# BAYESIAN QUICKEST DETECTION WITH UNKNOWN POST-CHANGE PARAMETER

Jun Geng<sup>1</sup> and Lifeng Lai<sup>2</sup>

<sup>1</sup> School of Elec. and Info. Engr., Harbin Inst. of Tech., Harbin, China, jgeng@hit.edu.cn <sup>2</sup> Dept. of Elec. and Comp. Engr., Worcester Poly. Inst., Worcester, MA, llai@wpi.edu

# ABSTRACT

In this paper, Bayesian quickest change-point detection problem with incomplete post-change information is considered. In particular, the observer knows that the post-change distribution belongs to a parametric distribution family, but he does not know the true value of the post-change parameter. Two problem formulations are considered in this paper. In the first formulation, we assume no additional prior information about the post-change parameter. In this case, the observer aims to design a detection algorithm to minimize the average (over the change-point) detection delay for all possible post-change parameters simultaneously subject to a worst case false alarm constraint. In the second formulation, we assume that there is a prior distribution on the possible value of the unknown parameter. For this case, we propose another formulation that minimizes the average (over both the change-point and the post-change parameter) detection delay subject to an average false alarm constraint. We propose a noval algorithm, which is termed as M-Shiryaev procedure, and show that the proposed algorithm is first order asymptotically optimal for both formulations considered in this paper.

*Index Terms*— Bayesian quickest change detection; M-Shiryaev procedure; sequential detection; unknown post-change parameter

### 1. INTRODUCTION

Quickest change-point detection problem aims to detect the abrupt change in the probability distribution of a random sequence as quickly and reliably as possible [1, 2, 3]. This technique has found a lot of applications in wireless sensor networks for network intrusion detection [4], seismic sensing, structural health monitoring, etc. In classic quickest changepoint detection problem, it assumes that both pre-change and post-change distributions are known by the observer. In practice, the pre-change distribution is likely to be known by the observer as he can collect a large amount of data to make an accurate estimation of the pre-change distribution. However, the post-change distribution is often unknown or known only to belong to a parametric distribution family. Hence, to explore the quickest detection algorithm for the problem with incomplete post-change information is of practical interest.

In this paper, we consider the Bayesian quickest detection problem with an unknown post-change parameter. In partic-

ular, the observer sequentially obtains a sequence of random observations whose distribution changes at an unknown time. The observer knows the pre-change distribution completely but the post-change distribution incompletely. Specifically, we assume that the post-change distribution belongs to a parametric distribution family, but the true post-change parameter, which comes from a known finite set  $\Xi$ , is unknown to the observer. The goal of the observer is to design an online algorithm to quickly detect the change in the observation sequence. Two problem formulations are considered in this paper. If there is no prior distribution over  $\Xi$ , then we aim to design an algorithm that minimizes the average (only over change-point) detection delay simultaneously for all possible post-change parameters subject to a worst case false alarm constraint. On the other hand, if the prior distribution of  $\Xi$ is available, then we aim to design an algorithm to minimize the average (over both post-change parameter and the changepoint) detection delay subject to an average false alarm constraint. In this paper, we also propose a new multi-chart detection algorithm termed as M-Shiryaev procedure. In the proposed algorithm, for each possible post-change parameter, the observer runs a Shiryaev detection procedure; hence the observer runs multiple Shiryaev procedures simultaneously. The observer declares that a change has occurred when any one of these parallelled procedures stops. We show that this proposed algorithm is first order asymptotically optimal for both formulations.

There have been several works on the quickest changepoint detection problems that take the unknown post-change parameter into consideration. To authors best knowledge, all these existing works are in the context of non-Bayesian quickest detection [2, 5, 6, 7, 8, 9]. In particular, [2, 5, 8] show that the generalized likelihood ratio (GLR) based CUSUM is asymptotically optimal over all post-change parameters. [7] adopted a shrinkage estimator to estimate the unknown postchange parameter. [9] considered a distributed sensor network setup. One may refer to a recent book [10] for more detailed results of this topic.

