matrix), and they are assumed to belong to the same family of distribution.

Note that since the framework of our study is the estimation performance of algorithms such as the Maximum Likelihood estimator (or the Maximum *A Posteriori* estimator), the number of changes Q has to be assumed known. This assumption is usually made in the mathematical change-point literature, dealing with the distribution of such estimators [4], [6], [26]. There are also practical examples in which Q is *a priori* known [27], [28], [2, page 221]. In addition, one can consider that a preliminary detection step giving the number of changes has been operated as in [29], [30].

The unknown parameter vector is $\boldsymbol{t} = [t_1, \ldots, t_Q]^T$, where $t_q \in \mathbb{N} \setminus \{0\}$, $q = 1, \ldots, Q$. In the Bayesian framework, the unknown parameters are random variables whose *a priori* distribution has to be properly chosen. Since the total number of changes is known, a practical distribution that is compatible with this assumption is a uniform random walk, i.e., for all $q \in \{1, \ldots, Q\}$:

$$t_q = t_{q-1} + \varepsilon_q \tag{2}$$

where ε_q are i.i.d. variables following a discrete uniform distribution on the finite set of integers $\{1, \ldots, \tau\}$, $\tau \in \mathbb{N}\setminus\{0\}$ being the maximum gap between two consecutive change-points: $\Pr(\varepsilon_q = k) = 1/\tau$ if $k \in \{1, \ldots, \tau\}$, otherwise it is zero. As a consequence, τ is defined in such a way that the last change-point t_Q does not exceed the total number of observations, i.e., $t_Q < N$. This leads to the fact that the product τQ does not exceed $N-1$, or equivalently,

$$\tau \leq \left\lfloor \frac{N-1}{Q} \right\rfloor \tag{3}$$

where $\lfloor . \rfloor$ denotes the floor function.

In order to obtain the distribution of $\boldsymbol{t}$, we write the distribution of each t_q conditionally to t_{q-1} as:

$$\begin{aligned} &\Pr(t_q = k_q | t_{q-1} = k_{q-1}) \\ &\quad = \Pr(t_{q-1} + \varepsilon_q = k_q | t_{q-1} = k_{q-1}) \\ &\quad = \Pr(\varepsilon_q = k_q - k_{q-1}) \\ &\quad = \begin{cases} \frac{1}{\tau} & \text{if } k_{q-1}+1 \leq k_q \leq k_{q-1}+\tau \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{4}$$

The *a priori* joint distribution of $\boldsymbol{t}$ can then be written as:

$$\Pr(\boldsymbol{t} = \boldsymbol{k}) = \begin{cases} \left(\frac{1}{\tau}\right)^Q & \text{if } k_q \in \mathcal{J}_q \text{ for all } q \in \{1, \ldots, Q\} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where, for any integer $q \in \{1, \ldots, Q\}$, we define the set of integers $\mathcal{J}_q \triangleq \{k_{q-1}+1, \ldots, k_{q-1}+\tau\}$.

In the next section, we recall the definition of the WWB in a general framework.

III. Background on the WWB

For any Bayesian estimator $\hat{\boldsymbol{t}}(\boldsymbol{X})$ of the parameter vector $\boldsymbol{t}$ lying in the parameter space Θ (that can be a subset of $\mathbb{Z}^Q$ or $\mathbb{R}^Q$, depending on the application), the multiparameter version of the WWB as given by [31] satisfies the following matrix inequality:

$$\mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\left\{\left[\hat{\boldsymbol{t}}(\boldsymbol{X}) - \boldsymbol{t}\right]\left[\hat{\boldsymbol{t}}(\boldsymbol{X}) - \boldsymbol{t}\right]^T\right\} \succeq \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^T \tag{6}$$

in which $\boldsymbol{A} \succeq \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is a non negative matrix, $\mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\{.\}$ denotes the expectation operation w.r.t. the joint distribution between the random vectors $\boldsymbol{X}$ and $\boldsymbol{t}$, $\boldsymbol{H} = \left[\tilde{\boldsymbol{t}}_1 - \boldsymbol{t}, \ldots, \tilde{\boldsymbol{t}}_Q - \boldsymbol{t}\right] \in \Theta^Q$ where, for all $q \in \{1, \ldots, Q\}$, $\tilde{\boldsymbol{t}}_q$

denote the so-called test-points. Those test points lie in Θ, and their choice is left to the user. For simplicity, we define the vectors $\boldsymbol{h}_q = \tilde{\boldsymbol{t}}_q - \boldsymbol{t}$ so that $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_Q]$. Finally, the $Q \times Q$ matrix $\boldsymbol{G}$ is defined by:

$$[\boldsymbol{G}]_{m,n} = \frac{\mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\left\{\begin{array}{l}\left[L^{s_m}(\boldsymbol{X},\boldsymbol{t}+\boldsymbol{h}_m,\boldsymbol{t})-L^{1-s_m}(\boldsymbol{X},\boldsymbol{t}-\boldsymbol{h}_m,\boldsymbol{t})\right]\\ \times\left[L^{s_n}(\boldsymbol{X},\boldsymbol{t}+\boldsymbol{h}_n,\boldsymbol{t})-L^{1-s_n}(\boldsymbol{X},\boldsymbol{t}-\boldsymbol{h}_n,\boldsymbol{t})\right]\end{array}\right\}}{\mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\left\{L^{s_m}(\boldsymbol{X},\boldsymbol{t}+\boldsymbol{h}_m,\boldsymbol{t})\right\}\mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\left\{L^{s_n}(\boldsymbol{X},\boldsymbol{t}+\boldsymbol{h}_n,\boldsymbol{t})\right\}} \quad (7)$$

where $L(\boldsymbol{X},\boldsymbol{\varphi},\boldsymbol{\psi}) \triangleq p(\boldsymbol{X},\boldsymbol{\varphi})/p(\boldsymbol{X},\boldsymbol{\psi})$. The inequality (6) holds for any combination of $\boldsymbol{h}_q$, $s_q \in]0,1[$ such that $\boldsymbol{G}$ is invertible, and maximizing the right side of (6) w.r.t. these variables leads to the tightest WWB, i.e.,

$$\mathbf{WWB} = \sup_{\substack{\boldsymbol{h}_1,\ldots,\boldsymbol{h}_Q\\ s_1,\ldots,s_Q}} \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^T \quad (8)$$

where the supremum operation is taken w.r.t. Loewner partial ordering [32] (see Section VI).

Note that with this formulation, the computation of the supremum has to be made over a total of (Q^2+Q) parameters, which is computationally prohibitive and leads to unfeasible calculations. For these reasons, and since inequality (6) is valid for any matrix $\boldsymbol{H}$, we will use a restricted version of the WWB where we assume, for $q = 1, \ldots, Q$, $\boldsymbol{h}_q = [0, \ldots, h_q, \ldots, 0]^T$, i.e., $\boldsymbol{H} = \mathrm{diag}(h_1, \ldots, h_Q)$. As explained in [18], this form makes it possible to derive closed-form expressions of $\boldsymbol{G}$, and then to compute the WWB efficiently. Under this assumption, the maximization has to be performed w.r.t. a reduced number of parameters of $2Q$. This number could even be reduced to Q by making the choice $s_q = 1/2$, $q = 1, \ldots, Q$, since it has been noticed, for some applications, that this value leads to the tightest bound [33]. However, in order to be extensive and contrary to [21], we do not make such a restrictive assumption. Specifically, the following derivation of the bound is done for any value of s_q in order to verify whether the optimal value of s_q can differ from $1/2$ for the specific problem of multiple change-points.

In order to derive $\boldsymbol{G}$ in a compact form, let us introduce the function ζ defined, for $(\alpha,\beta) \in]0,1[\times]0,1[$ and $\boldsymbol{u},\boldsymbol{v}$ being $Q \times 1$ vectors such that $\boldsymbol{t}+\boldsymbol{u} \in \Theta$ and $\boldsymbol{t}+\boldsymbol{v} \in \Theta$, by:

$$\zeta(\alpha,\beta,\boldsymbol{u},\boldsymbol{v}) \triangleq \mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\left\{L^{\alpha}(\boldsymbol{X},\boldsymbol{t}+\boldsymbol{u},\boldsymbol{t})\,L^{\beta}(\boldsymbol{X},\boldsymbol{t}+\boldsymbol{v},\boldsymbol{t})\right\}. \quad (9)$$

With this definition, we rewrite (7) as:

$$[\boldsymbol{G}]_{m,n} = \frac{\left(\begin{array}{l}\zeta(s_m,s_n,\boldsymbol{h}_m,\boldsymbol{h}_n)\\ +\zeta(1-s_m,1-s_n,-\boldsymbol{h}_m,-\boldsymbol{h}_n)\\ -\zeta(s_m,1-s_n,\boldsymbol{h}_m,-\boldsymbol{h}_n)\\ -\zeta(1-s_m,s_n,-\boldsymbol{h}_m,\boldsymbol{h}_n)\end{array}\right)}{\zeta(s_m,0,\boldsymbol{h}_m,\boldsymbol{0})\,\zeta(s_n,0,\boldsymbol{h}_n,\boldsymbol{0})}. \quad (10)$$

Since matrix $\boldsymbol{G}$ is symmetric, it is possible to only derive the diagonal and upper triangle terms, i.e., to assume $m \leq n$. The remaining terms (those of the lower triangle of $\boldsymbol{G}$) are directly deduced by symmetry.

Then, finding a closed-form expression of $\zeta(\alpha,\beta,\boldsymbol{u},\boldsymbol{v})$ directly implies finding that of $\boldsymbol{G}$, which is the cornerstone of the WWB derivation. In the next section, we give the main steps to achieve this task in our multiple change-points context.

IV. Derivation of the WWB

In this section, it is assumed that the observations $\boldsymbol{X}$ have a probability density function (p.d.f.), although the equations remain valid if their range space is discrete. Taking model (1) into account (particularly the discrete nature of the unknown parameters), we rewrite ζ as:

$$\zeta(\alpha,\beta,\boldsymbol{u},\boldsymbol{v}) = \sum_{\boldsymbol{k}\in\mathcal{I}}\int_{\Omega}\frac{p^{\alpha}(\boldsymbol{X},\boldsymbol{t}=\boldsymbol{k}+\boldsymbol{u})\,p^{\beta}(\boldsymbol{X},\boldsymbol{t}=\boldsymbol{k}+\boldsymbol{v})}{p^{\alpha+\beta-1}(\boldsymbol{X},\boldsymbol{t}=\boldsymbol{k})}\mathrm{d}\boldsymbol{X} \quad (11)$$

where Ω is the observation space, and $p(\boldsymbol{X},\boldsymbol{t}=\boldsymbol{k})$ denotes the joint distribution between the random vectors $\boldsymbol{X}$ and $\boldsymbol{t}$, and where $\boldsymbol{k} = [k_1, \ldots, k_Q]^T$ is the value taken by the random vector $\boldsymbol{t}$. In other words, $p(\boldsymbol{X},\boldsymbol{t}=\boldsymbol{k}) = p(\boldsymbol{X}|\boldsymbol{t}=\boldsymbol{k})\Pr(\boldsymbol{t}=\boldsymbol{k})$. Notice that the sum $\sum_{\boldsymbol{k}\in\mathcal{I}}(.)$ in (11) has to be understood as the Q-fold sum $\sum_{k_1\in\mathcal{I}_1}\cdots\sum_{k_Q\in\mathcal{I}_Q}(.)$, in which we define the set $\mathcal{I} \subset \mathbb{Z}^Q$ as the range of vector $\boldsymbol{k}$ such that the terms in the sum $\sum_{\boldsymbol{k}\in\mathcal{I}}(.)$ are nonzero, and the sets $\mathcal{I}_q$ are such that their cartesian product $\prod_{q=1}^{Q}\mathcal{I}_q = \mathcal{I}$. We give an explicit expression of these sets in Appendix.

Since we assumed the matrix $\boldsymbol{H}$ diagonal, we restrain to vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ under the form $\boldsymbol{u} = \boldsymbol{u}_m = [0, \ldots, u_m, \ldots, 0]^T$, whose elements are all zero except the m-th element, and $\boldsymbol{v} = \boldsymbol{v}_n = [0, \ldots, v_n, \ldots, 0]^T$ in which v_n can be 0, as in the denominator of (10). Then, with abuse of notation, we replace vectors $\boldsymbol{u}_m$ and $\boldsymbol{v}_n$ with their only nonzero value u_m and v_n in the arguments of the function ζ and everywhere it is not necessary.

