Quickest Search for a Change Point

Javad Heydari
LG Electronics USA
Santa Clara, CA 95054
Email: javad.heydari@ieee.org

Ali Tajer
Rensselaer Polytechnic Institute
Troy, NY 12180
Email: tajer@ecse.rpi.edu

***Abstract*—This paper considers a sequence of random variables that undergo periods of transient changes at an unknown set of time instants, referred to as *transient* change-points. The objective is to constantly monitor the sequence in order to detect *one* of the change-points subject to a hard constraint on the detection delay, while in parallel, the rate of false alarms is controlled. This setting is fundamentally different from the conventional change-point detection problems, in which there exists at most *one* change-point that can be either persistent or transient. In this paper, the exact optimal decision rules are characterized. Furthermore, it is shown that in the special case that the objective is detecting a transient change-point at exactly the instant that a change occurs (i.e., no detection delay), the test reduces to the well-known Shewhart test. Numerical evaluations are also provided to assess the performance of the decision rules.**

***Index Terms*—Change-point detection, quickest detection, transient changes.**

I. Introduction

Real-time monitoring of a system or process for detecting a change in its behavior arises in many application domains such as detecting faults or security breaches in networks, and searching for under-utilized spectrum bands for opportunistic spectrum access. It is often of interest to detect changes with minimal delay after they occur. At the same time, detection rules that are exceedingly sensitive to fluctuations in the data are susceptible to raising frequent false alarms. This creates an inherent tension between the quickness and the reliability of the decisions.

The classical change-point detection problems, generally, focus on detecting a *permanent* change in the statistical model of a given sequence of random variables. Specifically, in such problems, a random sequence is generated according to a nominal distribution, which at an unknown time permanently changes to a different distribution [1]. In such problems, a decision-maker designs a stopping rule to detect the change with a minimal delay after the change occurs, while in parallel, controlling the rate of false alarms. The setting and objective of this paper has major distinctions from the classical quickest change-point detection. First, the change is not persistent, i.e., after a period of time following a change-point, the distribution returns to the nominal model. Secondly, the change-points are not unique, and a sequence can potentially undergo multiple transient change-points. Finally, our objective is to search and identify a change-point within a *hard* deadline after it occurs, while in the classical problems the focus is on *soft* deadline, i.e., the *average* detection delay. The advantage of enforcing a hard deadline on the delay is that it prevents arbitrarily large delays, which is not the case when enforcing soft deadlines. This vision is in line with the settings in [2]–[5], with the distinction that in this paper multiple change-points can occur and they are transient. This problem is, for instance, motivated by application in spectrum sensing, in which we are interested in quickly identifying free spectrum bands that become available for only short periods.

Quickest detection of transient changes in a sequence has gained recent interest. For instance, the study in [6] aims to characterize the shortest duration of a transient change that can be detected reliably as the false alarm rate approaches zero, and [7] aims to minimize the rate of missing transient changes. Besided the objectives, the main distinction of our setting with those of [6] and [7] is that they consider only one transient change, while we consider having an unknown number of transient change-points. The studies in [8]–[11] consider a setting in which only one permanent change occurs in the sequence, but the change does not occur abruptly, but it rather does through a series of changes, after which it settles to a permanent steady state. In this setting, the steady-state distribution is different from the pre-change one. In [8] the transient duration is a single sample, while in [9] it is a deterministic unknown constant. Quickest change-point detection under multiple transient changes is also considered in [10] and [11], in which the state of the system is assumed to be a Markov process and only one of the states, which is an absorbing state, is considered as the desirable change state.

Besides the distinction in the data model, the ultimate goal of this paper also differs from the classical settings. Specifically, instead of minimizing the average detection delay, the probability of stopping at or within a deadline after a change-point is maximized. This approach was first used in [2] in a Bayesian setting for detecting a persistent change in a sequence of independent and identically distributed (i.i.d.) random variables. The results were extended to dependent random variables in [12], and composite post-change models in [3]. In [4] and [5], the objective is detecting a persistent change immediately by using the first sample under the change state. Under both Bayesian and minimax regimes the exactly optimal detection rules have been characterized, and the results have been extended to independent non-identically distributed

This research was supported in part by the U. S. National Science Foundation under the CAREER Award ECCS-1554482 and the grant DMS-1737976.

samples and composite post-change models in [4], and to Markovian samples in [5].

