

Bornes bayésiennes pour la localisation d'un point de rupture : Application à des processus exponentiels

Lucien BACHARACH¹, Geoffrey BIBICHE¹, Alexandre RENAUX¹, Mohammed Nabil EL KORSO²

¹Université Paris-Sud, Laboratoire des Signaux et Systèmes (L2S)
3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France

²Université Paris-Ouest Nanterre-La Défense, Laboratoire Énergétique Mécanique Électromagnétisme (LEME)
50 rue de Sèvres, 92410 Ville d'Avray, France

lucien.bacharach@u-psud.fr, geoffrey.bibiche@u-psud.fr, alexandre.renaux@u-psud.fr
m.elkorso@u-paris10.fr

Résumé – Dans cette communication, nous donnons une expression analytique de la borne de Weiss-Weinstein sur l'erreur quadratique moyenne pour la localisation d'un point de rupture dans des séries temporelles indépendantes identiquement distribuées, quel que soit leur type de distribution. L'expression est ensuite spécifiée dans le cas où les observations suivent une loi exponentielle. Enfin, le comportement de la borne est illustré avec des simulations numériques, en la comparant avec l'erreur quadratique moyenne de l'estimateur du maximum *a posteriori*.

Abstract – In this paper, closed-form expressions are given for the Weiss-Weinstein bound on the mean square error, in the case of independent, identically distributed time series submitted to a change-point, i.e., an abrupt change in their distribution. Expressions are first given whatever the observations distribution, and then specified for exponentially distributed observations. Finally, numerical results are provided to illustrate the behavior of the bound regarding to the mean square error of the maximum *a posteriori* estimator.

1 Introduction

Dans de nombreuses applications pratiques, il se peut que des changements abrupts, appelés points de rupture, surviennent dans la distribution du signal d'intérêt. C'est le cas par exemple en contrôle de qualité, en traitement de la parole ou bien pour la maintenance de systèmes de navigation [1, 3]. Dans la plupart des cas, ces signaux comportent une composante aléatoire et les instants d'occurrence des changements sont inconnus ; ils doivent donc être estimés. De plus, puisque l'on considère des signaux échantillonnés, l'instant de rupture à estimer est un paramètre discret.

Dans ce contexte, l'estimation de la localisation d'un point de rupture a été largement étudiée au cours des cinquante dernières années. Beaucoup d'algorithmes ont été proposés dans ce but précis, et parmi eux, ceux basés sur le maximum de vraisemblance (EMV) ont fait l'objet d'une attention particulière [6–8]. Cela s'explique par ses propriétés d'optimalité asymptotique pour une large classe de problèmes [9], auquel cas il est asymptotiquement normal et atteint en particulier la borne de Cramér-Rao (BCR). Notons que ce n'est pas le cas pour l'estimation d'un point de rupture puisque ce type de problème ne respecte pas les conditions de régularité classiques. Une première expression asymptotique (i.e., pour un nombre d'échantillons avant et après la rupture infiniment grands) de l'EMV a été donnée par Hinkley, démontrant également sa non-consistance [8]. Des formes exactes de cette distribution asymptotique, plus raison-

nables en termes calculatoires, ont été obtenues plus tard pour des données suivant une loi exponentielle [6] puis normale [7]. Comme en témoigne le temps écoulé entre l'obtention de ces différents résultats, il n'est pas aisé en général de connaître précisément le comportement d'un estimateur tel que l'EMV, pour ce problème. C'est d'autant plus le cas en contexte non asymptotique, et pour des distributions moins couramment rencontrées que l'exponentielle ou la gaussienne.

Pour pallier à ces difficultés, une alternative est de s'intéresser aux bornes inférieures de la variance d'estimateurs d'un point de rupture, comme la borne de Cramér-Rao (BCR) mentionnée plus haut. Définie simplement comme l'inverse de l'information de Fisher, elle permet de définir une notion d'efficacité [9]. Elle requiert cependant des conditions de régularité pour être calculée, comme la différentiabilité de la fonction de vraisemblance par rapport au paramètre inconnu à estimer. Or dans notre cas, comme celui-ci est discret, il est impossible d'obtenir directement une expression de la BCR. Comme cela a été fait dans la littérature, on peut la calculer de façon approchée [10], donnant un aperçu des performances d'estimation malheureusement trop vague. Il existe cependant une inégalité plus générale que celle de Cramér-Rao, qui met en jeu la borne dite de Chapman et Robbins [2]. Celle-ci peut être utilisée pour l'estimation de paramètres discrets, et il a été démontré récemment que pour ce type de problème, l'EMV n'atteint pas cette borne à fort nombre d'échantillons [4]. Quelques années plus tôt, Ferrari et Tourneret [5] avaient d'ailleurs observé un écart asymptotique