A. Expression of $\zeta(\alpha,\beta,u_m,v_n)$

In this section, we give the expression of $\zeta(\alpha,\beta,u_m,v_n)$, which enables to obtain $[\boldsymbol{G}]_{m,n}$ using (10). This expression can actually be written in several different manners, depending on the three cases (i) $n = m$ (case named "D", for "diagonal"); (ii) $n > m+1$ (case "UT", for "upper triangle"); and (iii) $n = m+1$ (case "FSD", for "first superdiagonal"). All the technical details leading to the following results are given in Appendix.

1) Case D ($n = m$): In this case, as it can be seen from the numerator of (10), u_m and v_m can either be the same, i.e., $u_m = v_m$, or opposed, i.e., $u_m = -v_m$. In addition, the terms in the denominator of (10) can also be derived under this scope using the special case $v_m = 0$. Then, by using the following discrete step-function U_γ defined for $\gamma \in \mathbb{R}$ as

$$U_\gamma(n) = \begin{cases}\gamma & \text{if } n \geq 0\\ 1-\gamma & \text{if } n < 0,\end{cases} \quad (12)$$

we obtain (see Appendix)

$$\zeta(\alpha,\beta,u_m,u_m) = f_D(\tau,u_m,u_m)\,\rho_m^{|u_m|}\big(U_{\alpha+\beta}(u_m)\big) \quad (13)$$

$$\zeta(\alpha,\beta,u_m,-u_m) = f_D(\tau,u_m,-u_m)\,\rho_m^{|u_m|}\big(U_{\alpha}(u_m)\big) \times \rho_m^{|u_m|}\big(U_{\beta}(-u_m)\big) \quad (14)$$

$$\zeta(\alpha,0,u_m,0) = f_D(\tau,u_m,0)\,\rho_m^{|u_m|}\big(U_{\alpha}(u_m)\big) \quad (15)$$

where $f_D(\tau, u_m, v_m)$ is defined as

$$f_D(\tau, u_m, v_m) \triangleq \begin{cases} \left[\frac{(\tau-|u_m|)^+}{\tau}\right]^2 & \text{if } m \leq Q-1 \\ & \text{and } v_m = u_m \text{ or } v_m = 0, \\ \left[\frac{(\tau-|u_m|-|v_m|)^+}{\tau}\right]^2 & \text{if } m \leq Q-1 \\ & \text{and } v_m = -u_m, \\ \frac{(\tau-|u_Q|)^+}{\tau} & \text{if } m = Q \\ & \text{and } v_Q = u_Q \text{ or } v_Q = 0, \\ \frac{(\tau-|u_Q|-|v_Q|)^+}{\tau} & \text{if } m = Q \\ & \text{and } v_Q = -u_Q, \end{cases} \quad (16)$$

in which the function $(x)^+ \triangleq \max(x, 0)$ is introduced; and $\rho_m(\alpha)$, $m = 1, \ldots, Q$, is defined as

$$\rho_m(\alpha) \triangleq \int_{\Omega'} \frac{p^{\alpha}_{\boldsymbol{\eta}_m}(\boldsymbol{x})}{p^{\alpha-1}_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} \quad (17)$$

in which Ω' denotes the observation space in one single time instant, i.e., $\prod_{i=1}^{N} \Omega' = \Omega$ (Cartesian product).

As already mentioned, the choices of u_m and v_n are left to the user. Nevertheless, note that these choices must be made in an admissible range so that the matrix $\boldsymbol{G}$ is invertible. Specifically, we deduce from (15) and (16) that values satisfying $|u_m| \geq \tau$ or $|v_n| \geq \tau$ vanish the denominator of (10). This is incompatible with the invertibility of $\boldsymbol{G}$ and these values for the test-points are consequently dismissed. As a consequence, in the following, we assume that $\max(|u_m|, |v_n|) \leq \tau - 1$, $\forall m, n \in \{1, \ldots, Q\}$.

2) Case UT ($n > m+1$): In this case, quite similarly as in the previous one, we obtain

$$\zeta(\alpha, \beta, u_m, v_n) = f_{UT}(\tau, u_m, v_n)\, \rho_m^{|u_m|}\big(U_\alpha(u_m)\big) \rho_n^{|v_n|}\big(U_\beta(v_n)\big) \quad (18)$$

where $f_{UT}(\tau, u_m, v_n)$ is defined as

$$f_{UT}(\tau, u_m, v_n) = \begin{cases} \frac{\big((\tau-|u_m|)^+\big)^2\big((\tau-|v_n|)^+\big)^2}{\tau^4} & \text{if } n \leq Q-1 \\ \frac{\big((\tau-|u_m|)^+\big)^2(\tau-|v_n|)^+}{\tau^3} & \text{if } n = Q. \end{cases} \quad (19)$$

Note that if $\max(|u_m|, |v_n|) > \tau - 1$, $f_{UT}(\tau, u_m, v_n) = 0$, as well as that f_{UT} does not depend on the signs of u_m nor v_n.

3) Case FSD ($n = m+1$): As reflected in Appendix, this case is more complicated than the previous two, because of what we name the possible "*overlap between the test-points*". As explained in Appendix, this situation occurs only if $u_m > 0$ and $v_n < 0$ and if, in that case, $\min(|u_m|, |v_{m+1}|) \geq 2$. Let us first handle the case where there is *no* overlap.

a) Case without overlap, i.e., $[u_m > 0$ and $v_{m+1} > 0]$, or $[u_m < 0$ and $v_{m+1} < 0]$, or $[u_m < 0$ and $v_{m+1} > 0]$: This situation is similar to the previous ones, and we obtain

$$\zeta(\alpha, \beta, u_m, v_{m+1}) = f_{FSD}(\tau, u_m, v_{m+1})\, \rho_m^{|u_m|}\big(U_\alpha(u_m)\big) \rho_{m+1}^{|v_{m+1}|}\big(U_\beta(v_{m+1})\big) \quad (20)$$

where $f_{FSD}(\tau, u_m, v_{m+1})$ is defined as

$$f_{FSD}(\tau, u_m, v_{m+1}) \triangleq \begin{cases} \frac{(\tau-|u_m|)\big(\tau-(-u_m,v_{m+1})^+-(u_m,-v_{m+1})^+\big)^+(\tau-|v_{m+1}|)}{\tau^3} \\ \qquad \text{if } m+1 < Q \\ \frac{(\tau-|u_{Q-1}|)\big(\tau-(-u_{Q-1},v_Q)^+-(u_{Q-1},-v_Q)^+\big)^+}{\tau^2} \\ \qquad \text{if } m+1 = Q. \end{cases} \quad (21)$$

One can notice in particular that $f_{FSD}(\tau, u_m, v_{m+1}) \neq 0$ if $(-u_m, v_{m+1})^+ + (u_m, -v_{m+1})^+ \leq \tau - 1$ (which implies that $\max(|u_m|, |v_{m+1}|) \leq \tau - 1$).

b) Case with possible overlap, i.e., $[u_m > 0$ and $v_{m+1} < 0]$: Here, we obtain:

$$\zeta(\alpha, \beta, u_m, v_{m+1}) = f_{\mathrm{ov}}(\tau, \alpha, \beta, u_m, v_{m+1})\, \rho_m^{|u_m|}(\alpha)\, \rho_{m+1}^{|v_{m+1}|}(1-\beta) \quad (22)$$

where $f_{\mathrm{ov}}(\tau, \alpha, \beta, u_m, v_{m+1})$ is defined as

$$f_{\mathrm{ov}}(\tau, \alpha, \beta, u_m, v_{m+1}) \triangleq \begin{cases} \frac{(\tau-|u_m|)(\tau-|v_{m+1}|)}{\tau^3}\Bigg[(\tau-|u_m|-|v_{m+1}|+1)^+ \\ \qquad -\frac{1-\big(R_m(\alpha,\beta)\big)^{1-(|u_m|,|v_{m+1}|)^-}}{1-R_m(\alpha,\beta)}\Bigg] \quad \text{if } m+1 < Q \\ \frac{(\tau-|u_{Q-1}|)}{\tau^2}\Bigg[(\tau-|u_{Q-1}|-|v_Q|+1)^+ \\ \qquad -\frac{1-\big(R_{Q-1}(\alpha,\beta)\big)^{1-(|u_{Q-1}|,|v_Q|)^-}}{1-R_{Q-1}(\alpha,\beta)}\Bigg] \quad \text{if } m+1 = Q, \end{cases} \quad (23)$$

in which

$$R_m(\alpha, \beta) \triangleq \frac{\rho_m(\alpha)\, \rho_{m+1}(1-\beta)}{\kappa_m(\alpha, 1-\alpha-\beta)}, \quad (24)$$

with function $\kappa_m(\alpha_1, \alpha_2)$ defined, for $m \in \{1, \ldots, Q-1\}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$, by

$$\kappa_m(\alpha_1, \alpha_2) \triangleq \int_{\Omega'} p^{\alpha_1}_{\boldsymbol{\eta}_m}(\boldsymbol{x})\, p^{\alpha_2}_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x})\, p^{1-\alpha_1-\alpha_2}_{\boldsymbol{\eta}_{m+2}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \quad (25)$$

and the function $(x, y)^- \triangleq \min(x, y, 0)$ is used in (23).

It is worth noticing that, if $\min(|u_m|, |v_{m+1}|) = 1$, we have $f_{\mathrm{ov}}(\tau, \alpha, \beta, u_m, v_{m+1}) = f_{FSD}(\tau, u_m, v_{m+1})$, then (22) remains valid whether $\min(|u_m|, |v_{m+1}|) \geq 2$, or $\min(|u_m|, |v_{m+1}|) = 1$, i.e., whether there is indeed overlap or not.

Since we gave closed-form expressions of ζ in all possible cases (as given by (13), (14), (15), (18), (20) and (22)), we are now able to deduce a closed-form expression for matrix $\boldsymbol{G}$, as presented in the next section.

B. Expression of matrix $\boldsymbol{G}$

First, we will give the expression of the diagonal terms of $\boldsymbol{G}$ (case $n = m$), then the first superdiagonal terms (case $n = m+1$), and we will finish with the upper triangle terms (case $n > m+1$).

1) Diagonal terms (case $n = m$): By plugging (13)–(15) into (10), and by using the fact that $U_{1-s_m}(-h_m) = U_{s_m}(h_m)$, we obtain the following closed-form of the diagonal of $\boldsymbol{G}$:

$$[\boldsymbol{G}]_{m,m} = \frac{\begin{pmatrix} f_D(\tau, h_m, h_m) \\ \times \left[\rho_m^{|h_m|}\big(U_{2s_m}(h_m)\big) + \rho_m^{|h_m|}\big(U_{2s_m-1}(h_m)\big)\right] \\ -2 f_D(\tau, h_m, -h_m)\, \rho_m^{2|h_m|}\big(U_{s_m}(h_m)\big) \end{pmatrix}}{f_D^2(\tau, h_m, 0)\, \rho_m^{2|h_m|}\big(U_{s_m}(h_m)\big)}. \quad (26)$$

Notice that in the special case of one single change-point, i.e., $Q = m = 1$, with $s_1 = 1/2$ and $\tau = N-1$, by plugging (26) into (8), we find:

$$\begin{aligned} \mathrm{WWB} &= \sup_h \frac{h^2}{[\boldsymbol{G}]_{1,1}\left(\frac{1}{2}, h\right)} \\ &= \sup_h \frac{h^2\left(1 - \frac{|h|}{N-1}\right)^2 \rho_1^{2|h|}\left(\frac{1}{2}\right)}{2\left(1 - \frac{|h|}{N-1} - \left(1 - \frac{2|h|}{N-1}\right)\rho_1^{2|h|}\left(\frac{1}{2}\right)\right)} \end{aligned} \quad (27)$$

which is our proposed bound in [21].

2) Upper triangle terms (case $n > m+1$): We notice from (18) that, for this case $n > m+1$

$$\begin{aligned} \zeta(s_m, s_n, h_m, h_n) &= \zeta(1-s_m, 1-s_n, -h_m, -h_n) \\ &= \zeta(s_m, 1-s_n, h_m, -h_n) \\ &= \zeta(1-s_m, s_n, -h_m, h_n) \end{aligned} \quad (28)$$

which is obtained by using the fact that $U_{1-s_m}(-h_m) = U_{s_m}(h_m)$. This directly implies that

$$[\boldsymbol{G}]_{m,n} = 0. \quad (29)$$

3) First superdiagonal terms (case $n = m+1$): In order to avoid a tedious sign analysis of h_m and h_{m+1}, let us rewrite the numerator of (10) as

$$\begin{aligned} &\mathrm{num}\left([\boldsymbol{G}]_{m,m+1}\right) \\ &= \mathrm{sign}(h_m h_{m+1}) \\ &\quad \times \Big[\zeta\big(U_{s_m}(h_m), U_{s_{m+1}}(h_{m+1}), |h_m|, |h_{m+1}|\big) \\ &\qquad + \zeta\big(U_{s_m}(-h_m), U_{s_{m+1}}(-h_{m+1}), -|h_m|, -|h_{m+1}|\big) \\ &\qquad - \zeta\big(U_{s_m}(h_m), U_{s_{m+1}}(-h_{m+1}), |h_m|, -|h_{m+1}|\big) \\ &\qquad - \zeta\big(U_{s_m}(-h_m), U_{s_{m+1}}(h_{m+1}), -|h_m|, |h_{m+1}|\big)\Big] \end{aligned} \quad (30)$$

where, for any real x, $\mathrm{sign}(x) = 1$ if $x > 0$ and $\mathrm{sign}(x) = -1$ if $x < 0$. Since $|h_m| > 0 > -|h_{m+1}|$, according to results from Section C3b, the third term $\zeta\left(U_{s_m}(h_m), U_{s_{m+1}}(-h_{m+1}), |h_m|, -|h_{m+1}|\right)$ in (30) has to be written according to (22), whereas the remaining three terms are given by (20).