The remainder of the paper is organized as follows. Section II provides the data model and formalizes the search problem of interest. The quickest search rule for detecting a change immediately is characterized in Section III, where the exact decision rules are established. The results of Section III are extended to the setting in which a change should be detected within a hard deadline after it occurs in Section IV. Section V provides the numerical evaluation of the quickest search approach, and concluding remarks are provided in Section VI.

II. Problem Statement

A. Data Model

Consider a sequence of random variables denoted by $\mathcal{X} \triangleq \{X_t : t \in \mathbb{N}\}$. As shown in Fig. 1, these random variables have a common nominal probability distribution that undergoes periods of *transient* changes at an unknown and non-random set of time instants. Specifically, the elements of $\mathcal{X}$ are *nominally* generated according to a probabilistic distribution with the cumulative density function (cdf) F_0. There potentially exist a finite but unknown number of time instants $\gamma \triangleq \{\gamma_i : i \in \{1, \dots, s\}\}$, referred to as transient change-points, at which the distribution changes from the nominal cdf F_0 to a distinct one with cdf F_1. It is assumed that the number of change-points $|\gamma| = s \in \mathbb{N}$ is unknown, and the duration of each transient change is a known constant denoted by T. The transient intervals are assumed to be non-overlapping, i.e., $|\gamma_i - \gamma_j| > T$, for all distinct $i, j \in \{1, \dots, s\}$. We define $\mathcal{S}$ as the set of all time instants $t \in \mathbb{N}$ at which X_t is generated by F_1, i.e.,

$$\mathcal{S} \triangleq \{t \ : \ X_t \sim F_1\} \ . \tag{1}$$

Hence, for the elements of $\mathcal{X}$ we have the following dichotomous model

$$\begin{array}{ll} X_t \sim F_0 \ , & t \in \mathbb{N} \backslash \mathcal{S} \\ X_t \sim F_1 \ , & t \in \mathcal{S} \end{array} \ . \tag{2}$$

We also assume that there exist well-defined probability density functions (pdfs) corresponding to F_0 and F_1, which we denote by f_0 and f_1, respectively. Subsequently, we denote the probability measure governing sequence $\mathcal{X}$ and the expectation with respect to this measure by $\mathbb{P}_\gamma$ and $\mathbb{E}_\gamma$, respectively. We also use $\mathbb{P}_\infty$ and $\mathbb{E}_\infty$ for the case that no change occurs in the data under consideration, i.e., $s = 0$.

B. Problem Formulation

The objective is to sequentially collect samples from sequence $\mathcal{X}$ and design a sequential decision rule for the quickest detection and identification of one of the transient changes with a delay not exceeding $\xi \in \{1, \dots, T\}$, while the rate of false alarms is controlled. Hence, the sequential decision-making process continually collects samples until the stopping time of the process, at which point it is confident enough that a change has occured within the past ξ samples. It is noteworthy that the setting in which there exists only one change-point

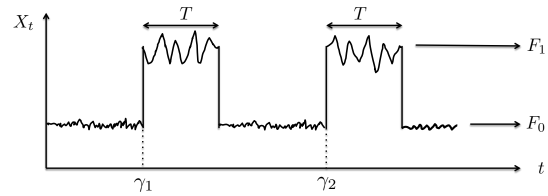

Fig. 1: Data model.

(i.e., $s = 1$), which can be either persistent or transient, is studied extensively in the literature (c.f. [1]–[8]). In contrast, in this paper we assume that the number of change-points s is unknown and can exceed one.

The information generated by the data sequentially up to time t generates the filtration $\{\mathcal{F}_t : t \in \mathbb{N}\}$, where

$$\mathcal{F}_t \triangleq \sigma\big(X_1, \dots, X_t\big) \ . \tag{3}$$

We define a coarser filtration, which at time $t \in \mathbb{N}$ is generated by only the samples generated from the end of the preceding change period up to time t. This filtration is denoted by

$$\mathcal{G}_t \triangleq \sigma\big(X_{r(t)+T}, X_{r(t)+T+1}, \dots, X_t\big) \ , \tag{4}$$

where we have defined $r(t) \triangleq \sup \ \{\gamma_i \in \gamma : \gamma_i \leq t - T\}$, and adopt the convention that the supremum of an empty set is zero. The sequential sampling process continues until the stopping time, denoted by τ, after which no further samples are collected and a change is declared. The stopping time τ is set to be a $\mathcal{G}_t$-measurable function.