The remainder of the paper is organized as follows. The mathematical model is given in Section 2. Section 3 proposes the M-Shiryaev procedure and shows its asymptotic optimality. Numerical examples are given in Section 4. Section 5 offers concluding remarks.

# 2. PROBLEM FORMULATION

We consider a random observation sequence  $\{X_k, k = 1, 2, ...\}$ whose distribution changes at an unknown time t. Before

The work of J. Geng was supported by the Fundamental Research Funds for the Central Universities under grant AUGA5710013915. The work of L. Lai was supported by the National Science Foundation under grants DMS-12-65663 and CNS-14-57076.

the change-point  $t, X_1, X_2, \ldots, X_{t-1}$  are independent and identically distributed (i.i.d.) with probability density function (pdf)  $f(x;\xi_0)$ ; after the change-point t, the density of  $X_t, X_{t+1}, \ldots$  changes to  $f(x;\xi)$ . In the Bayesian quickest detection, the change-point t is modeled as a geometric random variable with parameter  $\rho$ , i.e., for  $0 < \rho < 1$ ,

$$P(t=k) = \rho(1-\rho)^{k-1}, \quad k=1,2,\dots$$
 (1)

Observations  $X_k$ 's generate the filtration  $\{\mathcal{F}_k\}_{k\in\mathbb{N}}$  with

$$\mathcal{F}_k = \sigma(\{t = 0\}, X_1, \cdots, X_k), \quad k = 1, 2, \dots,$$

and  $\mathcal{F}_0$  contains the sample space  $\Omega$ .

In the classic Bayesian quickest detection problem, both the pre-change distribution and the post-change distribution are perfectly known by the observer. However, in this paper, we consider the case that the observer only knows partial information of the post-change distribution. Specifically, the post-change distribution  $f(x;\xi)$  contains an *unknown* parameter  $\xi$ . The observer knows that  $\xi$  is taken from a finite set  $\Xi = \{\xi_1, \xi_2, \ldots, \xi_M\}$  but he does not know the true value of  $\xi$ . The pre-change distribution  $f(x;\xi_0)$  is perfectly known by the observer, i.e.,  $\xi_0$  is a known parameter, and  $\xi_0 \notin \Xi$ . In this paper,  $f(x;\xi)$  and  $f_{\xi}(x)$  are used interchangeably.

To facilitate the presentation, we denote  $P_{k,\xi_i}$  as the conditional probability measure of the observation sequence given  $\{t = k; \xi = \xi_i\}$ . For a measurable event F, we define probability measure  $P_{\pi,\xi_i}$  as

$$P_{\pi,\xi_i}(F) := \sum_{k=1}^{\infty} P_{k,\xi_i}(F) P(t=k).$$

We use  $\mathbb{E}_{k,\xi_i}$  and  $\mathbb{E}_{\pi,\xi_i}$  to denote the expectations with respect to probability measures  $P_{k,\xi_i}$  and  $P_{\pi,\xi_i}$ , respectively.

The observer aims to detect the change as quickly and accurately as possible. Let  $\mathcal{T}$  be the set of all finite stopping times with respect to filtration  $\{\mathcal{F}_k\}$ . The observer wants to find a stopping time  $\tau \in \mathcal{T}$ , at which the observer declares the change, to minimize the average detection delay (ADD) subject to certain false alarm constraints. Based on the availability of the prior information on  $\Xi$ , we consider two problem formulations in this paper. Specifically, if the observer has no prior information on  $\Xi$ , we consider the following problem setup

(P1) 
$$\inf_{\tau \in \mathcal{T}} \mathbb{E}_{\pi,\xi}[(\tau - t)^+] \quad \text{for all } \xi \in \Xi,$$
  
subject to 
$$\sup_{\xi \in \Xi} P_{\pi,\xi}(\tau < t) \le \alpha.$$
(2)

That is, the observer aims to find a stopping time  $\tau$  that minimizes ADDs simultaneously for all possible post-change parameters. In the mean while,  $\tau$  should achieve a small probability of false alarm (PFA) for all possible post-change parameters. In general, there is no optimal solution for this multiobjective optimization problem. However, in the sequel, we will propose an algorithm that is simultaneously asymptotically optimal for all  $\xi \in \Xi$ .