Then, by appropriately plugging (20) and (22) into (30), we find:

$$\begin{aligned} &\mathrm{num}\left([\boldsymbol{G}]_{m,m+1}\right) = \\ &\quad \mathrm{sign}(h_m h_{m+1})\, g_{\mathrm{ov}}(\tau, s_m, s_{m+1}, h_m, h_{m+1}) \\ &\quad \times \rho_m^{|h_m|}\big(U_{s_m}(h_m)\big)\rho_{m+1}^{|h_{m+1}|}\big(U_{s_{m+1}}(h_{m+1})\big) \end{aligned} \quad (31)$$

where $g_{\mathrm{ov}}(\tau, h_m, h_{m+1})$ is defined, for $m+1 < Q$, by

$$\begin{aligned} g_{\mathrm{ov}}(\tau, h_m, h_{m+1}) &\triangleq f_{FSD}(\tau, |h_m|, |h_{m+1}|) \\ &\quad + f_{FSD}(\tau, -|h_m|, -|h_{m+1}|) \\ &\quad - f_{\mathrm{ov}}(\tau, s_m, s_{m+1}, |h_m|, -|h_{m+1}|) \\ &\quad - f_{FSD}(\tau, -|h_m|, |h_{m+1}|) \\ &= \frac{(\tau - |h_m|)(\tau - |h_{m+1}|)}{\tau^3}\Bigg[2(\tau - |h_m| - |h_{m+1}|)^+ \\ &\quad - (\tau - |h_m| - |h_{m+1}| + 1)^+ - \Big(\tau - (|h_m|, |h_{m+1}|)^+\Big)^+ \\ &\quad + \frac{1 - \left(R_m\left(U_{s_m}(h_m), U_{s_{m+1}}(-h_{m+1})\right)\right)^{1-(|h_m|, |h_{m+1}|)^-}}{1 - R_m\left(U_{s_m}(h_m), U_{s_{m+1}}(-h_{m+1})\right)}\Bigg] \end{aligned} \quad (32)$$

and for $m+1 = Q$, $g_{\mathrm{ov}}(\tau, h_{Q-1}, h_Q)$ is obtained by replacing the term $\frac{(\tau - |h_m|)(\tau - |h_{m+1}|)}{\tau^3}$ in (32) with $\frac{\tau - |h_{Q-1}|}{\tau^2}$. The denominator of $[\boldsymbol{G}]_{m,m+1}$ is obtained by plugging (15) twice into (10), i.e., once with $\alpha = s_m$ and $u_m = h_m$, and once with $\alpha = s_{m+1}$ and $u_m = h_{m+1}$. We finally obtain, for any $m \in \{1, \ldots, Q-1\}$

$$\begin{aligned} [\boldsymbol{G}]_{m,m+1} &= \frac{\tau\, \mathrm{sign}(h_m h_{m+1})}{(\tau - |h_m|)(\tau - |h_{m+1}|)}\Bigg[2(\tau - |h_m| - |h_{m+1}|)^+ \\ &\quad - (\tau - |h_m| - |h_{m+1}| + 1)^+ - \Big(\tau - (|h_m|, |h_{m+1}|)^+\Big)^+ \\ &\quad + \frac{1 - \left(R_m\left(U_{s_m}(h_m), U_{s_{m+1}}(-h_{m+1})\right)\right)^{1-(|h_m|, |h_{m+1}|)^-}}{1 - R_m\left(U_{s_m}(h_m), U_{s_{m+1}}(-h_{m+1})\right)}\Bigg]. \end{aligned} \quad (33)$$

Notice (33) is valid also if $|h_m| = |h_{m+1}| = 1$, as well as if $m+1 = Q$.

We conclude, due to (26), (33) and (29), that matrix $\boldsymbol{G}$ is tridiagonal, i.e.,

$$\boldsymbol{G} = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 \\ B_1 & A_2 & B_2 & \ddots & \vdots \\ 0 & B_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & B_{Q-1} \\ 0 & \cdots & 0 & B_{Q-1} & A_Q \end{bmatrix} \quad (34)$$

where, for any $m \in \{1, \ldots, Q\}$, A_m is given by the right hand side of (26), and for $m \in \{1, \ldots, Q-1\}$, B_m is given by the right hand side of (33).

Note that the same kind of result was obtained in [18], where the Barankin information matrix plays the same role as $\boldsymbol{G}$ in this paper. As in [18], the fact that there exists some m such that $B_m \neq 0$ confirms that the joint estimation of Q change-points is not equivalent to estimating one single change-point Q times. However, a noticeable difference with results from [18] is that B_m generally never equals zero, consequently matrix $\boldsymbol{G}$ is not block diagonal. Even if its inversion is not as simple, it can still be computed efficiently [34].

These results have been obtained whatever the distribution of the observations. In the next section, we apply these

results to the changes in Gaussian and Poisson distributed observations. Note that from equations (26) and (33), for a given distribution of the observations, the only terms that need to be calculated are $\rho_m(\alpha)$ as defined in (17), and $\kappa_m(\alpha,\beta)$ as defined in (25), whereas $R_m(\alpha,\beta)$ is deduced from them, using (24).

V. Cases of Gaussian and Poisson distributions

Gaussian and Poisson distributions both are widely encountered in signal processing applications, see e.g., [2], [35], [36]. For this reason, this section is dedicated to the derivation of the terms of matrix $\boldsymbol{G}$ in these two cases.

A. Gaussian Case

In this section, we assume the distribution of the observations is piecewise Gaussian i.i.d.: $\boldsymbol{x}_i \sim \mathcal{N}\left(\boldsymbol{\mu}_q, \sigma_q^2 \boldsymbol{I}\right)$ for $i = t_{q-1}+1,\ldots,t_q$, with $q = 1,\ldots,Q+1$ (we recall that, by definition, $t_0 = 0$ and $t_{Q+1} = N$). It is then possible to derive $\rho_m(\alpha)$ for $1 \le m \le Q$ and $0 < \alpha < 1$, using (17). In the case where both mean and variance are likely to change, after some manipulations we obtain

$$\rho_m(\alpha) = \left(\frac{\left(\mathrm{snr}^{\mathrm{v}}_{m,m+1}\right)^{\alpha}}{\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+1} + 1 - \alpha}\right)^{\frac{M}{2}} \times \exp\left\{-\frac{\alpha(1-\alpha)\,\mathrm{snr}^{\mathrm{m}}_{m,m+1}}{2\left(\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+1} + 1 - \alpha\right)}\right\} \tag{35}$$

where the following signal-to-noise ratios (SNR) have been defined:

$$\mathrm{snr}^{\mathrm{m}}_{m,n} = \frac{\left\|\boldsymbol{\mu}_n - \boldsymbol{\mu}_m\right\|^2}{\sigma_m^2} \quad \text{and} \quad \mathrm{snr}^{\mathrm{v}}_{m,n} = \frac{\sigma_n^2}{\sigma_m^2} \tag{36}$$

for $m,n = 1,\ldots,Q$. The adaptation to the expression of $\rho_m\big(U_\alpha(h_m)\big)$ is straightforward. If the changes occur either in the mean only (i.e., $\mathrm{snr}^{\mathrm{v}}_{m,m+1} = 1$, $\forall m$), or in the variance only (i.e., $\mathrm{snr}^{\mathrm{m}}_{m,m+1} = 0$, $\forall m$), the expression of $\rho_m(\alpha)$ can be simplified accordingly:

$$\rho_m(\alpha) = \exp\left\{-\frac{\alpha(1-\alpha)}{2}\mathrm{snr}^{\mathrm{m}}_{m,m+1}\right\} \tag{37}$$

if the variance is constant, and

$$\rho_m(\alpha) = \left(\frac{\left(\mathrm{snr}^{\mathrm{v}}_{m,m+1}\right)^{\alpha}}{\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+1} + 1 - \alpha}\right)^{\frac{M}{2}} \tag{38}$$

if the mean is constant.

Regarding $\kappa_m(\alpha,\beta)$ defined in (25), the same kind of manipulations leads to

$$\kappa_m(\alpha,\beta) = \left[\frac{\left(\mathrm{snr}^{\mathrm{v}}_{m,m+2}\right)^{\alpha}\left(\mathrm{snr}^{\mathrm{v}}_{m+1,m+2}\right)^{\beta}}{\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+2} + \beta \mathrm{snr}^{\mathrm{v}}_{m+1,m+2} + 1 - \alpha - \beta}\right]^{\frac{M}{2}} \times \exp\left\{-\frac{\begin{array}{c}\alpha\beta \mathrm{snr}^{\mathrm{v}}_{m+1,m+2}\mathrm{snr}^{\mathrm{m}}_{m,m+1} \\ + \alpha(1-\alpha-\beta)\,\mathrm{snr}^{\mathrm{m}}_{m,m+2} \\ + \beta(1-\alpha-\beta)\,\mathrm{snr}^{\mathrm{m}}_{m+1,m+2}\end{array}}{2(\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+2} + \beta \mathrm{snr}^{\mathrm{v}}_{m+1,m+2} + 1 - \alpha - \beta)}\right\}. \tag{39}$$

Finally, R_m can be calculated by plugging (35) and (39) into (24), and we obtain

$$\begin{aligned} R_m(\alpha,\beta) &= \left[\frac{\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+2} + (1-\alpha-\beta)\,\mathrm{snr}^{\mathrm{v}}_{m+1,m+2} + \beta}{\left(\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+1} + 1 - \alpha\right)\left((1-\beta)\,\mathrm{snr}^{\mathrm{v}}_{m+1,m+2} + \beta\right)}\right]^{\frac{M}{2}} \\ &\times \exp\left\{-\frac{\alpha\beta}{2\left(\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+2} + (1-\alpha-\beta)\,\mathrm{snr}^{\mathrm{v}}_{m+1,m+2} + \beta\right)}\right. \\ &\times \left(\frac{\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+2} + 1 - \alpha}{\alpha \mathrm{snr}^{\mathrm{v}}_{m,m+1} + 1 - \alpha}\mathrm{snr}^{\mathrm{m}}_{m,m+1}\right. \\ &+ \frac{(1-\beta)\,\mathrm{snr}^{\mathrm{v}}_{m,m+2} + \beta}{(1-\beta)\,\mathrm{snr}^{\mathrm{v}}_{m+1,m+2} + \beta}\mathrm{snr}^{\mathrm{m}}_{m+1,m+2} \\ &\left.\left.- \mathrm{snr}^{\mathrm{m}}_{m,m+2}\right)\right\}. \end{aligned} \tag{40}$$

B. Poisson case

Let us now assume that the observations follow a piecewise, i.i.d., Poisson distribution, i.e., $x_i \sim \mathcal{P}(\lambda_q)$ for $i = t_{q-1}+1,\ldots,t_q$ and $q = 1,\ldots,Q+1$. As in the previous section, it is possible to obtain closed-form expressions for $\rho_m(\alpha)$, $\kappa_m(\alpha,\beta)$ and $R_m(\alpha,\beta)$, which we give hereafter. Since the distribution is discrete, one has to replace the integral operator in (17) and (25) with a discrete sum. Then, we have

$$\begin{aligned} \rho_m(\alpha) &= \sum_{k=0}^{+\infty}\frac{\mathrm{Pr}^{\alpha}_{\lambda_m}(x=k)}{\mathrm{Pr}^{\alpha-1}_{\lambda_{m+1}}(x=k)} \\ &= \exp\left\{-\alpha\lambda_m - (1-\alpha)\lambda_{m+1}\right\}\sum_{k=0}^{+\infty}\frac{\left(\lambda_m^{\alpha}\lambda_{m+1}^{1-\alpha}\right)^k}{k!} \\ &= \exp\left\{-\alpha\lambda_m - (1-\alpha)\lambda_{m+1} + \lambda_m^{\alpha}\lambda_{m+1}^{1-\alpha}\right\}. \end{aligned} \tag{41}$$