tique (cette fois-ci en terme de rapport signal à bruit) entre l'erreur quadratique moyenne (EQM) de l'EMV et la borne de Chapman-Robbins, dans le cas de l'estimation d'un point de rupture.

Afin d'essayer de fournir une meilleure approximation du comportement d'estimation, nous nous intéressons à la borne de Weiss-Weinstein (WWB) pour ce problème, impliquant de se placer dans le cadre bayésien. Cette borne est reconnue pour être l'une des plus proches du comportement réel d'estimation, mais également pour le peu de conditions de régularité qu'elle nécessite. Nous rappelons donc ici son expression générale après avoir posé formellement le problème, nous donnons son expression dans le cas de l'estimation d'un point de rupture quelle que soit la famille de distribution des observations avant et après la rupture, et enfin nous appliquons ces résultats au cas d'observations distribuées exponentiellement, pour lequel nous fournissons des résultats numériques.

2 Position du problème

On considère le cas général d'un vecteur de N observations indépendantes $\mathbf{x} = [x_1, x_2, \dots, x_N]^\top \in \Omega \subset \mathbb{R}^N$, comportant un point de rupture unique noté t . Ces observations sont modélisées par :

$$\begin{cases} x_n \sim p_{\eta_1}(x_n) & \text{pour } n = 1, \dots, t, \\ x_n \sim p_{\eta_2}(x_n) & \text{pour } n = t + 1, \dots, N, \end{cases} \quad (1)$$

où $p_{\eta_{j=1,2}}$ désignent les densités de probabilité de paramètres $\eta_{j=1,2}$. On suppose que ces densités appartiennent à la même famille de distribution, et que les observations sont stationnaires avant et après le point de rupture. C'est ce dernier paramètre qui est à estimer : $t \in \llbracket 1, N-1 \rrbracket$. Dans notre cadre d'étude bayésien, on suppose donc que t est une variable aléatoire (discrète) sur l'espace de paramètres $\mathcal{T} = \mathbb{N}$. Dans un premier temps, nous considérons une loi *a priori* quelconque pour t , que nous notons $\pi_t(k) = \Pr(t = k)$.

3 Étude de la borne de Weiss-Weinstein

3.1 Expression générale dans le cas discret

Nous rappelons ici l'expression de la borne de Weiss-Weinstein (BWW), notée WWB, pour un paramètre scalaire t . On a l'inégalité suivante pour tout estimateur bayésien $\hat{t}(\mathbf{x})$ du paramètre inconnu t (comme par exemple celui du MAP) [12]

$$\mathbb{E}_{\mathbf{x},t} \left\{ [\hat{t}(\mathbf{x}) - t]^2 \right\} \geq \sup_{h,s} W(h,s) \triangleq \text{WWB} \quad (2)$$

où $\mathbb{E}_{\mathbf{x},t} \{ \cdot \}$ représente l'espérance par rapport à la loi jointe $p_{\mathbf{x},t}(\mathbf{x}, k) = p(\mathbf{x} | t = k) \pi_t(k)$ – c'est-à-dire que le premier membre de (2) représente l'EQM globale de l'estimateur $\hat{t}(\mathbf{x})$,

et où, au second membre de (2), on définit [12]

$$W(h, s) = \frac{h^2 \mathbb{E}_{\mathbf{x},t}^2 \left\{ \left(\frac{p_{\mathbf{x},t}(\mathbf{x}, k+h)}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right)^s \right\}}{\mathbb{E}_{\mathbf{x},t} \left\{ \left[\left(\frac{p_{\mathbf{x},t}(\mathbf{x}, k+h)}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right)^s - \left(\frac{p_{\mathbf{x},t}(\mathbf{x}, k-h)}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right)^{1-s} \right]^2 \right\}} \quad (3)$$