The second problem formulation considers the case that the observer knows the prior distribution of  $\Xi$ . In particular, the prior distribution is denoted as

$$\varpi_i = P(\xi = \xi_i),\tag{3}$$

with  $0 < \varpi_i < 1$  for all i = 1, ..., M. We further assume that  $\xi$  is independent of t. For measurable event F, we define probability measure  $P_{\pi,\varpi}$  as

$$P_{\pi,\varpi}(F) := \sum_{i=1}^{M} P_{\pi,\xi_i}(F) P(\xi = \xi_i),$$

and denote  $\mathbb{E}_{\pi,\varpi}$  as the expectations with respect to  $P_{\pi,\varpi}$ . The following formulation is of interest:

(P2) 
$$\inf_{\tau \in \mathcal{T}} \mathbb{E}_{\pi,\varpi}[(\tau - t)^+],$$
  
subject to  $P_{\pi,\varpi}(\tau < t) \le \alpha.$  (4)

We note that both ADD and PFA in (P2) are measured under  $P_{\pi,\varpi}$ , which takes average over both the change-point and the post-change parameter.

# 3. THE M-SHIRYAEV PROCEDURE AND ITS ASYMPTOTIC OPTIMALITY

#### 3.1. M-Shiryaev Procedure

For 
$$i = 1, \ldots, M$$
, let

$$R_{\rho,n,i} := \sum_{k=1}^{n} \prod_{j=k}^{n} \frac{1}{1-\rho} \frac{f_{\xi_i}(X_j)}{f_{\xi_0}(X_j)}$$
(5)

be the detection statistic, we propose the following multichart detection procedure:

$$\tau_{S,i} = \inf\{n \ge 1 | \log R_{\rho,n,i} > \log B_i\},$$
(6)

$$\tau_{MS} = \min \tau_{S,i}.\tag{7}$$

In the proposed algorithm, the observer updates M statistics  $R_{\rho,n,i}$  for  $i = 1, \ldots, M$  at each time slot, and each statistic is compared with its own threshold  $B_i$ . The procedure stops when either of statistics exceeds its threshold. Since  $R_{\rho,n,i}$  is the statistic used in the Shiryaev procedure, we term this proposed procedure as M-Shiryaev procedure.

Shiryaev's statistic is related to the posterior probability. For (P1), we define the posterior probability as

$$\pi_{i,n} := P_{\pi,\xi_i} (t \le n | \mathcal{F}_n), \quad i = 1, \dots, M.$$

it is easy to verify

$$\pi_{i,n} = \frac{\varrho_{i,n}(X_1, \dots, X_n)}{\varrho_{0,n}(X_1, \dots, X_n) + \varrho_{i,n}(X_1, \dots, X_n)},$$
(8)

where

$$\varrho_{0,n}(X_1,\ldots,X_n) = (1-\rho)^n \prod_{j=1}^n f_{\xi_0}(X_j),$$
  
$$\varrho_{i,n}(X_1,\ldots,X_n) = \rho \sum_{k=1}^n (1-\rho)^{k-1} \prod_{j=1}^{k-1} f_{\xi_0}(X_j) \prod_{j=k}^n f_{\xi_i}(X_j).$$

Then, it is easy to verify that

$$\Lambda_{n,i}^{(1)} := \log \frac{\pi_{i,n}}{1 - \pi_{i,n}} = \log \rho + \log R_{\rho,n,i}.$$
 (9)