In a similar way, we obtain the expression of $\kappa_m(\alpha,\beta)$, that is

$$\begin{aligned} \kappa_m(\alpha,\beta) &= \sum_{k=0}^{+\infty}\frac{\mathrm{Pr}^{\alpha}_{\lambda_m}(x=k)\,\mathrm{Pr}^{\beta}_{\lambda_{m+1}}(x=k)}{\mathrm{Pr}^{\alpha+\beta-1}_{\lambda_{m+2}}(x=k)} \\ &= \exp\left\{-\alpha\lambda_m - \beta\lambda_{m+1} - (1-\alpha-\beta)\lambda_{m+2}\right. \\ &\left.+ \frac{\lambda_m^{\alpha}\lambda_{m+1}^{\beta}}{\lambda_{m+2}^{\alpha+\beta-1}}\right\}, \end{aligned} \tag{42}$$

as well as that of $R_m(\alpha,\beta)$:

$$\begin{aligned} R_m(\alpha,\beta) = \exp\Bigg\{-\lambda_{m+1}\Bigg[1 &- \left(\frac{\lambda_m}{\lambda_{m+1}}\right)^{\alpha} - \left(\frac{\lambda_{m+2}}{\lambda_{m+1}}\right)^{\beta} \\ &+ \left(\frac{\lambda_m}{\lambda_{m+1}}\right)^{\alpha}\left(\frac{\lambda_{m+2}}{\lambda_{m+1}}\right)^{\beta}\Bigg]\Bigg\}. \end{aligned} \tag{43}$$

VI. Computation of the tightest WWB

In this section, we aim to present a fast and efficient procedure in order to compute the tightest WWB, w.r.t. the global mean square error, given by

$$\mathbb{E}_{\boldsymbol{X},\boldsymbol{t}}\left\{\left[\hat{\boldsymbol{t}}\left(\boldsymbol{X}\right)-\boldsymbol{t}\right]\left[\hat{\boldsymbol{t}}\left(\boldsymbol{X}\right)-\boldsymbol{t}\right]^T\right\}\succeq \mathbf{WWB} = \sup_{\substack{\boldsymbol{h},\\ s_1,\ldots,s_Q}} \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{H}^T \tag{44}$$

in which $\boldsymbol{h}=[h_1,\ldots,h_Q]^T$. It is worth mentioning that, we notice by extensive simulations, see Section VII, that $s_q=\frac{1}{2}$ for $q=1,\ldots,Q$ seems to be the optimal value $\forall h_1,\ldots,h_Q\in\mathcal{A}$, in which the set of the possible test points $\mathcal{A}$, satisfies $\boldsymbol{G}$ to be a positive definite matrix. Consequently, in the remaining of this section, we focus on the computation of the supremum of the finite set $\mathcal{W}=\{\mathbf{WWB}(\boldsymbol{h})\mid(h_1,\ldots,h_Q)\in\mathcal{A}\}$ w.r.t. Loewner partial ordering [32] by using a similar methodology to [18], [37], in which, for sake of clarity the WWB matrices are indexed by the vector containing the test point $\boldsymbol{h}$ for fixed $s_q=\frac{1}{2}$ for $q=1,\ldots,Q$.

First, let us recall some useful definitions: The greatest element $\mathbf{WWB}(\boldsymbol{h}^*)$ w.r.t. $(\mathcal{S},\preceq)$ of $\mathcal{W}$, if it exists, is defined as the element of $\mathcal{W}$ satisfying $\mathbf{WWB}(\boldsymbol{h})\preceq\mathbf{WWB}(\boldsymbol{h}^*)$, $\forall\boldsymbol{h}\in\mathcal{A}$ in which $\mathcal{S}$ denotes the set of square matrices of size $Q\times Q$ and $\boldsymbol{A}\succeq\boldsymbol{B}$ means that $\boldsymbol{A}-\boldsymbol{B}$ is a non-negative definite matrix. Whereas, the supremum is a minimal-upper bound on $\mathcal{W}$ that could be not contained in the set $\mathcal{W}$. Meaning that, the greatest element may not exist (in the case of the presence of at least two non comparable matrices w.r.t. Loewner partial ordering and belonging to $\mathcal{W}$), but if it exists, thus, it is also the supremum. Finally, a maximal element, w.r.t. $(\mathcal{S},\preceq)$, $\mathbf{WWB}(\boldsymbol{h}')$, of $\mathcal{W}$ insures that there is no element in $\mathcal{W}$ such that $\mathbf{WWB}(\boldsymbol{h}')\preceq\mathbf{WWB}(\boldsymbol{h}'')$.

Second, let us recall *Lemma 3* of [18], which sates that for any two positive definite matrices $\boldsymbol{A}$ and $\boldsymbol{A}'$, $\boldsymbol{A}\succeq\boldsymbol{A}'$ if and only if $\varepsilon(\boldsymbol{A}')\subseteq\varepsilon(\boldsymbol{A})$ in which the hyper-ellipsoid $\varepsilon(\boldsymbol{A})$ is defined as $\varepsilon(\boldsymbol{A})=\left\{\boldsymbol{y}\mid\boldsymbol{y}^T\boldsymbol{A}^{-1}\boldsymbol{y}\leq 1\right\}$. Consequently, $\mathbf{WWB}(\boldsymbol{h}^*)$ is given as the generating matrix of the minimum volume hyper-ellipsoid containing the set constituted of the hyper-ellipsoids generated by all matrices belonging to $\mathcal{W}$. This means that, the computation of the supremum is done by finding $\varepsilon(\mathbf{WWB}(\boldsymbol{h}^*))$ which contains $\mathcal{E}_{\mathcal{W}}=\{\varepsilon(\mathbf{WWB}(\boldsymbol{h}))\mid\boldsymbol{h}\in\mathcal{A}\}$. Furthermore, for a lower computational cost, $\mathcal{E}_{\mathcal{W}}$ can be reduced to Q' elements stacked in the set $\mathcal{E}'_{\mathcal{W}}$ in which only the maximal elements of $\mathcal{E}_{\mathcal{W}}$ are retained (this can be done by a simple Matlab routine).

Finally, the computation of the minimum volume ellipsoid covering an union of ellipsoids can be solved efficiently using the following convex optimization problem (see page 411 of [37], for more details)

$$\begin{aligned}&\text{minimize} && \log\left(\det\left(\mathbf{WWB}(\boldsymbol{h})^{\frac{1}{2}}\right)\right) \\ &\text{subject to} && b_1\geq 0, b_2\geq 0,\ldots,b_m\geq 0,\\ &&& \begin{bmatrix}\mathbf{WWB}(\boldsymbol{h})^{-1}-b_i\mathbf{WWB}(\boldsymbol{h}_i)^{-1} & \mathbf{0}\\ \mathbf{0} & b_i-1\end{bmatrix}\preceq 0\\ &&& (i=1,\ldots,Q'),\end{aligned} \tag{45}$$

in which $\mathbf{WWB}(\boldsymbol{h}_i)\in\mathcal{E}'_{\mathcal{W}}$.

VII. Numerical Results

In order to illustrate the previous theoretical results, we now apply them to simulated data. First of all, we present the conditions of these simulations. We then investigate the three cases introduced in the previous section, namely changes in i) the mean of a Gaussian time series with constant variance, ii) the variance of a Gaussian time series with constant mean and iii) the parameter of a Poisson distribution. These are the cases treated in [18], however the approach is different since we assume the changes locations are random variables, as explained in section II. Consequently, it is important to keep in mind that the WWB holds on the *global* mean square error (GMSE). The presented results thus essentially consist in the comparison between the GMSE of the Maximum *A Posteriori* (MAP) estimator and the proposed bound.

A. MAP estimator and simulation parameters

By definition, the MAP estimator is given by

$$\hat{\boldsymbol{t}}_{\mathrm{MAP}}=\arg\max_{\boldsymbol{k}}\ln\Pr\left(\boldsymbol{t}=\boldsymbol{k}|\boldsymbol{X}\right) \tag{46}$$

which reduces to, using the fact that $\Pr\left(\boldsymbol{t}=\boldsymbol{k}\right)$ does not depend on $\boldsymbol{k}$ (see (5)) and the Bayes' rule:

$$\begin{aligned}\hat{\boldsymbol{t}}_{\mathrm{MAP}}&=\arg\max_{\boldsymbol{k}}\ln p\left(\boldsymbol{X}|\boldsymbol{t}=\boldsymbol{k}\right)\\ &=\arg\max_{\boldsymbol{k}}\sum_{q=1}^{Q+1}\ln p_{\boldsymbol{\eta}_q}\left(\boldsymbol{x}_{k_{q-1}+1},\ldots,\boldsymbol{x}_{k_q}\right)\end{aligned} \tag{47}$$

where $k_0=t_0=0$ and $k_{Q+1}=t_{Q+1}=N$. Its computation is done numerically.

The following simulations were conducted with $M=1$ (i.e., single time series), $N=80$ and a number of changes $Q=2$ or 3. We plot the average GMSE of the MAP computed over 1000 Monte-Carlo experiments. It is traced in function of SNRs, whose definitions are recalled hereafter for each case. The prior law parameter was fixed at its maximum value, i.e., $\tau=\left\lfloor\frac{N-1}{Q}\right\rfloor$, given that the number of changes Q is assumed to be known (as we stated in (3)).

B. Mean changes of a Gaussian distribution

As stated in (8), the WWB is obtained by solving (45) using CVX package [38] w.r.t. h_1,h_2,h_3, and for different values of s_1,s_2,s_3 if necessary. We took the test-points in their maximum acceptable range, that is $|h_q|\leq\tau-1$ for all q.

We consider the first particular case described in section V-A, namely multiple changes in the mean of a Gaussian distribution with constant variance. The corresponding matrix $\boldsymbol{G}$ from (8) is given by (26) and (33), in which one has to plug (37) and (40) with $\mathrm{snr}^{\mathrm{v}}_{m,n}=1$, $\forall m,n$. In our example, we chose $\sigma^2=1$, $\mu_1=1$, and μ_2, μ_3 and μ_4 are set in such a way that $\mathrm{snr}^{\mathrm{m}}_{1,2}=\mathrm{snr}^{\mathrm{m}}_{2,3}=\mathrm{snr}^{\mathrm{m}}_{3,4}=\mathrm{snr}$. More precisely, we set $\mu_{m+1}=\mu_m+(-1)^m\sqrt{\sigma^2\mathrm{snr}}$, $m=1,2,3$.

The influence of parameters s_q is first studied on a reduced simulation with $Q=2$ and for a large number of couple values $(s_1,s_2)\in\left]0,1\right[^2$. Our simulation results reveal that $s_1=s_2=$

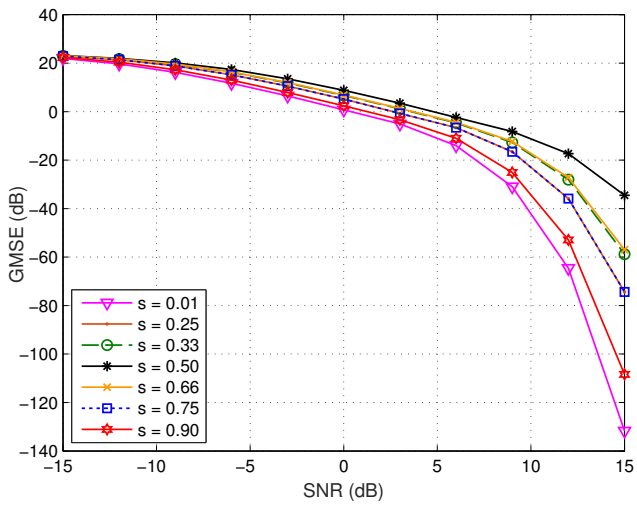

Fig. 1. Plot of $\mathrm{trace}(\mathbf{WWB})$ in function of the SNR for various values of parameter s_q in the particular case $s_1 = s_2$, for $Q = 2$ changes in the mean of a Gaussian distribution.

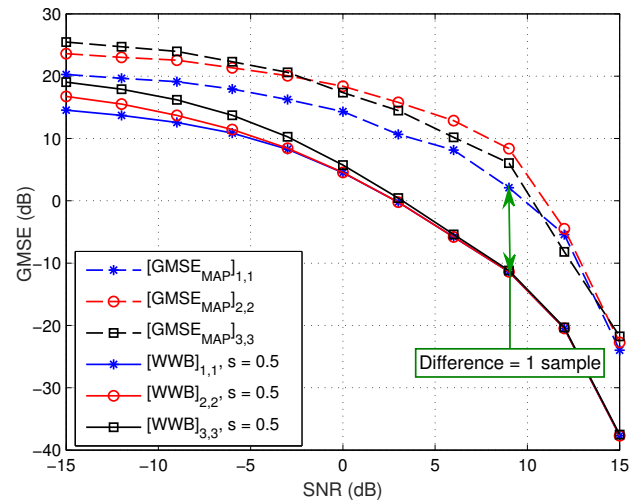

Fig. 2. WWBs and estimated GMSEs w.r.t. each change-point versus SNR, for $Q = 3$ changes in the mean of a Gaussian distribution. Green arrow: gap between the GMSE and the bound in terms of root mean square error.