avec h et s tels que $k \pm h \in \mathcal{T}$ et $s \in]0, 1[$, sous réserve que le dénominateur de (3) soit non nul. Le fait de maximiser le second membre de (2) par rapport à ces deux paramètres permet d'obtenir une borne précise. Concernant le paramètre s , la valeur $s = 1/2$ est souvent notée comme étant celle conduisant au maximum [11–13], c'est pourquoi nous choisirons cette valeur dans la suite et donnerons donc l'expression de $W(h, 1/2)$. Dès lors, en notant $\mathcal{T}' = \{k \in \mathcal{T} \mid p_{\mathbf{x},t}(\mathbf{x}, k) > 0\}$, le terme au numérateur de (3) peut s'écrire

$$\mathbb{E}_{\mathbf{x},t} \left\{ \sqrt{\frac{p_{\mathbf{x},t}(\mathbf{x}, k+h)}{p_{\mathbf{x},t}(\mathbf{x}, k)}}} \right\} = \sum_{k \in \mathcal{T}'} B(k+h, k) \sqrt{\pi_t(k+h) \pi_t(k)} \quad (4)$$

où on a défini, pour $i, j \in \mathcal{T}'$

$$B(i, j) \triangleq \int_{\Omega} \sqrt{p(\mathbf{x} | t = i) p(\mathbf{x} | t = j)} d\mathbf{x} \quad (5)$$

qui est la distance de Bhattacharyya entre les densités conditionnelles $p(\mathbf{x} | t = i)$ et $p(\mathbf{x} | t = j)$.

Au dénominateur de (3), on a

$$\begin{aligned} \mathbb{E}_{\mathbf{x},t} \left\{ \left[\sqrt{\frac{p_{\mathbf{x},t}(\mathbf{x}, k+h)}{p_{\mathbf{x},t}(\mathbf{x}, k)}}} - \sqrt{\frac{p_{\mathbf{x},t}(\mathbf{x}, k-h)}{p_{\mathbf{x},t}(\mathbf{x}, k)}}} \right]^2 \right\} \\ = \mathbb{E}_{\mathbf{x},t} \left\{ \frac{p_{\mathbf{x},t}(\mathbf{x}, k+h)}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right\} + \mathbb{E}_{\mathbf{x},t} \left\{ \frac{p_{\mathbf{x},t}(\mathbf{x}, k-h)}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right\} \\ - 2 \mathbb{E}_{\mathbf{x},t} \left\{ \frac{\sqrt{p_{\mathbf{x},t}(\mathbf{x}, k+h) p_{\mathbf{x},t}(\mathbf{x}, k-h)}}}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right\}. \end{aligned} \quad (6)$$

Les deux premiers termes au second membre de (6) s'écrivent

$$\begin{aligned} \mathbb{E}_{\mathbf{x},t} \left\{ \frac{p_{\mathbf{x},t}(\mathbf{x}, k \pm h)}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right\} &= \sum_{k \in \mathcal{T}'} \left[\pi_t(k \pm h) \int_{\Omega} p(\mathbf{x} | t = k \pm h) d\mathbf{x} \right] \\ &= \sum_{k \in \mathcal{T}'} \pi_t(k \pm h) \end{aligned} \quad (7)$$

et de manière analogue, le troisième terme peut être développé comme suit

$$\begin{aligned} \mathbb{E}_{\mathbf{x},t} \left\{ \frac{\sqrt{p_{\mathbf{x},t}(\mathbf{x}, k+h) p_{\mathbf{x},t}(\mathbf{x}, k-h)}}}{p_{\mathbf{x},t}(\mathbf{x}, k)} \right\} &= \\ \sum_{k \in \mathcal{T}'} B(k+h, k-h) \sqrt{\pi_t(k+h) \pi_t(k-h)}. \end{aligned} \quad (8)$$

On remarque que (8) fait intervenir le terme $B(k+h, k-h)$, or, à l'aide du changement de variable $k' = k-h$, on trouve que $B(k+h, k-h) = B(k'+2h, k')$. Ainsi il suffira d'obtenir l'expression de $B(k+h, k)$ pour en déduire directement celle de $B(k+h, k-h)$, en remplaçant h par $2h$ dans la première.