For (P2), with a little abuse of notations, we define the posterior probability as

$$\pi_{i,k} := P_{\pi,\varpi}(t \le k, \xi = \xi_i | \mathcal{F}_k), \quad i = 1, \dots, M,$$
  
$$\pi_{0,k} := P_{\pi,\varpi}(t > k | \mathcal{F}_k) = 1 - \sum_{i=1}^M \pi_{i,k}.$$

It is easy to verify that

$$\pi_{i,n} = \frac{\varrho_{i,n}(X_1, \dots, X_n)}{\sum_{j=0}^{M} \varrho_{j,n}(X_1, \dots, X_n)},$$
(10)

where

$$\varrho_{0,n}(X_1,\ldots,X_n) = (1-\rho)^n \prod_{j=1}^n f_{\xi_0}(X_j),$$

$$\varrho_{i,n}(X_1,\ldots,X_n) = \rho \varpi_i \sum_{k=1}^n (1-\rho)^{k-1} \prod_{j=1}^{k-1} f_{\xi_0}(X_j) \prod_{j=k}^n f_{\xi_i}(X_j).$$

In this case, we have

$$\Lambda_{n,i}^{(2)} := \log \frac{\pi_{i,n}}{\pi_{0,n}} = \log \varpi_i \rho + \log R_{\rho,n,i}.$$
 (11)

Hence,  $\tau_{S,i}$  in (6) can be equivalently expressed in terms of  $\Lambda_{n,i}^{(1)}$  and  $\Lambda_{n,i}^{(2)}$  for (P1) and (P2), respectively.

### 3.2. Asymptotic optimality

We first present asymptotic lower bounds of detection delays for all possible post-change parameters as the worst case PFA or the average PFA vanishes. In particular, we have the following result:

**Lemma 3.1.** As  $\alpha \rightarrow 0$ ,

$$\inf_{\tau \in \mathcal{T}} \{ \mathbb{E}_{\pi,\xi} [(\tau - t)^+] : \sup_{\xi \in \Xi} P_{\pi,\xi} (\tau < t) \le \alpha \} \\
\ge \frac{|\log \alpha|}{D(f_{\xi} || f_{\xi_0}) + |\log(1 - \rho)|} (1 + o(1)), \quad (12)$$

and

$$\inf_{\tau \in \mathcal{T}} \{ \mathbb{E}_{\pi, \xi} [(\tau - t)^+] : P_{\pi, \varpi}(\tau < t) \le \alpha \} \\
\geq \frac{|\log \alpha|}{D(f_{\xi} || f_{\xi_0}) + |\log(1 - \rho)|} (1 + o(1)), \quad (13)$$

where  $D(f_{\xi}||f_{\xi_0})$  is the KL divergence between  $f(x;\xi)$  and  $f(x;\xi_0)$ .

*Proof.* (12) can be proved as follows:

$$\inf_{\tau \in \mathcal{T}} \{ \mathbb{E}_{\pi,\xi_i} [(\tau - t)^+] : \sup_{\xi \in \Xi} P_{\pi,\xi}(\tau < t) \le \alpha \}$$

$$\geq \inf_{\tau \in \mathcal{T}} \{ \mathbb{E}_{\pi,\xi_i} [(\tau - t)^+] : P_{\pi,\xi_i}(\tau < t) \le \alpha \}$$

$$\geq \frac{|\log \alpha|}{D(f_{\xi_i}||f_0) + |\log(1 - \rho)|} (1 + o(1)). \tag{14}$$

The last inequality is from the lower bound of the average detection delay in the classic Bayesian quickest detection (see Theorem 1 in [11]).

(13) can be proved similarly. In particular, the condition  $P_{\pi,\varpi}(\tau < t) \leq \alpha$  can be relaxed to  $\varpi_i P_{\pi,\xi_i}(\tau < t) \leq \alpha$ , then (13) follow the facts that  $\alpha/\varpi_i \to 0$  for all  $i = 1, \ldots, M$ .