0.5 seems to be the optimal value. For sake of clarity, we present in Fig. 1 only a sample of this values, namely, $s_1 = s_2$ lying in the set $\{0.01, 0.25, 0.33, 0.5, 0.66, 0.75, 0.90\}$.

In Fig. 2, we illustrate the global mean square error and the associated WWB for each change-point t_1, t_2, t_3, as a function of snr, by plotting the diagonal terms of the matrices $\mathbf{GMSE}$ and $\mathbf{WWB}$, respectively. One can note that, even though the gap between the GMSE and the bound increases for SNRs between -20 dB and 10 dB, it decreases for SNR higher than 10 dB to less than 10 dB of difference. In addition, it is worth mentioning that a gap of 13 dB between the estimated GMSE and the bound, as it is the case at $\mathrm{snr} = 9$ dB regarding to t_1, corresponds to a difference of 1 sample in terms of the *root* mean square error, which is reasonable. Such an asymptotic gap can be explained by the fact that the estimated parameters are *discrete*, and in this context the classical theorems of convergence, consistence and efficiency of estimators like the MLE are no longer valid: one can refer to [39] for more details on discrete parameters estimation. One can also remark that, at $\mathrm{snr} = -15$ dB, both the GMSE and the WWB are higher regarding to t_1 than they are regarding to t_3. This is somewhat intuitive since the error made at estimating the first change is likely to accumulate with that made when estimating the subsequent ones. The derived bound consequently reflects this behavior. The study is restricted to a SNR range under 15 dB, since for higher SNR, the bound and the GMSE drastically tend to zero. This is partly due to the integer nature of the parameters: at high SNR, the average GMSE over all the Monte-Carlo simulations is more and more likely to equal zero, then its logarithm tends to $-\infty$, consequently the corresponding points cannot be plotted. One way to overcome this problem is to increase the number of Monte-Carlo experiments.

C. Variance changes of a Gaussian distribution

In this section, we study the second particular case described in section V-A, namely multiple changes in the variance of a Gaussian distribution with constant mean. The matrix $\boldsymbol{G}$ in (8) is once again obtained from (26) and (33), in which one plugs (38) and (40) with $\mathrm{snr}^{\mathrm{m}}_{m,n} = 0$, $\forall m, n$. In this example, we chose $\mu = 0$, $\sigma_1^2 = 1$, and σ_2^2, σ_3^2 and σ_4^2 are set such that $\sigma_{m+1}^2 = \sigma_m^2 \mathrm{snr}$, $m = 1, 2, 3$, where $\mathrm{snr} = \mathrm{snr}^{\mathrm{v}}_{1,2} = \mathrm{snr}^{\mathrm{v}}_{2,3} = \mathrm{snr}^{\mathrm{v}}_{3,4}$. Fig. 3 illustrates the behavior of the GMSE and the WWB for each change-point t_1, t_2, t_3 as a function of snr, which is approximately the same as in the previous case, except the gap between them seems slightly smaller.

D. Mean Rate changes of a Poisson distribution

In the same way as in the previous examples, we study the case of three change-points in the mean rate of a Poisson distribution, as formulated in section V-B. One here uses equations (41) and (43) to obtain (26) and (33), which enables us to compute the WWB as stated in (8). In this context, we define the SNR for the mth change-point as $\mathrm{snr}^{\lambda}_{m,m+1} = (\lambda_{m+1} - \lambda_m)^2 / (\lambda_m)^2$. In practice, we set $\lambda_1 = 1$ and the followings are set in such way that $\mathrm{snr}^{\lambda}_{1,2} = \mathrm{snr}^{\lambda}_{2,3} = \mathrm{snr}^{\lambda}_{3,4} = \mathrm{snr}$. More precisely, $\lambda_{m+1} = \lambda_m (1 + \sqrt{\mathrm{snr}})$. In Fig. 4, we illustrate the behavior of the WWB compared to the GMSE performance of the MAP estimator. The main difference with the previous cases is that neither the WWB nor the GMSE performance for each change-point tend to be equivalent at high SNR: there exists a cross point ($\mathrm{snr} = -6$ dB) below which the behavior is similar to the Gaussian case (i.e., the GMSE – and the WWB – is higher for late changes than for the early ones) and above which the errors relative order is inverted, i.e., the error relative to late changes becomes smaller than that relative to early changes. This different behavior can be explained by two reasons:

1) The fact that a change in λ affects both the mean and the variance of the distribution. In the Poisson case, only one parameter (the mean rate λ_m) is likely to change from one segment to another, but unlike in the Gaussian case, it drives *both* mean and variance. Yet, it can intuitively be understood that a composite change that affects both mean and variance is

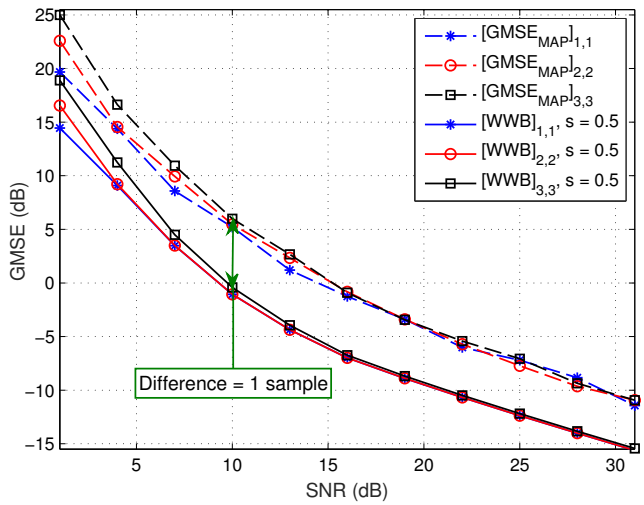

Fig. 3. WWBs and estimated GMSEs w.r.t. each change-point versus SNR, for $Q = 3$ changes in the variance of a Gaussian distribution. Green arrow: gap between the GMSE and the bound in terms of root mean square error.

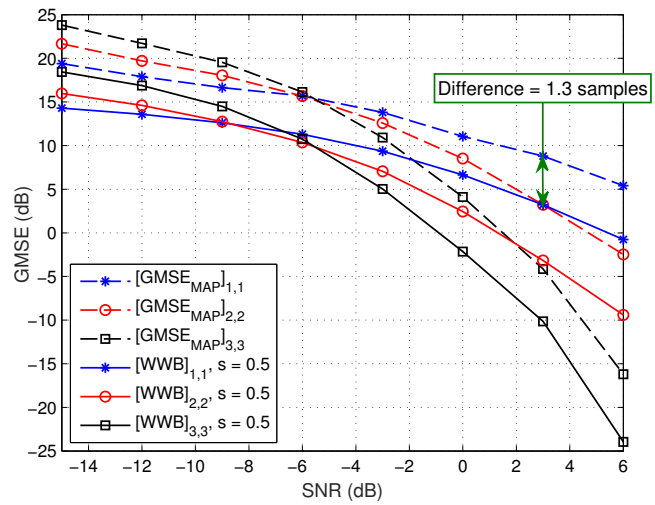

Fig. 4. WWBs and estimated GMSEs w.r.t. each change-point versus SNR, for $Q = 3$ changes in the mean rate of a Poisson distribution. Green arrow: gap between the GMSE and the bound in terms of root mean square error.

more easily estimated than a simple change in the mean only, for example (a mathematical justification of this point can be found in [1]).

2) The definition of the SNR used for this simulation. Indeed, the relation $\lambda_{m+1} = \lambda_m \left(1 + \sqrt{\text{snr}}\right)$ defines an increasing geometric progression $(\lambda_m)_{m \geq 1}$, which makes the difference between the means of consecutive segments higher for late changes (for instance for the last one t_3 between segments $[t_2 + 1, t_3]$ and $[t_3 + 1, N]$) than for early changes (for example the one before t_2 between segments $[t_1 + 1, t_2]$ and $[t_2 + 1, t_3]$).

VIII. Conclusion

In this paper, we analyzed the Weiss-Weinstein bound (WWB) for the estimation of multiple change-points whose total number is assumed to be known, completing [21]. Particularly, we have given a closed-form expression for the matrix $\boldsymbol{G}$ that is involved in its computation. The tridiagonal structure of this matrix shows that, in this Bayesian context as well, the estimation of consecutive change-points interfere one with another. We then applied the proposed bound to two types of distribution: first in the Gaussian case with changes either in the mean or in the variance, secondly in the Poisson case. Furthermore, a semi-definite problem formulation is given in order to solve efficiently the optimization problem leading to the tightest WWB. Simulations were conducted for each of these examples, which led to similar conclusions: even if there is a slight gap between the WWB and the estimates errors, the proposed bound is in good agreement with the estimation behavior.

Appendix
Technical details on the derivation of $\zeta(\alpha, \beta, u_m, v_n)$

In this appendix, we compute $\zeta\left(\alpha, \beta, u_m, v_n\right)$ in order to obtain $[\boldsymbol{G}]_{m,n}$. Using the conditional probabilities chain rule, we rewrite (11) as:

$$\begin{aligned}
&\zeta\left(\alpha, \beta, u_m, v_n\right) \\
&= \sum_{\boldsymbol{k} \in \mathcal{I}} \left(\frac{\Pr^{\alpha}\left(\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m\right) \Pr^{\beta}\left(\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{v}_n\right)}{\Pr^{\alpha+\beta-1}\left(\boldsymbol{t} = \boldsymbol{k}\right)} \right. \\
&\qquad \left. \times \int_{\Omega} \frac{p^{\alpha}\left(\boldsymbol{X} | \boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m\right) p^{\beta}\left(\boldsymbol{X} | \boldsymbol{t} = \boldsymbol{k} + \boldsymbol{v}_n\right)}{p^{\alpha+\beta-1}\left(\boldsymbol{X} | \boldsymbol{t} = \boldsymbol{k}\right)} \mathrm{d}\boldsymbol{X} \right) \\
&= \sum_{\boldsymbol{k} \in \mathcal{I}} \left(\pi\left(\alpha, \beta, u_m, v_n, \boldsymbol{k}\right) \int_{\Omega} \ell\left(\boldsymbol{X}, \alpha, \beta, u_m, v_n, \boldsymbol{k}\right) \mathrm{d}\boldsymbol{X} \right) \quad (48)
\end{aligned}$$

in which we define

$$\pi\left(\alpha, \beta, u_m, v_n, \boldsymbol{k}\right) \triangleq \frac{\Pr^{\alpha}(\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m) \Pr^{\beta}(\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{v}_n)}{\Pr^{\alpha+\beta-1}\left(\boldsymbol{t} = \boldsymbol{k}\right)} \quad (49)$$

and

$$\ell\left(\boldsymbol{X}, \alpha, \beta, u_m, v_n, \boldsymbol{k}\right) \triangleq \frac{p^{\alpha}\left(\boldsymbol{X} | \boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m\right) p^{\beta}\left(\boldsymbol{X} | \boldsymbol{t} = \boldsymbol{k} + \boldsymbol{v}_n\right)}{p^{\alpha+\beta-1}\left(\boldsymbol{X} | \boldsymbol{t} = \boldsymbol{k}\right)}. \quad (50)$$

In the following, we first give the expression of $\pi\left(\alpha, \beta, u_m, v_n, \boldsymbol{k}\right)$. Then, we derive $\int_{\Omega} \ell\left(\boldsymbol{X}, \alpha, \beta, u_m, v_n, \boldsymbol{k}\right) \mathrm{d}\boldsymbol{X}$, and we finally deduce a closed-form expression of ζ.