On peut donc déduire, en réinjectant (4), (7) et (8) dans (3), une expression de la borne de Weiss-Weinstein pour un paramètre scalaire discret :

$$\text{WWB} = \sup_h \frac{h^2 \left(\sum_{k \in \mathcal{G}'} B(k+h, k) \sqrt{\pi_t(k+h)\pi_t(k)} \right)^2}{\left(\sum_{k \in \mathcal{G}'} \pi_t(k+h) + \sum_{k \in \mathcal{G}'} \pi_t(k-h) - \sum_{k \in \mathcal{G}'} \left[B(k+2h, k) \sqrt{\pi_t(k+2h)\pi_t(k)} \right] \right)} \quad (9)$$

en utilisant également le changement de variable $k' = k - h$ pour le terme $\sqrt{\pi_t(k+h)\pi_t(k-h)}$ de (8).

3.2 Expression pour l'estimation d'un point de rupture

On cherche désormais à obtenir l'expression de $B(k+h, k)$ à partir de (5), dans le cas du modèle (1). En supposant d'abord $h > 0$, par indépendance des observations, on a

$$\begin{aligned} & \sqrt{p(\mathbf{x} | t = k+h) p(\mathbf{x} | t = k)} \\ &= \prod_{n=1}^{k+h} \sqrt{p_{\eta_1}(x_n)} \prod_{n=k+h+1}^N \sqrt{p_{\eta_2}(x_n)} \prod_{n=1}^k \sqrt{p_{\eta_1}(x_n)} \prod_{n=k+1}^N \sqrt{p_{\eta_2}(x_n)} \\ &= \prod_{n=1}^k p_{\eta_1}(x_n) \prod_{n=k+1}^{k+h} \sqrt{p_{\eta_1}(x_n) p_{\eta_2}(x_n)} \prod_{n=k+h+1}^N p_{\eta_2}(x_n). \quad (10) \end{aligned}$$

On intègre ensuite sur l'espace des observations Ω , que l'on note comme le produit cartésien des espaces d'observation Ω' en chaque instant n : $\Omega \triangleq \bigotimes_{n=1}^N \Omega' = (\Omega')^N$. Les N variables $(x_n)_{n \in \llbracket 1, N \rrbracket}$ étant séparées, on obtient

$$\begin{aligned} & B(k+h, k) \\ &= \int_{\Omega} \left[\prod_{n=1}^k p_{\eta_1}(x_n) \prod_{n=k+1}^{k+h} \sqrt{p_{\eta_1}(x_n) p_{\eta_2}(x_n)} \prod_{n=k+h+1}^N p_{\eta_2}(x_n) \right] dx \\ &= \prod_{n=1}^k \left[\int_{\Omega'} p_{\eta_1}(x_n) dx_n \right] \prod_{n=k+1}^{k+h} \left[\int_{\Omega'} \sqrt{p_{\eta_1}(x_n) p_{\eta_2}(x_n)} dx_n \right] \\ & \quad \times \prod_{n=k+h+1}^N \left[\int_{\Omega'} p_{\eta_2}(x_n) dx_n \right] \quad (11) \end{aligned}$$

Dans (11), de manière évidente, les intégrales du premier produit sont toutes égales à 1, de même que dans le dernier produit. Au final, on obtient donc

$$B(k+h, k) = \left(\int_{\Omega'} \sqrt{p_{\eta_1}(x) p_{\eta_2}(x)} dx \right)^h \triangleq \rho_{1,2}^h \quad (12)$$

où il est intéressant de remarquer que $B(k+h, k)$ ne dépend en fait pas de k . Pour $h < 0$, on obtient par le même raisonnement $B(k+h, k) = \rho_{1,2}^{-h}$, d'où finalement, quel que soit h , $B(k+h, k) = \rho_{1,2}^{|h|}$. Ainsi, en injectant (12) dans (9), on obtient l'expression générale de la borne de Weiss-Weinstein

pour l'estimation d'un point de rupture et pour une loi *a priori* quelconque :

$$\text{WWB} = \sup_h \frac{h^2 \rho_{1,2}^{2h} \left(\sum_{k \in \mathcal{G}'} \sqrt{\pi_t(k+h)\pi_t(k)} \right)^2}{\left(\sum_{k \in \mathcal{G}'} \pi_t(k+h) + \sum_{k \in \mathcal{G}'} \pi_t(k-h) - 2 \rho_{1,2}^{2h} \sum_{k \in \mathcal{G}'} \sqrt{\pi_t(k+2h)\pi_t(k)} \right)} \quad (13)$$