The achieveability of the M-Shiryaev procedure is presented in Theorem 3.2 and Theorem 3.4.

**Theorem 3.2.** For (P1), the M-Shiryaev procedure defined in (7) is asymptotically optimal as  $\alpha \to 0$  by setting  $B_1 = \ldots = B_M = M(\rho\alpha)^{-1}$ .

$$\mathbb{E}_{\pi,\xi_i}\left[\left(\tau_{MS} - t\right)^+\right] \le \frac{|\log \alpha|}{D(f_{\xi_i}||f_{\xi_0}) + |\log(1-\rho)|} (1+o(1)).$$
(15)

*Proof.* Since  $\tau_{S,i}$  is asymptotically optimal for the Bayesian quickest detection when  $f_{\xi_i}(x)$  is the true post-change distribution, we have

$$\mathbb{E}_{\pi,\xi_i}\left[(\tau_{S,i}-t)^+\right] = \frac{|\log B|}{D(f_{\xi_i}||f_{\xi_0}) + |\log(1-\rho)|}(1+o(1)).$$

Since  $\tau_{MS} \leq \tau_{S,i}$  and  $|\log B| = |\log \alpha|(1+o(1))$ , we know that  $\tau_{MS}$  satisfies (15).

We then show that  $\tau_{MS}$  satisfies the false alarm constraint. By setting  $B_i = M(\rho \alpha)^{-1}$  and using (9), (6) can be equivalently written as

$$\tau_{S,i} = \inf\left\{n \ge 1 \left| \pi_{i,n} > 1 - \left(1 + \frac{\alpha}{M}\right)^{-1} \frac{\alpha}{M}\right\}.$$
 (16)

Hence

$$P_{\pi,\xi_i}(\tau_{S,i} < t) = \mathbb{E}_{\pi,\xi_i}[1 - \pi_{i,\tau_{S,i}}] \le \alpha/M.$$

Further,  $P_{\pi,\xi_i}(\tau_{MS} < t; \tau_{MS} = \tau_{S,i}) \le \alpha/M$ . We note that  $P_{\pi,\xi_j}(\tau_{MS} < t; \tau_{MS} = \tau_{S,i}) = P_{\pi,\xi_i}(\tau_{MS} < t; \tau_{MS} = \tau_{S,i}).$ 

This is true because all observations are generated from  $f_{\xi_0}(x)$ , regardless of the true post-change parameter, on the event  $\{\tau_{MS} < t\}$ . As a result, we have

$$P_{\pi,\xi_j}(\tau_{MS} < t) = \sum_{i=1}^M P_{\pi,\xi_j}(\tau_{MS} < t; \tau_{MS} = \tau_{S,i}) \le \alpha.$$
(17)

Since  $\xi_j$  in (17) is arbitrarily selected in  $\Xi$ , then we have  $\sup_{\xi \in \Xi} P_{\pi,\xi}(\tau_{MS} < t) \le \alpha$ . That ends the proof.

To show the proposed algorithm is asymptotically optimal for (P2), we first present the following lemma:

**Lemma 3.3.** (Lemma 2.3 in [12]) Let  $\tau$  be a stopping time with respect to  $\{\mathcal{F}_k\}$ . Let F be an  $\mathcal{F}_{\tau}$  measurable event, we have

$$P_{\pi,\varpi}(F \cap \{\tau < t\}) = \varpi_i \mathbb{E}_{\pi,\xi_i} \left[ 1_{F \cap \{t \le \tau < \infty\}} e^{-\Lambda_{\tau,i}^{(2)}} \right].$$
(18)