Two relevant performance measures for evaluating the quality of these sampling and decision-making processes are the quickness of the process as well as the frequency of false alarms. To account for the quickness, we are interested in the real-time detection of a change with a hard deadline after the change has occurred. Hence, the conventional average detection delay is ineffective as it does not impose a hard constraint on the detection delay and the delay can become arbitrarily large. To circumvent this, for quantifying the agility of the process we adopt a probability-based approach similar to [2] and [4]. Specifically, we investigate two minimax settings in which we consider probability maximization criteria mimicking Pollak's [13] and Lorden's [14] approaches. In particular, when the hard constraint on the delay is ξ samples, for some $\xi \in \{1, \dots, T\}$, we define a Pollak-like criterion as

$$\mathcal{L}_{\rm P}(\tau) \triangleq \inf_{\gamma} \sum_{\gamma_i \in \gamma} \sum_{k=1}^{\xi} \mathbb{P}_\gamma(\tau = \gamma_i + k - 1 \mid \tau \geq \gamma_i) \ . \tag{5}$$

Similarly, we define a Lorden-like worst case criterion as

$$\mathcal{L}_{\rm L}(\tau) \triangleq \tag{6}$$

$$\inf_{\gamma} \sum_{\gamma_i \in \gamma} \sum_{k=1}^{\xi} \operatorname*{essinf}_{\mathcal{F}_{\gamma_i - 1}} \mathbb{P}_\gamma(\tau = \gamma_i + k - 1 \mid \mathcal{F}_{\gamma_i - 1}, \tau \geq \gamma_i) \ .$$

It can be readily verified that

$$\mathcal{L}_{\rm L}(\tau) \leq \mathcal{L}_{\rm P}(\tau) \ . \tag{7}$$

In order to account for the frequency of the false alarms, we use $\mathbb{E}_\infty\{\tau\}$, which captures the average run length to a false alarm before the first change-point γ_1 occurs.

There exists an inherent tension between the rate of false alarms on the one hand, and the measures $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$, on the other hand, as improving these two measures penalizes the false alarm rate. An optimal sampling strategy can be obtained by striking a balance between the false alarm rate and the detection probability. Hence, under the Pollak-like and Lorden-like criteria in (5) and (6), respectively, the sampling strategy is the solution to

$$\begin{array}{ll} \sup_\tau & \mathcal{L}_{\rm P}(\tau) \\ \text{s.t.} & \mathbb{E}_\infty\{\tau\} \geq \eta \end{array} , \quad \text{and} \quad \begin{array}{ll} \sup_\tau & \mathcal{L}_{\rm L}(\tau) \\ \text{s.t.} & \mathbb{E}_\infty\{\tau\} \geq \eta \end{array} , \tag{8}$$

where $\eta \geq 1$ in both settings controls the false alarm rate. In the next section, we first focus on these two problems for $\xi = 1$, i.e., the setting in which we aim to identify one exact change-point. By leveraging the insight gained from this special case, we solve the problems in (8) in their general forms in Section IV.

III. Quickest Search Rules for $\xi = 1$

In this section, we characterize the optimal stopping rules for the problems in (8) when $\xi = 1$. For this purpose, we first find upper bounds on the objective functions $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$ in Section III-A. Then we briefly review the Shewhart test in Section III-B, and in Section III-C we show that by using the Shewhart test as the decision rule the values of $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$ achieve their upper bounds determined in Section III-A, thereby establishing that the Shewhart test is an optimal solution to (8).

A. Upper Bounds on the Objective Functions

In order to facilitate finding upper bounds on the objective functions in (5) and (6), we denote the likelihood ratio of the sample collected at time t by ℓ_t, i.e.,

$$\ell_t \triangleq \frac{f_1(X_t)}{f_0(X_t)} . \tag{9}$$

The following theorem characterizes an upper bound on both Pollak-like and Lorden-like criteria defined in (5) and (6), respectively.