A. Derivation of $\pi\left(\alpha, \beta, u_m, v_n, \boldsymbol{k}\right)$

We directly deduce the expression of the denominator of (49) from (5) as:

$$\Pr^{\alpha+\beta-1}\left(\boldsymbol{t} = \boldsymbol{k}\right) = \begin{cases} \left(\frac{1}{\tau}\right)^{Q(\alpha+\beta-1)} & \text{if } k_q \in \mathcal{J}_q \text{ for all } q \in \{1, \ldots, Q\} \\ 0 & \text{otherwise.} \end{cases} \quad (51)$$

TABLE I
RANGE OF VALUES FOR k_q INDICES IN THE CASE D ($m = n$)

q	$(k_q)_{\min}$	$(k_q)_{\max}$
m	$k_{m-1} + 1 + (-u_m, -v_m)^+$	$k_{m-1} + \tau - (u_m, v_m)^+$
$m+1$	$k_m + 1 + (u_m, v_m)^+$	$k_m + \tau - (-u_m, -v_m)^+$
Other	$k_{q-1} + 1$	$k_{q-1} + \tau$

The two terms on the numerator of (49) are obtained similarly by writing the elements of vectors $\boldsymbol{u}_m$ and $\boldsymbol{v}_n$ as follows:

$$[\boldsymbol{u}_m]_i = u_m \delta_{m,i} \qquad \text{and} \qquad [\boldsymbol{v}_n]_i = v_n \delta_{n,i} \tag{52}$$

where $[\boldsymbol{u}_m]_i$ denotes the i-th element of vector $\boldsymbol{u}_m$ and $\delta_{m,i}$ is the Krönecker delta, i.e., $\delta_{m,i} = 1$ iff $i = m$, otherwise $\delta_{m,i} = 0$. Then, from (51) and (52), we deduce, on the one hand:

$$\Pr{}^{\alpha}(\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m) = \begin{cases} \left(\frac{1}{\tau}\right)^{Q\alpha} & \text{if } k_q \in \mathcal{J}_{q,u_m} \text{ for all } q \in \{1, \ldots, Q\} \\ 0 & \text{otherwise,} \end{cases} \tag{53}$$

where, for any integers $q \in \{1, \ldots, Q\}$ and $u_m \in \mathbb{Z}$, we define the sets of integers $\mathcal{J}_{q,u_m}$ as $\mathcal{J}_{q,u_m} \triangleq \mathcal{J}_q + u_m(\delta_{m,q-1} - \delta_{m,q})$, in which the "+" sign denotes the translation of the set. More explicitly:

$$\begin{cases} \mathcal{J}_{q,u_m} = \mathcal{J}_q = \{k_{q-1} + 1, \ldots, k_{q-1} + \tau\} \text{ if } q \neq m, m+1 \\ \mathcal{J}_{m,u_m} = \{k_{m-1} - u_m + 1, \ldots, k_{m-1} - u_m + \tau\} \\ \mathcal{J}_{m+1,u_m} = \{k_m + u_m + 1, \ldots, k_m + u_m + \tau\}. \end{cases} \tag{54}$$

On the other hand, in the same way, we have:

$$\Pr{}^{\beta}(\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{v}_n) = \begin{cases} \left(\frac{1}{\tau}\right)^{Q\beta} & \text{if } k_q \in \mathcal{J}_{q,v_n} \text{ for all } q \in \{1, \ldots, Q\} \\ 0 & \text{otherwise.} \end{cases} \tag{55}$$

It is necessary that the terms in (51), (53) and (55) are simultaneously nonzero to ensure that $\pi(\alpha, \beta, u_m, v_n, \boldsymbol{k})$ in (49) is nonzero as well. Consequently, for all $q = 1, \ldots, Q$, the summation domain $\mathcal{I}_q$ w.r.t. k_q in (48) is given by:

$$\mathcal{I}_q = \mathcal{J}_{q,0} \cap \mathcal{J}_{q,u_m} \cap \mathcal{J}_{q,v_n}. \tag{56}$$

Tables I, II and III give the smallest element $(k_q)_{\min} \triangleq \min \mathcal{I}_q$ and the greatest element $(k_q)_{\max} \triangleq \max \mathcal{I}_q$ of each set $\mathcal{I}_q$ in the three cases $n = m$ (D), $n > m + 1$ (UT) and $n = m + 1$ (FSD), respectively. In these tables, we also use the functions $(x)^+ \triangleq \max(x, 0)$, defined earlier, and $(x, y)^+ \triangleq \max(x, y, 0)$.

Finally, the complete expression of $\pi(\alpha, \beta, u_m, v_n, \boldsymbol{k})$ is deduced by plugging (51), (53) and (55) into (49):

$$\pi(\alpha, \beta, u_m, v_n, \boldsymbol{k}) = \begin{cases} \left(\frac{1}{\tau}\right)^{Q} & \text{if } k_q \in \mathcal{I}_q \text{ for all } q \in \{1, \ldots, Q\} \\ 0 & \text{otherwise,} \end{cases} \tag{57}$$

where the sets $\mathcal{I}_q$ are given by tables I to III.

TABLE II
RANGE OF VALUES FOR k_q INDICES IN THE CASE UT ($m + 1 < n$)

q	$(k_q)_{\min}$	$(k_q)_{\max}$
m	$k_{m-1} + 1 + (-u_m)^+$	$k_{m-1} + \tau - (u_m)^+$
$m+1$	$k_m + 1 + (u_m)^+$	$k_m + \tau - (-u_m)^+$
n	$k_{n-1} + 1 + (-v_n)^+$	$k_{n-1} + \tau - (v_n)^+$
$n+1$	$k_n + 1 + (v_n)^+$	$k_n + \tau - (-v_n)^+$
Other	$k_{q-1} + 1$	$k_{q-1} + \tau$

TABLE III
RANGE OF VALUES FOR k_q INDICES IN THE CASE FSD ($m + 1 = n$)

q	$(k_q)_{\min}$	$(k_q)_{\max}$
m	$k_{m-1} + 1 + (-u_m)^+$	$k_{m-1} + \tau - (u_m)^+$
$m+1$	$k_m + 1 + (u_m, -v_{m+1})^+$	$k_m + \tau - (-u_m, v_{m+1})^+$
$m+2$	$k_{m+1} + 1 + (v_{m+1})^+$	$k_{m+1} + \tau - (-v_{m+1})^+$
Other	$k_{q-1} + 1$	$k_{q-1} + \tau$

B. Derivation of $\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, v_n, \boldsymbol{k})\, \mathrm{d}\boldsymbol{X}$

Let us now derive the second term of (48). Given the model (1), and particularly using the independency of the observations one from another, the likelihood $p(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k})$ of the observations is given by:

$$p(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k}) = \prod_{q=1}^{Q+1} \prod_{i=k_{q-1}+1}^{k_q} p_{\boldsymbol{\eta}_q}(\boldsymbol{x}_i) \tag{58}$$

and in the same way,

$$p(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m) = \prod_{i=1}^{k_1} p_{\boldsymbol{\eta}_1}(\boldsymbol{x}_i) \prod_{i=k_1+1}^{k_2} p_{\boldsymbol{\eta}_2}(\boldsymbol{x}_i) \ldots \prod_{i=k_{m-1}+1}^{k_m+u_m} p_{\boldsymbol{\eta}_m}(\boldsymbol{x}_i) \times \prod_{i=k_m+u_m+1}^{k_{m+1}} p_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x}_i) \ldots \prod_{i=k_Q+1}^{N} p_{\boldsymbol{\eta}_{Q+1}}(\boldsymbol{x}_i). \tag{59}$$

Of course, the expression of $p(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{v}_n)$ is the same as that in (59) by only replacing m with n and u_m with v_n. As already mentioned in the main body of this paper, since matrix $\boldsymbol{G}$ is symmetric, it is possible to consider the case $m \leq n$ only (diagonal and upper triangle terms). We separate the study of the diagonal elements (case D $m = n$) from that of the upper triangle ones (case UT $m < n$), as done in the next two sections.

1) Case D ($m = n$): In this case, as it can be seen from the numerator of (10), u_m and v_m can either be the same, i.e., $u_m = v_m$, or opposed, i.e., $u_m = -v_m$. Then, depending on the sign of u_m, the study of the expression of $\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, v_m, \boldsymbol{k})\, \mathrm{d}\boldsymbol{X}$ can be split into the following four cases ("D" stands for "diagonal"):

$$\begin{cases} \textit{a) Case D1: } u_m = v_m > 0 \\ \textit{b) Case D2: } u_m = v_m < 0 \\ \textit{c) Case D3: } u_m = -v_m > 0 \\ \textit{d) Case D4: } u_m = -v_m < 0. \end{cases} \tag{60}$$

a) Case D1 ($u_m = v_m > 0$): Using (58) and (59), we rewrite $\ell(\boldsymbol{X}, \alpha, \beta, u_m, u_m, \boldsymbol{k})$ as:

$$\begin{aligned}
&\ell(\boldsymbol{X}, \alpha, \beta, u_m, u_m, \boldsymbol{k}) \\
&= \frac{p^{\alpha+\beta}(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m)}{p^{\alpha+\beta-1}(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k})} \\
&= \frac{\prod_{i=1}^{k_1} p_{\boldsymbol{\eta}_1}^{\alpha+\beta}(\boldsymbol{x}_i) \dots \prod_{i=k_{m-1}+1}^{k_m+u_m} p_{\boldsymbol{\eta}_m}^{\alpha+\beta}(\boldsymbol{x}_i)}{\prod_{i=1}^{k_1} p_{\boldsymbol{\eta}_1}^{\alpha+\beta-1}(\boldsymbol{x}_i) \dots \prod_{i=k_{m-1}+1}^{k_m} p_{\boldsymbol{\eta}_m}^{\alpha+\beta-1}(\boldsymbol{x}_i)} \\
&\quad \times \frac{\prod_{i=k_m+u_m+1}^{k_{m+1}} p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta}(\boldsymbol{x}_i) \dots \prod_{i=k_Q+1}^{N} p_{\boldsymbol{\eta}_{Q+1}}^{\alpha+\beta}(\boldsymbol{x}_i)}{\prod_{i=k_m+1}^{k_{m+1}} p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta-1}(\boldsymbol{x}_i) \dots \prod_{i=k_Q+1}^{N} p_{\boldsymbol{\eta}_{Q+1}}^{\alpha+\beta-1}(\boldsymbol{x}_i)} \\
&= \prod_{i=1}^{k_1} p_{\boldsymbol{\eta}_1}(\boldsymbol{x}_i) \dots \prod_{i=k_{m-1}+1}^{k_m} p_{\boldsymbol{\eta}_m}(\boldsymbol{x}_i) \\
&\quad \times \prod_{i=k_m+1}^{k_m+u_m} \frac{p_{\boldsymbol{\eta}_m}^{\alpha+\beta}(\boldsymbol{x}_i)}{p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta-1}(\boldsymbol{x}_i)} \\
&\quad \times \prod_{i=k_m+u_m+1}^{k_{m+1}} p_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x}_i) \dots \prod_{i=k_Q+1}^{N} p_{\boldsymbol{\eta}_{Q+1}}(\boldsymbol{x}_i). \qquad (61)
\end{aligned}$$

By integrating over Ω, we notice that all variables $\boldsymbol{x}_i$, $i = 1, \dots, N$ are separated. Thus, given (61) the integrals w.r.t. $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k_m}, \boldsymbol{x}_{k_m+u_m+1}, \dots, \boldsymbol{x}_{k_Q}$ equal 1, then we have:

$$\begin{aligned}
&\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, u_m, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X} \\
&= \int_{(\Omega')^{u_m}} \prod_{i=k_m+1}^{k_m+u_m} \frac{p_{\boldsymbol{\eta}_m}^{\alpha+\beta}(\boldsymbol{x}_i)}{p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta-1}(\boldsymbol{x}_i)} \mathrm{d}\boldsymbol{x}_{k_m+1} \dots \mathrm{d}\boldsymbol{x}_{k_m+u_m} \\
&= \prod_{i=k_m+1}^{k_m+u_m} \left(\int_{\Omega'} \frac{p_{\boldsymbol{\eta}_m}^{\alpha+\beta}(\boldsymbol{x}_i)}{p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta-1}(\boldsymbol{x}_i)} \mathrm{d}\boldsymbol{x}_i \right) \\
&= \left(\int_{\Omega'} \frac{p_{\boldsymbol{\eta}_m}^{\alpha+\beta}(\boldsymbol{x})}{p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta-1}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} \right)^{u_m} \triangleq \rho_m^{u_m}(\alpha+\beta) \qquad (62)
\end{aligned}$$

in which Ω' denotes the observation space in one single time instant, i.e., $\prod_{k=1}^{N} \Omega' = \Omega$, and where we define

$$\rho_q(\alpha) \triangleq \int_{\Omega'} \frac{p_{\boldsymbol{\eta}_q}^{\alpha}(\boldsymbol{x})}{p_{\boldsymbol{\eta}_{q+1}}^{\alpha-1}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} \qquad (63)$$

with $q, r \in \{1, \dots, Q+1\}$ and $\alpha \in \mathbb{R}$.