Cette expression est valable quelles que soient les densités avant et après la rupture, et quel que soit la loi *a priori* choisie. En choisissant un *a priori* uniforme, i.e., $\pi_t(k) = \mathbb{1}_{\llbracket 1, N-1 \rrbracket}(k)$, les sommes discrètes se simplifient de la manière suivante :

$$\begin{aligned} \sum_{k \in \mathcal{G}'} \pi_t(k \pm h) &= \frac{1}{N-1} \sum_{k=1}^{N-1} \mathbb{1}_{\llbracket 1, N-1 \rrbracket}(k \pm h) \\ &= 1 - \frac{|h|}{N-1} \quad (14) \end{aligned}$$

et de la même manière,

$$\sum_{k \in \mathcal{G}'} \sqrt{\pi_t(k+h)\pi_t(k)} = 1 - \frac{|h|}{N-1}. \quad (15)$$

Avec cette loi *a priori* uniforme, on obtient donc une forme simple de la borne de Weiss-Weinstein (13), qui s'écrit donc

$$\text{WWB} = \max_{h \in \mathcal{H}} \frac{h^2 \rho_{1,2}^{2h} \left(1 - \frac{|h|}{N-1}\right)^2}{2 \left(1 - \frac{|h|}{N-1} - \left(1 - \frac{2|h|}{N-1}\right) \rho_{1,2}^{2h}\right)} \quad (16)$$

où $\mathcal{H} = \llbracket 1, N-2 \rrbracket$, les valeurs négatives de h n'étant pas nécessaires au vu de la parité de (16). Ainsi, le seul terme restant à calculer explicitement est $\rho_{1,2}$, qui va dépendre du signal considéré.

3.3 Application à des processus exponentiels

Nous supposons maintenant qu'en chaque instant, les observations x_n sont suivent une loi exponentielle de moyenne $1/\lambda_1$ pour $n \leq t$, et de moyenne $1/\lambda_2$ pour $n > t$, c'est-à-dire : $p_{\lambda_j}(x_n) = \lambda_j \exp\{-\lambda_j x_n\}$, $\lambda_j > 0$, $j = 1, 2$ (ici les vecteurs $\boldsymbol{\eta}_{j=1,2}$ de (1) deviennent les scalaires $\lambda_{j=1,2}$). Rappelons que ce type de signaux est rencontré par exemple lorsqu'on observe le module au carré d'un signal complexe gaussien circulaire, dont les parties réelles et imaginaires sont toutes deux de variance $1/(2\lambda_1)$ avant rupture et $1/(2\lambda_2)$ après rupture.

On peut alors obtenir une expression explicite de $\rho_{1,2}$ à partir de (12) :

$$\begin{aligned} \rho_{1,2} &= \int_0^{+\infty} \sqrt{\lambda_1 \lambda_2} \exp\left\{-\frac{\lambda_1 + \lambda_2}{2} x\right\} dx \\ &= \frac{2\sqrt{\lambda_1 \lambda_2}}{\lambda_1 + \lambda_2} = \frac{2\sqrt{\text{rsb}}}{1 + \text{rsb}} \quad (17) \end{aligned}$$

où rsb désigne le rapport signal à bruit, que l'on définit comme $\text{rsb} = \lambda_1/\lambda_2$, de la même façon que dans [5]. La borne de Weiss-Weinstein pour ce problème est donc obtenue en injectant (17) dans (16).