Theorem 1 (Upper Bound): For the objective functions $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$, and for $s \in \mathbb{N}$, we have

$$\mathcal{L}_{\rm L}(\tau) \leq \mathcal{L}_{\rm P}(\tau) \leq s \cdot \frac{\mathbb{E}_\infty\{\ell_\tau\}}{\mathbb{E}_\infty\{\tau\}} . \tag{10}$$

B. Shewhart Test

The form of Shewhart test that we adopt in this paper consists in a dynamic and sequential likelihood ratio test. Formally, at each time t and based on the observation X_t we form the likelihood ratio value ℓ_t defined in (9). The Shewhart test compares ℓ_t with a pre-specified and deterministic threshold α and declares a change when ℓ_t exceeds α. Specifically, the stopping time of the Shewhart test is found via

$$\tau_{\rm s} \triangleq \inf\{t \ : \ \ell_t \geq \alpha\} . \tag{11}$$

The value of the threshold α is chosen such that the average run-length to a false alarm is guaranteed not to be smaller than η, and it can be computed by solving

$$\mathbb{P}_\infty(\ell_1 \geq \alpha) = \eta^{-1} . \tag{12}$$

C. Optimality of Shewhart Test

We prove the exact optimality of the Shewhart test formalized in (11) and (12) for problems in (8). For this purpose, we start by proving that corresponding to any feasible[1] decision rule with the stopping time ν and the associated ratio

$$\frac{\mathbb{E}_\infty\{\ell_\nu\}}{\mathbb{E}_\infty\{\nu\}} , \tag{13}$$

we can construct an alternative feasible decision rule that achieves the false alarm constraint with equality, and its stopping time, denoted by ν', achieves the same ratio. Specifically, corresponding to any feasible stopping time ν there exists ν' such that

$$\mathbb{E}\{\nu'\} = \eta \quad \text{and} \quad \frac{\mathbb{E}_\infty\{\ell_{\nu'}\}}{\mathbb{E}_\infty\{\nu'\}} = \frac{\mathbb{E}_\infty\{\ell_\nu\}}{\mathbb{E}_\infty\{\nu\}} . \tag{14}$$

This observation is formalized in the following lemma.

Lemma 1: Corresponding to any given feasible decision rule with the stopping time ν, there always exists a feasible decision rule that satisfies the false alarm constraint with equality, and its stopping time, denoted by ν', yields

$$\frac{\mathbb{E}_\infty\{\ell_{\nu'}\}}{\mathbb{E}_\infty\{\nu'\}} = \frac{\mathbb{E}_\infty\{\ell_\nu\}}{\mathbb{E}_\infty\{\nu\}} . \tag{15}$$

Next, we leverage the result of Lemma 1 and prove the following properties for the Shewhart test:

1) It is a feasible test.
2) It maximizes the upper bound on $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$ established in Theorem 1.
3) The objective functions $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$ meet this maximum upper bound when using the Shewhart test.

These properties are formalized in the following two lemmas.

Lemma 2 (Feasibility of Shewhart): Shewhart test achieves the false alarm constraints of (8) with equality.

In the next lemma, we establish that the Shewhart test maximizes the upper bound on the objective function.

Lemma 3: The Shewhart test is the solution to

$$\sup_\tau \frac{\mathbb{E}_\infty\{\ell_\tau\}}{\mathbb{E}_\infty\{\tau\}} , \quad \text{subject to} \ \ \mathbb{E}_\infty\{\tau\} = \eta . \tag{16}$$

The following theorem proves that for the Shewhart test, the objective function meets its upper bound and, consequently, the Shewhart test is an optimal solution to (8) for $\xi = 1$.

Theorem 2: The Shewhart test with the stopping time and threshold given in (11) and (12), respectively, is an optimal solution to both problems in (8), i.e.,

$$\mathcal{L}_{\rm L}(\tau_s) = \mathcal{L}_{\rm P}(\tau_s) = \sup_{\tau \,:\, \mathbb{E}_\infty\{\tau\} \geq \eta} \mathcal{L}_{\rm P}(\tau) . \tag{17}$$

[1]A decision rule with stopping time ν is called feasible if it satisfies the false alarm constraint, i.e., $\mathbb{E}_\infty\{\nu\} \geq \eta$.

Shewhart test is not only an optimal test, but it is also simple to implement; at each time t, it takes a new sample from the sequence, forms its likelihood ratio, and compares the likelihood ratio with a fixed pre-specified threshold. It stops the process and declares a change the first time the likelihood ratio exceeds the threshold.

IV. Quickest Search within a Window ($\xi > 1$)

In this section, we consider the general setting in which the goal is to detect a change within a window of ξ samples after its occurrence. For simplicity in notations, we present the detailed results for $\xi = 2$, and the generalization for $\xi > 2$ follows the same line of arguments. The following lemma provides an upper bound on the objective functions of (8) for $\xi = 2$.