*Proof.* To make the paper self-contain, we rewrite the proof as follows:

$$P_{\pi,\varpi}(F \cap \{\tau < t\}) = \sum_{n=0}^{\infty} \mathbb{E}_{\pi,\varpi} \left[ \mathbf{1}_{F \cap \{\tau=n;n < t\}} \right]$$
$$= \sum_{n=0}^{\infty} \mathbb{E}_{\pi,\varpi} \left[ \mathbf{1}_{F \cap \{\tau=n\}} \pi_{0,n} \right] = \sum_{n=0}^{\infty} \mathbb{E}_{\pi,\varpi} \left[ \mathbf{1}_{F \cap \{\tau=n\}} \pi_{i,n} \frac{\pi_{0,n}}{\pi_{i,n}} \right]$$
$$= \sum_{n=0}^{\infty} \mathbb{E}_{\pi,\varpi} \left[ \mathbf{1}_{F \cap \{\tau=n,t \le n;\xi=\xi_i\}} \frac{\pi_{0,n}}{\pi_{i,n}} \right] = \varpi_i \mathbb{E}_{\pi,\xi_i} \left[ \mathbf{1}_{F \cap \{t \le \tau < \infty\}} \frac{\pi_{0,\tau}}{\pi_{i,\tau}} \right]$$

**Theorem 3.4.** For (P2), the M-Shiryaev procedure is asymptotically optimal as  $\alpha \to 0$  by setting  $B_i = (\rho \alpha \varpi_i)^{-1}$ .

$$\mathbb{E}_{\pi,\xi_i}\left[\left(\tau_{MS} - t\right)^+\right] \le \frac{|\log \alpha|}{D(f_{\xi_i}||f_{\xi_0}) + |\log(1-\rho)|} (1 + o(1)).$$
(19)

*Proof.* (19) can be shown by using the same argument of proving (15). We then show that  $\tau_{MS}$  satisfies the false alarm constraint. By setting  $B_i = (\rho \alpha \varpi_i)^{-1}$  and using (11), (6) can be written equivalently as

$$\tau_{S,i} = \inf\left\{n \ge 1 |\Lambda_{n,i}^{(2)} > -\log\alpha\right\}.$$
(20)

Hence we have

$$P_{\pi,\varpi}(\tau_{MS} < t; \tau = \tau_{S,i})$$

$$\stackrel{(a)}{=} \varpi_i \mathbb{E}_{\pi,\xi_i} \left[ \mathbbm{1}_{\{\tau_{MS} = \tau_{S,i}\} \cap \{t \le \tau_{MS} < \infty\}} e^{-\Lambda_{\tau_{MS},i}^{(2)}} \right]$$

$$\leq \varpi_i \mathbb{E}_{\pi,\xi_i} \left[ \mathbbm{1}_{\{\tau_{MS} = \tau_{S,i}\} \cap \{t \le \tau < \infty\}} \alpha \right] \le \alpha \varpi_i, \quad (21)$$

where (a) is because of Lemma 3.3. Hence,  $P_{\pi,\varpi}(\tau_{MS} \le t) = \sum_{i=1}^{M} P_{\pi,\varpi}(\tau_{MS} < t; \tau_{MS} = \tau_i) \le \alpha$ . From (13) in Lemma 3.1 and (19), we conclude that the

From (13) in Lemma 3.1 and (19), we conclude that the proposed M-Shiryaev procedure is asymptotically optimal for (P2) since  $\tau_{MS}$  achieves the lower bound of detection delay for every individual post-change parameter  $\xi \in \Xi$ .

### 4. NUMERICAL SIMULATION

In this section, we provide two numerical examples to illustrate the results obtained in this paper. In these numerical examples, we assume that the pre-change distribution  $f_{\xi_0}$  is  $\mathcal{N}(0, 1)$  and the post-change distribution  $f_{\xi}$  is  $\mathcal{N}(0, \xi^2)$ , where  $\xi$  takes value in  $\Xi = \{1.5, 2, 2.5\}$ . The geometrically distributed change-point has parameter  $\rho = 0.01$ .