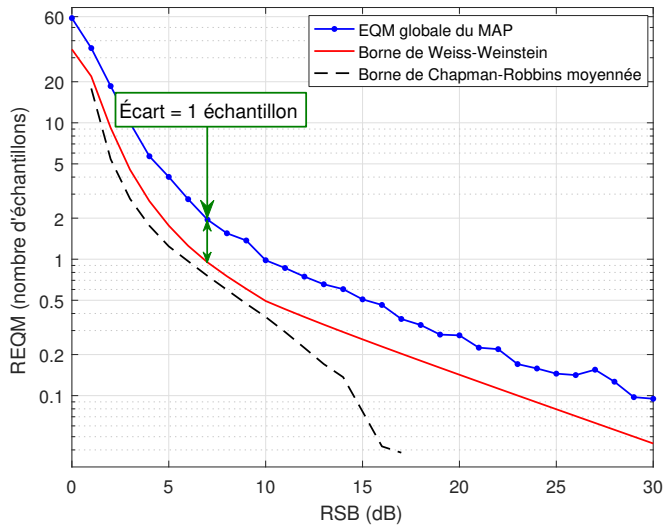


FIGURE 1 – Comparaison en terme de nombre d'échantillons d'erreur entre l'EQM de l'estimateur du MAP (courbe et points bleus), la borne proposée (WWB, en rouge, trait continu), et la borne de Chapman-Robbins initialement présentée dans [5] (ChRB, en noir, pointillés), moyennée par rapport à la loi *a priori*. La borne de Weiss-Weinstein proposée est plus précise, bien qu'il reste un écart entre l'EQM et la borne, dû à la nature discrète du paramètre à étudier [4].

4 Résultats numériques

Nous donnons dans cette section des résultats de simulation obtenus pour le cas de processus exponentiels, décrits dans la section précédente. Ces résultats ont été obtenus pour $N = 128$ observations et en choisissant $\lambda_1 = 1$. On compare la borne proposée à l'EQM globale empirique de l'estimateur du maximum *a posteriori* $\hat{t}^{\text{MAP}}(x)$, obtenue après 1000 simulations de Monte-Carlo. Plus précisément, on trace la racine carrée de l'EQM (REQM), afin de permettre une comparaison d'erreur en terme de nombre d'échantillons, plus significatifs ici que des écarts en dB. À titre de comparaison également, nous traçons aussi la borne de Chapman-Robbins, proposée dans [5], que nous moyennons par rapport à la loi *a priori* pour que la comparaison ait du sens. Ces résultats sont donnés sur la figure 1.

On peut d'une part remarquer que la borne proposée est plus précise que la borne de Chapman-Robbins, et permet donc de mieux rendre compte du comportement de l'estimation. Néanmoins, on note aussi qu'un écart subsiste, même à fort rapport signal à bruit, entre l'EQM du MAP et la borne. Comme mentionné plus haut dans la Section 1, ceci est dû au fait que le paramètre estimé t est discret, et que par conséquent ses estimateurs ne peuvent recouvrir les mêmes notions d'efficacité que pour des paramètres continus [4]. En particulier, dans le cas du MAP, il n'atteint pas la borne.

L'intérêt d'une telle borne reste néanmoins de fournir une référence commune pour une famille d'estimateurs (en l'occurrence les estimateurs bayésiens), permettant ainsi de les comparer objectivement. De plus, elle est également un moyen de

connaître les performances d'estimation réalisables pour ce problème, à un coût calculatoire réduit comparé à de nombreuses simulations de Monte-Carlo.

Références

- [1] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes, Theory and Application*. Englewood Cliffs, NJ : Prentice-Hall, 1993.
- [2] D. G. Chapman and H. Robbins. Minimum variance estimation without regularity assumptions. *The Annals of Mathematical Statistics*, 22(4) :581–586, Dec. 1951.
- [3] J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis*. Birkhäuser Basel, 2000.
- [4] C. Choirat and R. Seri. Estimation in discrete parameter models. *Statistical Science*, 27(2) :278–293, 2012.
- [5] A. Ferrari and J. Tournet. Barankin lower bound for change points in independent sequences. In *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, pages 557–560, St. Louis, MO, USA, Sept. 2003.
- [6] S. B. Fotopoulos and S. K. Jandhyala. Maximum likelihood estimation of a change-point for exponentially distributed random variables. *ELSEVIER Statistics and Probability Letters*, 51 :423–429, 2001.
- [7] 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*, 4(2) :1081–1104, Nov. 2010.
- [8] D. V. Hinkley. Inference about the change-point in a sequence of random variables. *Biometrika*, 57(1) :1–18, 1970.
- [9] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer Texts in Statistics. Springer, New-York, NY, USA, 2nd edition, Sept. 2003.
- [10] 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)*, pages 403–406, Portland, OR, USA, Sept. 1998.
- [11] H. L. Van Trees and K. L. Bell, editors. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley/IEEE Press, New-York, NY, USA, Sept. 2007.
- [12] E. Weinstein and A. J. Weiss. A general class of lower bounds in parameter estimation. *IEEE Transactions on Information Theory*, 34(2) :338–342, Mar. 1988.
- [13] W. Xu. *Performance bounds on matched-field methods for source localization and estimation of ocean environmental parameters*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2001.