Theorem 3 (General Upper Bound): The objective functions $\mathcal{L}_{\rm P}(\tau)$ and $\mathcal{L}_{\rm L}(\tau)$, defined in (5) and (6), respectively, satisfy

$$\mathcal{L}_{\rm L}(\tau) \leq \mathcal{L}_{\rm P}(\tau) \leq s \cdot \frac{\mathbb{E}_\infty\{\ell_\tau + \ell_{\tau-1}\ell_\tau\}}{\mathbb{E}_\infty\{\tau\}} \ . \qquad (18)$$

The results of the theorem above is intuitive: for stopping exactly at the change-point we already observed in Theorem 1 that the upper bound depends on the statistics of the last one sample. This lemma states that for $\xi = 2$ it depends on the statistics of the last two samples through the term $(\ell_\tau + \ell_{\tau-1}\ell_\tau)$. The next step is to find the stopping time that maximizes this upper bound, for which we later show that the objective function meets its upper bound. First, it can be readily verified that the upper bound in (18) can be achieved by a stopping time that meets the false alarm constraint with equality. Hence, for finding an optimal solution, we can equivalently solve

$$\sup_\tau \ \frac{\mathbb{E}_\infty\{\ell_\tau + \ell_{\tau-1}\ell_\tau\}}{\mathbb{E}_\infty\{\tau\}} \ , \quad \text{subject to } \mathbb{E}_\infty\{\tau\} = \eta \ . \qquad (19)$$

This problem, equivalently, can be converted to the following unconstrained problem,

$$\sup_\tau \ \ \mathbb{E}_\infty\{\ell_\tau + \ell_{\tau-1}\ell_\tau - \lambda\tau\} \ , \qquad (20)$$

where λ is a Lagrangian multiplier, which is chosen such that $\mathbb{E}_\infty\{\tau\} = \eta$. In order to solve this problem, we define a reward-to-go function at time t as

$$\tilde{G}_t(\mathcal{G}_t) \triangleq \max\left\{\ell_t + \ell_{t-1}\ell_t \ , \ -\lambda + \mathbb{E}_\infty\{\tilde{G}_{t+1}(\mathcal{G}_{t+1}) \mid \mathcal{G}_t\}\right\} \ ,$$

where $(\ell_t + \ell_{t-1}\ell_t)$ is the reward of stopping at time t while $(-\lambda + \mathbb{E}_\infty\{\tilde{G}_{t+1}(\mathcal{G}_{t+1}) \mid \mathcal{G}_t\})$ is the expected reward of collecting one more sample. The following lemma provides a sufficient statistic for calculating $\tilde{G}_t$.

Lemma 4 (Sufficient Statistic): The reward-to-go function $\tilde{G}_t$ is a function of $\mathcal{G}_t$ only through (ℓ_{t-1}, ℓ_t). Therefore, (ℓ_{t-1}, ℓ_t) is a sufficient statistic for calculating $\tilde{G}_t$.

Since the value of reward-to-go function only depends on the value of (ℓ_{t-1}, ℓ_t) and not t, we define the reward-to-go as a function of (ℓ_{t-1}, ℓ_t) by $G(\ell_{t-1}, \ell_t)$, i.e.,

$$G(\ell_{t-1}, \ell_t) \triangleq \max\{\ell_t + \ell_{t-1}\ell_t \ , \ -\lambda + \mathbb{E}_\infty\{G(\ell_t, \ell_{t+1}) \mid \ell_t\}\} \ . \qquad (21)$$

The stopping time that maximizes the objective function in (20) is the first time instant at which the reward of stopping, i.e., $(\ell_t + \ell_{t-1}\ell_t)$, exceeds the reward of collecting one more sample, i.e., $(-\lambda + \mathbb{E}_\infty\{G(\ell_t, \ell_{t+1}) \mid \ell_t\})$. To formalize this, we define

$$h(x) \triangleq \mathbb{E}_\infty\{G(x, y) \mid x\} \ , \qquad (22)$$

as the expected utility of collecting one more sample if we ignore the delay cost, and stop the first time instant at which

$$\ell_t + \ell_{t-1}\ell_t \geq -\lambda + h(\ell_t) \ . \qquad (23)$$

The following theorem establishes that such a stopping time optimizes the upper bound in (18).