b) Case D2 ($u_m = v_m < 0$): In this case, we follow the same method as in the previous case, and (62) becomes:

$$\begin{aligned}
&\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, u_m, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X} \\
&= \prod_{i=k_m+u_m+1}^{k_m} \left(\int_{\Omega'} \frac{p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta}(\boldsymbol{x}_i)}{p_{\boldsymbol{\eta}_m}^{\alpha+\beta-1}(\boldsymbol{x}_i)} \mathrm{d}\boldsymbol{x}_i \right) \\
&= \left(\int_{\Omega'} \frac{p_{\boldsymbol{\eta}_{m+1}}^{\alpha+\beta}(\boldsymbol{x})}{p_{\boldsymbol{\eta}_m}^{\alpha+\beta-1}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} \right)^{-u_m} \\
&= \rho_m^{-u_m}(1-\alpha-\beta). \qquad (64)
\end{aligned}$$

c) Case D3 ($u_m = -v_m > 0$): Using the same methodology, we obtain slightly different expressions since we have to rewrite $\ell(\boldsymbol{X}, \alpha, \beta, u_m, -u_m, \boldsymbol{k})$ as:

$$\begin{aligned}
&\ell(\boldsymbol{X}, \alpha, \beta, u_m, -u_m, \boldsymbol{k}) \\
&= \frac{p^{\alpha}(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k} + \boldsymbol{u}_m) \, p^{\beta}(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k} - \boldsymbol{u}_m)}{p^{\alpha+\beta-1}(\boldsymbol{X}|\boldsymbol{t} = \boldsymbol{k})} \\
&= \prod_{i=1}^{k_1} p_{\boldsymbol{\eta}_1}(\boldsymbol{x}_i) \dots \prod_{i=k_{m-1}+1}^{k_m-u_m} p_{\boldsymbol{\eta}_m}(\boldsymbol{x}_i) \\
&\quad \times \prod_{i=k_m-u_m+1}^{k_m} \frac{p_{\boldsymbol{\eta}_{m+1}}^{\beta}(\boldsymbol{x}_i)}{p_{\boldsymbol{\eta}_m}^{\beta-1}(\boldsymbol{x}_i)} \prod_{i=k_m+1}^{k_m+u_m} \frac{p_{\boldsymbol{\eta}_m}^{\alpha}(\boldsymbol{x}_i)}{p_{\boldsymbol{\eta}_{m+1}}^{\alpha-1}(\boldsymbol{x}_i)} \\
&\quad \times \prod_{i=k_m+u_m+1}^{k_{m+1}} p_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x}_i) \dots \prod_{i=k_Q+1}^{N} p_{\boldsymbol{\eta}_{Q+1}}(\boldsymbol{x}_i). \qquad (65)
\end{aligned}$$

which gives, by integrating over Ω:

$$\begin{aligned}
&\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, -u_m, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X} \\
&= \left(\int_{\Omega'} \frac{p_{\boldsymbol{\eta}_{m+1}}^{\beta}(\boldsymbol{x})}{p_{\boldsymbol{\eta}_m}^{\beta-1}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} \right)^{u_m} \left(\int_{\Omega'} \frac{p_{\boldsymbol{\eta}_m}^{\alpha}(\boldsymbol{x})}{p_{\boldsymbol{\eta}_{m+1}}^{\alpha-1}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} \right)^{u_m} \\
&= \rho_m^{u_m}(1-\beta) \, \rho_m^{u_m}(\alpha). \qquad (66)
\end{aligned}$$

d) Case D4 ($u_m = -v_m < 0$): This last case is once again very similar to the previous one: as in (66), we have

$$\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, -u_m, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X} = \rho_m^{-u_m}(\beta) \, \rho_m^{-u_m}(1-\alpha). \qquad (67)$$

Summary of cases D1–D4: Finally, by using the following discrete step-function U_γ defined for $\gamma \in \mathbb{R}$ as

$$U_\gamma(n) = \begin{cases} \gamma & \text{if } n \geq 0 \\ 1-\gamma & \text{if } n < 0 \end{cases} \qquad (68)$$

one can merge (62), (64), (66) and (67) into

$$\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, u_m, v_m, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X} = \begin{cases} \rho_m^{|u_m|}\big(U_{\alpha+\beta}(u_m)\big) & \text{if } v_m = u_m \\ \rho_m^{|u_m|}\big(U_\alpha(u_m)\big)\rho_m^{|v_m|}\big(U_\beta(v_m)\big) & \text{if } v_m = -u_m \end{cases} \qquad (69)$$

which reduces, for $v_m = 0$, $u_m \neq 0$ and $\beta = 0$, to:

$$\int_\Omega \ell(\boldsymbol{X}, \alpha, 0, u_m, 0, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X} = \rho_m^{|u_m|}\big(U_\alpha(u_m)\big). \qquad (70)$$

The same work now remains to be done for $n > m$.

2) Case UT ($m < n$): What differs from the previous case, is that u_m and v_n can differ not only in their sign, but also in their absolute value. As in the previous section, we split the study into four cases based on the respective signs of u_m and v_n ("UT" stands for "upper triangle"):

$$\begin{cases} \textit{a) Case UT1: } u_m > 0 \text{ and } v_n > 0 \\ \textit{b) Case UT2: } u_m < 0 \text{ and } v_n < 0 \\ \textit{c) Case UT3: } u_m < 0 \text{ and } v_n > 0 \\ \textit{d) Case UT4: } u_m > 0 \text{ and } v_n < 0. \end{cases} \qquad (71)$$

The study of the aforementioned four cases is similar to that of the previous section, i.e., we first write

$\ell(\boldsymbol{X},\alpha,\beta,u_m,v_m,\boldsymbol{k})$ in the same way as in (65), then we integrate over Ω. After some calculus, the obtained expressions in each case are given below.

a) Case UT1 ($u_m>0$ and $v_n>0$):

$$\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_n,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}=\rho_m^{u_m}(\alpha)\,\rho_n^{v_n}(\beta) \tag{72}$$

b) Case UT2 ($u_m<0$ and $v_n<0$):

$$\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_n,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}=\rho_m^{-u_m}(1-\alpha)\,\rho_n^{-v_n}(1-\beta) \tag{73}$$

c) Case UT3 ($u_m<0$ and $v_n>0$):

$$\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_n,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}=\rho_m^{-u_m}(1-\alpha)\,\rho_n^{v_n}(\beta) \tag{74}$$

d) Case UT4 ($u_m>0$ and $v_n<0$): In this case, it should be noticed that, despite we assumed $m<n$, which implies $k_m<k_n$, one could have $k_m+u_m>k_n+v_n$, meaning an *overlap* between the test-points, that we have to take into account. A careful inspection of tables II and III shows that this overlap can only occur if $n=m+1$ (i.e., for terms in the first superdiagonal of matrix $\boldsymbol{G}$, and by symmetry in the first subdiagonal). The necessary condition for at least one overlap point is then $k_m+u_m>(k_{m+1})_{\min}+v_{m+1}$, which is equivalent to $\min(u_m,-v_{m+1})\geq 2$. Thus, we obtain two expressions for $\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_n,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}$, depending on whether this condition is satisfied or not, i.e., whether there is at least one overlap point or there is none.

• *No overlap case:* $\min(u_m,-v_{m+1})=1$

In this situation, the calculations are similar to those in cases UT1–UT3, consequently we obtain:

$$\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_n,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}=\rho_m^{u_m}(\alpha)\,\rho_n^{-v_n}(1-\beta) \tag{75}$$

• *Overlap case:* $\min(u_m,-v_{m+1})\geq 2$

This situation requires special attention. In this case one must be careful about the order of the change-points in each p.d.f. when writing $\ell(\boldsymbol{X},\alpha,\beta,u_m,v_m,\boldsymbol{k})$ as in (65). In particular, one has $k_m<k_{m+1}+v_{m+1}<k_m+u_m<k_{m+1}$, then

$$\begin{aligned}
&\ell(\boldsymbol{X},\alpha,\beta,u_m,v_m,\boldsymbol{k})\\
&=\frac{p^\alpha(\boldsymbol{X}|\boldsymbol{t}=\boldsymbol{k}+\boldsymbol{u}_m)\,p^\beta(\boldsymbol{X}|\boldsymbol{t}=\boldsymbol{k}+\boldsymbol{v}_{m+1})}{p^{\alpha+\beta-1}(\boldsymbol{X}|\boldsymbol{t}=\boldsymbol{k})}\\
&=\prod_{i=1}^{k_1}p_{\boldsymbol{\eta}_1}(\boldsymbol{x}_i)\ldots\prod_{i=k_{m-1}+1}^{k_m}p_{\boldsymbol{\eta}_m}(\boldsymbol{x}_i)\\
&\quad\times\prod_{i=k_m+1}^{k_{m+1}+v_{m+1}}\frac{p^\alpha_{\boldsymbol{\eta}_m}(\boldsymbol{x}_i)}{p^{\alpha-1}_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x}_i)}\\
&\quad\times\prod_{i=k_{m+1}+v_{m+1}+1}^{k_m+u_m}\frac{p^\alpha_{\boldsymbol{\eta}_m}(\boldsymbol{x}_i)\,p^\beta_{\boldsymbol{\eta}_{m+2}}(\boldsymbol{x}_i)}{p^{\alpha+\beta-1}_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x}_i)}\\
&\quad\times\prod_{i=k_m+u_m+1}^{k_{m+1}}\frac{p^\beta_{\boldsymbol{\eta}_{m+2}}(\boldsymbol{x}_i)}{p^{\beta-1}_{\boldsymbol{\eta}_{m+1}}(\boldsymbol{x}_i)}\\
&\quad\times\prod_{i=k_{m+1}+1}^{k_{m+2}}p_{\boldsymbol{\eta}_{m+2}}(\boldsymbol{x}_i)\ldots\prod_{i=k_Q+1}^{N}p_{\boldsymbol{\eta}_{Q+1}}(\boldsymbol{x}_i)\,.
\end{aligned} \tag{76}$$

By integrating over Ω, we obtain:

$$\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_m,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}=\frac{\rho_m^{k_{m+1}+v_{m+1}-k_m}(\alpha)\,\rho_{m+1}^{k_{m+1}-(k_m+u_m)}(1-\beta)}{\kappa_m^{(k_{m+1}+v_{m+1})-(k_m+u_m)}(\alpha,1-\alpha-\beta)} \tag{77}$$

with function κ_m defined as in (25).

Summary of cases UT1–UT4 *(cases without overlap):* As in the previous section, we can summarize the previous results into:

$$\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,u_m,v_n,\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}=\rho_m^{|u_m|}\big(U_\alpha(u_m)\big)\rho_n^{|v_n|}\big(U_\beta(v_n)\big) \tag{78}$$

which is valid for $n>m+1$, and for $n=m+1$ only if $\min(u_m,-v_{m+1})=1$.

C. Final expression of $\zeta(\alpha,\beta,u_m,v_n)$

Finally, it is possible to deduce the closed-form expressions of $\zeta(\alpha,\beta,u_m,v_n)$ given in Section IV-A of the main body, by plugging results from Appendix A and B into (48). Since results from Appendix B depend upon the three cases D ($m=n$), UT ($m+1<n$) and FSD ($m+1=n$) (see (69), (78) and (77)), we need to separate them in the following as well.

1) Case D ($m=n$): Equations (13), (14) and (15) are obtained by plugging (57) and (69) or (70) into (48). Since neither (57) nor (69) depend on $\boldsymbol{k}$, they both can be taken out of the sum over $\mathcal{I}$ in (48). This leads to function $f_D(\tau,u_m,v_m)$ introduced in (16), whose actual basic definition is $f_D(\tau,u_m,v_m)\triangleq\frac{\mathrm{Card}(\mathcal{I})}{\tau^Q}=\frac{\sum_{\mathcal{I}}1}{\tau^Q}$, and whose expression is directly deduced from Table I.

2) Case UT ($m+1<n$): Equation (18) is obtained by plugging (57) and (78) into (48). Here again, since neither (57) nor (78) depend on $\boldsymbol{k}$, they both can be taken out of the sum over $\mathcal{I}$ in (48). This leads to function $f_{UT}(\tau,u_m,v_n)$ introduced in (19), whose actual basic definition is $f_{UT}(\tau,u_m,v_n)\triangleq\frac{\mathrm{Card}(\mathcal{I})}{\tau^Q}$ as well, and its expression is deduced from Table II.