We first illustrate the performance of the M-Shiryaev procedure for (P1). The simulation result is shown in Figure 1. In this simulation,  $\xi = 1.5$  is set to be the true post-change parameter. The black dot line is the theoretical asymptotic lower bound calculated by (12). The blue solid line is the performance of the M-Shiryaev procedure. We can see that ADD scales linearly with respect to  $|\log \alpha|$ . Moreover, the performance of the M-Shiryaev procedure is parallel to the lower bound, which indicates that the M-Shiryaev procedure is asymptotically optimal since the constant difference is negligible when the detection delay goes to infinity. Actually, we can obtain similar simulation results when we change the post-change parameter to  $\xi = 2$  and  $\xi = 2.5$ . This confirms our analysis that the proposed M-Shiryaev procedure is simultaneously optimal for all possible post-change parameters.



Fig. 1. ADD vs. PFA under  $\xi = 1.5$ 

In the second simulation, we examine the asymptotic optimality of the M-Shiryaev procedure for (P2). In the simulation, the prior distribution on  $\Xi$  is set to be  $\{0.1, 0.5, 0.4\}$ . The simulation result is shown in Figure 2. In the figure, the black dot line is the theoretical lower bound of  $\mathbb{E}_{\pi,\varpi}[(\tau-t)^+]$ , which is the average of the lower bounds presented in (13). The blue solid line is the performance of the M-Shiryaev procedure. We can see that the performance of the M-Shiryaev procedure is parallel to the lower bound, which indicates that the proposed procedure is asymptotically optimal for (P2).



Fig. 2. ADD vs. PFA under (P2)

# 5. CONCLUSION

In this paper, we have studied the Bayesian quickest detection problem with an unknown post-change parameter. We have proposed the M-Shiryaev procedure, and have shown the proposed procedure is first order asymptotically optimal for both formulations studied in this paper.

# 6. REFERENCES

- A. N. Shiryaev, "On optimal methods in quickest detection problems," *Theory of Probability and Its Applications*, vol. 8, pp. 22–46, 1963.
- [2] G. Lorden, "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, 1971.
- [3] M. Pollak, "Optimal detection of a change in distribution," *Annals of Statistics*, vol. 13, no. 1, pp. 206–227, Mar. 1985.
- [4] K. Premkumar and A. Kumar, "Optimal sleep-wake scheduling for quickest intrusion detection using sensor network," in *Proc. IEEE Conf. Computer Communications (Infocom)*, Phoenix, AZ, USA, Apr. 2008, pp. 1400–1408.
- [5] V. P. Dragalin, "Optimality of the generalized CUSUM procedure in quickest detection problem," *Proc. Steklov Institute of Mathematics*, vol. 202, pp. 107–119, Nov. 1994.
- [6] G. Lorden and M. Pollak, "Nonanticipating estimation applied to sequential analysis and changepoint detection," *Annals of Statistics*, vol. 33, no. 3, pp. 1422–1454, 2005.
- [7] Y. Wang and Y. Mei, "Montoring multiple data streams via shrinkage post-change estimation," *Annals of Statistics*, 2013, submitted.
- [8] T. L. Lai, "Information bounds and quickest detection of parameter changes in stochastic systems," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2917–2929, 1998.
- [9] A. G. Tartakovsky and A. S. Polunchenko, "Quickest changepoint detection in distributed multisensor systems under unknown parameters," in *Proc. Intl. Conf.* on Information Fusion, Cologne, Germany, Jul. 2008, pp. 878–885.
- [10] Alexander Tartakovsky, Igor Nikiforov, and Michele Basseville, Sequential Analysis: Hypothesis Testing and Changepoint Detection, CRC Press, New York, 2014.
- [11] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic Bayesian theory of quickest change detection," *Theory of Probability and Its Applications*, vol. 49, no. 3, pp. 458–497, 2005.
- [12] S. Dayanik, W. B. Powell, and K. Yamazaki, "Asymptotically optimal bayesian sequential change detection and identification rules," *Annals of Operations Research*, vol. 208, no. 1, pp. 337–370, 2013.