Lemma 5: An optimal solution to the optimization problem in (19) is

$$\tau^* \triangleq \inf\{t \ : \ \ell_t + \ell_{t-1}\ell_t \geq -\lambda + h(\ell_t)\} \ , \qquad (24)$$

where $\lambda > 0$ is chosen such that $\mathbb{E}_\infty\{\tau^*\} = \eta$.

Since τ^* optimizes the upper bound on the objective function and is a feasible solution, it only remains to show that for τ^* the objective function is equal to the upper bound. The following theorem asserts that this property holds, and therefore, τ^* is an optimal solution to (8).

Theorem 4: For the stopping time τ^*, defined in (24), where $\lambda > 0$ is chosen such that $\mathbb{E}_\infty\{\tau^*\} = \eta$, the objective function meets its upper bound, and therefore, it is an optimal solution to (8) for $\xi = 2$.

An optimal stopping time that solves (20) for $\xi = 2$ is a function of (ℓ_{t-1}, ℓ_t) and can be computed numerically. The generalization of the results to any $\xi > 2$ is straightforward. The following theorem summarizes this extension.

Theorem 5: For any $\xi \in \mathbb{N}$ and for the problems formulated in (8) we have the following results.

1) An upper bound on the objective functions is

$$\mathcal{L}_{\rm L}(\tau) \leq \mathcal{L}_{\rm P}(\tau) \leq s \cdot \frac{\mathbb{E}_\infty\left\{\sum_{k=1}^{\xi}\prod_{s=0}^{k-1}\ell_{\tau-s}\right\}}{\mathbb{E}_\infty\{\tau\}} \ . \qquad (25)$$

2) An optimal stopping time that maximizes the upper bound has the form

$$\tau_{\rm g}^* \triangleq \inf\left\{t \ : \ \sum_{k=1}^{\xi}\prod_{s=0}^{k-1}\ell_{t-s} \geq -\lambda + \hat{h}(\ell_{t-\xi+2}, \ldots, \ell_t)\right\}, \qquad (26)$$

where $\hat{h}(\ell_{t-\xi+2}, \ldots, \ell_t)$ is defined as

$$\hat{h}(\ell_{t-\xi+2}, \ldots, \ell_t) \triangleq \mathbb{E}_\infty\{G(\ell_{t-\xi+2}, \ldots, \ell_t, \ell_{t+1}) \mid \ell_{t-\xi+1}, \ldots, \ell_{t-1}, \ell_t\} \ . \qquad (27)$$

3) For the stopping time in (26), we have

$$\mathcal{L}_{\rm L}(\tau_{\rm g}^*) = \mathcal{L}_{\rm P}(\tau_{\rm g}^*) = s \cdot \frac{\mathbb{E}_\infty\left\{\sum_{k=1}^{\xi}\prod_{s=0}^{k-1}\ell_{\tau_{\rm g}^*-s}\right\}}{\mathbb{E}_\infty\{\tau_{\rm g}^*\}} \ , \qquad (28)$$

and therefore, it is an optimal solution to (8).

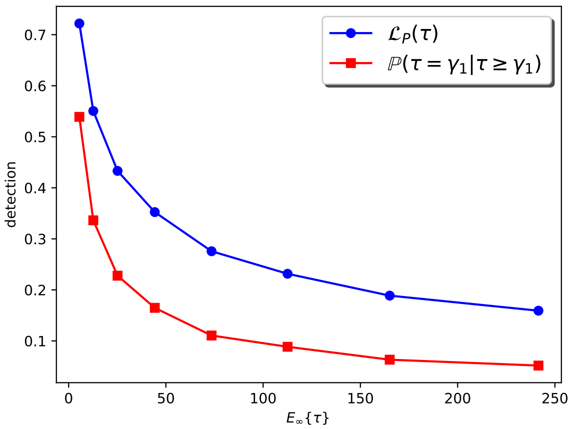

Fig. 2: The probability of detecting the first or any change-point.