3) Case $m+1=n$: This case is tricky because we need to distinguish between the cases without and with overlap between the test-points.

a) Case without overlap: Equation (20) is obtained by plugging (57) and (75) into (48). Here again, since neither (57) nor (75) depend on $\boldsymbol{k}$, they both can be taken out of the sum over $\mathcal{I}$ in (48). This leads to function $f_{FSD}(\tau,u_m,v_{m+1})$ introduced in (21), whose actual basic definition is $f_{FSD}(\tau,u_m,v_{m+1})\triangleq\frac{\mathrm{Card}(\mathcal{I})}{\tau^Q}$ as well, and its expression is deduced from Table III.

b) Case with possible overlap, i.e., $u_m>0$ and $v_{m+1}<0$: This case is different from the previous ones, since $\int_\Omega \ell(\boldsymbol{X},\alpha,\beta,\boldsymbol{u}_m,\boldsymbol{v}_{m+1},\boldsymbol{k})\,\mathrm{d}\boldsymbol{X}$ depends on $\boldsymbol{k}$ if $k_m+u_m>k_{m+1}+v_{m+1}$ (see (77)). Consequently, we split the discrete sum in (48) into two parts depending on whether $k_m+|u_m|\gtrless k_{m+1}-|v_{m+1}|$. More precisely, for $k_{m+1}\in\left\{k_m+(|u_m|,|v_{m+1}|)^+1,\ldots,k_m+|u_m|+|v_{m+1}|-1\right\}\triangleq\mathcal{I}^{\mathsf{ov}}_{m+1}$ there is overlap, leading to (77). On the other hand,

for $k_{m+1} \in \{k_m + |u_m| + |v_{m+1}|, \ldots, k_m + \tau\} \triangleq \mathcal{I}^{\mathsf{no}}_{m+1}$, there is no overlap, thus, $\int_\Omega \ell(\boldsymbol{X}, \alpha, \beta, \boldsymbol{u}_m, \boldsymbol{v}_{m+1}, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{X}$ is given by (78) with $n = m+1$. Notice that the only situation in which there is no overlap terms at all, i.e., $\mathrm{Card}\left(\mathcal{I}^{\mathsf{ov}}_{m+1}\right) = 0$, is when $\min(|u_m|, |v_{m+1}|) = 1$, then we can write $\zeta(\alpha, \beta, u_m, v_{m+1})$ according to (20).

Let us then derive ζ in a case with at least one overlap term, i.e., $\min(|u_m|, |v_{m+1}|) \geq 2$, or equivalently $\mathrm{Card}\left(\mathcal{I}^{\mathsf{ov}}_{m+1}\right) \geq 1$. Let us also first assume that $m+1 < Q$. We can then rewrite ζ as:

$$\zeta(\alpha, \beta, u_m, v_{m+1}) = \left(\frac{1}{\tau}\right)^Q \sum_{\boldsymbol{k}_{1:m} \in \mathcal{I}_{1:m}} \left[z_1(\alpha, \beta, u_m, v_m, \boldsymbol{k}_{1:m}) + z_2(\alpha, \beta, u_m, v_m, \boldsymbol{k}_{1:m})\right] \tag{79}$$

where $\boldsymbol{k}_{i:j}$ denotes the truncated vector $[k_i, \ldots, k_j]^T$, and $\mathcal{I}_{i:j}$ the cartesian product $\mathcal{I}_i \times \ldots \times \mathcal{I}_j$, and where the overlapping terms are included in

$$z_1(\alpha, \beta, u_m, v_m, \boldsymbol{k}_{1:m}) \triangleq \sum_{k_{m+1} \in \mathcal{I}^{\mathsf{ov}}_{m+1}} \left(\sum_{\boldsymbol{k}_{m+2:Q} \in \mathcal{I}_{m+2:Q}} \left[\rho_m^{k_{m+1} - |v_{m+1}| - k_m}(\alpha) \times \kappa_m^{(k_m + |u_m|) - (k_{m+1} - |v_{m+1}|)}(\alpha, 1 - \alpha - \beta) \times \rho_{m+1}^{k_{m+1} - (k_m + |u_m|)}(\beta) \right] \right) \tag{80}$$

and the non-overlapping terms are included in

$$z_2(\alpha, \beta, u_m, v_m, \boldsymbol{k}_{1:m}) = \rho_m^{|u_m|}(\alpha) \, \rho_{m+1}^{|v_{m+1}|}(1 - \beta) \sum_{k_{m+1} \in \mathcal{I}^{\mathsf{no}}_{m+1}} \left(\sum_{\boldsymbol{k}_{m+2:Q} \in \mathcal{I}_{m+2:Q}} 1 \right). \tag{81}$$

The expression of z_2 in (81) can be simplified in the same way as in the non-overlapping cases, and we obtain

$$z_2(\alpha, \beta, u_m, v_m, \boldsymbol{k}_{1:m}) = \tau^{Q-m-2}(\tau - |v_{m+1}|) \times (\tau - |u_m| - |v_{m+1}| + 1)^+ \times \rho_m^{|u_m|}(\alpha) \, \rho_{m+1}^{|v_{m+1}|}(1 - \beta), \tag{82}$$

and z_1 in (80) can be simplified as

$$z_1(\alpha, \beta, u_m, v_m, \boldsymbol{k}_{1:m}) = \tau^{Q-m-2}(\tau - |v_{m+1}|) \, \rho_m^{|u_m|}(\alpha) \times \rho_{m+1}^{|v_{m+1}|}(1 - \beta) \frac{\left(R_m(\alpha, \beta)\right)^{1 - (|u_m|, |v_{m+1}|)^-} - 1}{1 - R_m(\alpha, \beta)} \tag{83}$$

where $(x, y)^- \triangleq \min(x, y, 0)$, and $R_m(\alpha, \beta)$ is the common ratio of the geometric series appearing in (79), defined as in (24).

The manipulations are very similar in the case $m + 1 = Q$. Hence, by noticing that neither z_1 nor z_2 actually depend on $\boldsymbol{k}_{1:m}$, we find that (79) results in (22).

References

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

[2] J. Chen and A. K. Gupta, *Parametric Statistical Change Point Analysis*. Birkhäuser Basel, 2000.

[3] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., ser. Springer Texts in Statistics. New-York, NY, USA: Springer, Sep. 2003.

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

[5] S. B. Fotopoulos and S. K. Jandhyala, "Maximum likelihood estimation of a change-point for exponentially distributed random variables," *ELSEVIER Statistics and Probability Letters*, vol. 51, pp. 423–429, 2001.

[6] S. B. Fotopoulos, S. K. Jandhyala, and E. Khapalova, "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] D. Slepian, "Estimation of signal parameters in the presence of noise," *Trans. IRE Professional Group Inf. Theory*, vol. 3, no. 3, pp. 68–69, 1954.

[8] G. W. Bangs, "Array processing with generalized beamformers," Ph.D. dissertation, Yale University, New Haven, CT, USA, 1971.

[9] O. Besson and Y. Abramovich, "On the Fisher information matrix for multivariate elliptically contoured distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, Nov. 2013.

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

[11] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the high-SNR conditional maximum-likelihood estimator full statistical characterization," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4840–4843, Dec. 2006.

[12] 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.

[13] A. Swami and B. Sadler, "Cramér-Rao bounds for step-change localization in additive and multiplicative noise," in *Proc. of IEEE Workshop on Statistical Signal and Array Processing (SSAP)*, Portland, OR, USA, Sep. 1998, pp. 403–406.

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

[15] 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.

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

[17] 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.

[18] 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.

[19] M. A. Girshick and H. Rubbin, "A bayes approach to a quality control model," *The Annals of Mathematical Statistics*, vol. 23, no. 1, pp. 114–125, 1952.

[20] A. N. Shyriaev, "The problem of the most rapid detection of a disturbance in a stationary process," *Soviet Math. Dokl.*, no. 2, pp. 795–799, 1961.

[21] L. Bacharach, A. Renaux, M. N. El Korso, and E. Chaumette, "Weiss-Weinstein bound for change-point estimation," in *Proc. of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Cancún, Mexico, Dec. 2015, pp. 477–480.

[22] W. Xu, "Performance bounds on matched-field methods for source localization and estimation of ocean environmental parameters," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, Jun. 2001.

[23] W. Xu, A. B. Baggeroer, and K. L. Bell, "A bound on mean-square estimation error with background parameter mismatch," *IEEE Transactions on Information Theory*, vol. 50, no. 4, pp. 621–632, Apr. 2004.

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

[25] A. Renaux, P. Forster, P. Larzabal, C. D. Richmond, and A. Nehorai, "A fresh look at the Bayesian bounds of the Weiss-Weinstein family," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5334–5352, Nov. 2008.

[26] Y.-C. Yao, "Approximating the distribution of the maximum likelihood estimate of the change-point in a sequence of independent random variables," *The Annals of Statistics*, vol. 15, no. 3, pp. 1321–1328, 1987.

[27] R. E. Quandt, "The estimation of the parameters of a linear regression system obeying two separate regimes," *Journal of the American Statistical Association*, vol. 53, no. 284, pp. 873–880, Dec. 1958.

[28] K. J. Worsley, "Confidence regions and tests for a change-point in a sequence of exponential family random variables," *Biometrika*, vol. 73, no. 1, pp. 91–104, Apr. 1986.

[29] V. F. Pisarenko, A. F. Kushnir, and I. V. Savin, "Statistical adaptive algorithms for estimation of onset moments of seismic phases," *Physics of the Earth and Planetary Interiors*, vol. 47, pp. 4–10, Aug. 1987.

[30] Z. Harchaoui and O. Cappe, "Retrospective mutiple change-point estimation with kernels," in *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, Aug. 2007, pp. 768–772.

[31] 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.

[32] J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optimization*. 1st ed. New York: Springer-Verlag, 2000.

[33] 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.

[34] G. Meurant, "A review on the inverse of symmetric tridiagonal and block tridiagonal matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 3, pp. 707–728, Jul. 1992.

[35] H. Chernoff and S. Zacks, "Estimating the current mean of a normal distribution which is subjected to changes in time," *The Annals of Mathematical Statistics*, vol. 35, no. 3, pp. 999–1018, 1964.

[36] N. Dobigeon, J.-Y. Tourneret, and J. D. Scargle, "Joint segmentation of multivariate astronomical time series: Bayesian sampling with a hierarchical model," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 414–423, Feb. 2007.

[37] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[38] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming*. Dec. 2008 [Online]. Available: *http://stanford.edu/boyd/cvx*, Web page and software.

[39] C. Choirat and R. Seri, "Estimation in discrete parameter models," *Statistical Science*, vol. 27, no. 2, pp. 278–293, 2012.

Lucien Bacharach Lucien Bacharach was born in Cahors, France, in 1991. He studied aeronautical engineering at ISAE-ENSICA (Toulouse, France), where he got graduated in 2014, and received the Master Research degree in Signal and Image Processing from Paris-Sud University (Orsay, France) in 2015. He is currently pursuing the Ph.D. degree in signal processing with Paris-Sud University. His research interests and activities include estimation theory and performance analysis in statistical signal processing, especially lower bounds on the mean square error for signal models including change-points.

Alexandre Renaux Alexandre Renaux received the Ph.D. degrees in electrical engineering from École Normale Supérieure de Cachan, Cachan, France, in 2006. From 2006 to 2007, he was a Postdoctoral Research Associate at the Department of Electrical and Systems Engineering , Washington University in St. Louis, St. Louis, MO, USA. Since 2007, he has been an Assistant Professor (and currently Associate Professor since 2011) at the Physics Department of Paris-Sud University, Orsay, France. His research interests are in the field of estimation theory and lower bounds on the mean square error applied to statistical signal processing.

Mohammed Nabil El Korso Mohammed Nabil El Korso was born in Oran, Algeria. He received the M.Sc. in Electrical Engineering from the National Polytechnic School, Algeria in 2007. He obtained the Master Research degree in Signal and Image Processing from Paris-Sud XI University, France in 2008. In 2011, he obtained his Ph.D. degree from Paris-Sud XI University. From 2011 to 2012, he was a research scientist in the Communication Systems Group at Technische Universität Darmstadt, Germany. He was Assistant Professor at École Normale Supérieure de Cachan from 2012 to 2013. Currently, he is Assistant Professor at University of Paris Ouest Nanterre la Défense and a member of LEME (EA4416) laboratory. His research interests include statistical signal processing, estimation/detection theory with applications to array signal processing.

Éric Chaumette Éric Chaumette was born in 1965 at Chartres (France). He studied Electronics and Signal Processing both at ENAC (Toulouse, France) where he obtained an Engineer degree in 1989, and at Toulouse University where he obtained a M.Sc. degree in Signal Processing in 1989. From 1990 to 2007, he was with Thales in various radar studies departments. From 2007 to 2013, he was with the Electromagnetic and Radar Division of the French Aerospace Lab (ONERA), Palaiseau, France, as a research engineer. Simultaneously, from 2000 to 2014, he was a research associate at laboratory SATIE, CNRS, École Normale Supérieure de Cachan, France, where he received the PhD degree in 2004 and the "habilitation à diriger les recherches" in 2014. He is now with the Department of Electronics, Optronics and Signal of ISAE (Institut Supérieur de l'Aéronautique et de l'Espace). Main domains of interest are related to radar scene modelling, detection and estimation theory and navigation.