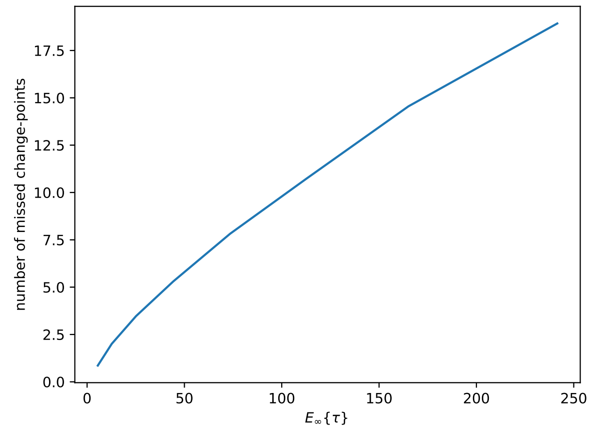

Fig. 3: The average number of missed change-points before detection.

V. Numerical Results

In this section, we numerically evaluate the performance of the optimal tests corresponding to the Pollak-like criterion characterized in sections III and IV. To this end, we consider a sequence with the nominal and the alternative distributions being the unit-variance Gaussian distributions with mean values 0 and 1, respectively. There exist 1000 change-points in the sequence, each with the duration $T = 1$. Figure 2 compares the conditional probability of detecting the first change-point with that of detecting any change-point. It is observed that when we have a more stringent constraint on the false alarm rates, i.e., the average run-length to a false alarm increases, the detection probability decreases since we want to raise fewer false alarms. Also, the ratio gap between these two objective function becomes more significant. This is due to the fact that in our objective function, we can afford to wait for a more reliable decision about the occurrence of a change-point. Figure 3 illustrates the average number of missed change-points in our setting. It is observed that for a larger average run length to a false alarm we miss more change-points in order to detect one of them more reliably with a hard deadline.

VI. Conclusion

We have analyzed the problem of quickest search for transient change-points. We have considered a setting in which a sequence of random variables might undergo multiple change-points and shortly after each change-point, they return to the nominal distribution. Both the nominal and alternative distributions are known and the objective is to identify one of these change-points in real-time no later than a pre-specified hard deadline, while controlling the false alarm rate. To this end, we have considered a probability maximizing approach in a minimax setting. We have characterized the exact optimal decision rules. Furthermore, we have shown that when the objective is detecting a change-point immediately after it occurs, the characterized decision rule simplifies to the well-known Shewhart test.

References

[1] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge, UK: Cambridge University Press, Dec. 2009.

[2] T. Bojdecki, "Probability maximizing approach to optimal stopping and its application to a disorder problem," *Stochastics*, vol. 3, no. 1, pp. 61–71, 1979.

[3] M. Pollak and A. M. Krieger, "Shewhart revisited," *Sequential Analysis*, vol. 32, no. 2, pp. 230–242, 2013.

[4] G. V. Moustakides, "Multiple optimality properties of the Shewhart test," *Sequential Analysis*, vol. 33, no. 3, pp. 318–344, 2014.

[5] ——, "Optimum Shewhart tests for Markovian data," in *Proc. Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep. 2015, pp. 822–826.

[6] E. Ebrahimzadeh and A. Tchamkerten, "Sequential detection of transient changes in stochastic systems under a sampling constraint," in *Proc. IEEE International Symposium on Information Theory*, Hong Kong, China, Jun. 2015, pp. 156–160.

[7] B. K. Guépié, L. Fillatre, and I. Nikiforov, "Sequential detection of transient changes," *Sequential Analysis*, vol. 31, no. 4, pp. 528–547, 2012.

[8] G. V. Moustakides and V. V. Veeravalli, "Sequentially detecting transitory changes," in *Proc. IEEE International Symposium on Information Theory*, Barcelona, Spain, Jul. 2016, pp. 26–30.

[9] S. Zou, G. Fellouris, and V. V. Veeravalli, "Quickest change detection under transient dynamics: Theory and asymptotic analysis," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1397–1412, Mar. 2019.

[10] V. Krishnamurthy, "Bayesian sequential detection with phase-distributed change time and nonlinear penalty—A POMDP lattice programming approach," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7096–7124, Oct. 2011.

[11] ——, "Quickest detection POMDPs with social learning: Interaction of local and global decision makers," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5563–5587, Aug. 2012.

[12] W. Sarnowski and K. Szajowski, "Optimal detection of transition probability change in random sequence," *Stochastics*, vol. 83, no. 4-6, pp. 569–581, 2011.

[13] M. Pollak, "Optimal detection of a change in distribution," *The Annals of Statistics*, vol. 13, no. 1, pp. 206–227, Mar. 1985.

[14] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, Dec. 1971.