Bandit Quickest Changepoint Detection

Anonymous Author(s)
Affiliation
Address
email

Abstract

Detecting abrupt changes in temporal behavior patterns is of interest in many industrial and security applications. Abrupt changes are often local and observable primarily through a well-aligned sensing action (e.g., a camera with a narrow field-of-view). Due to resource constraints, continuous monitoring of all of the sensors is impractical. We propose the bandit quickest changepoint detection framework as a means of balancing sensing cost with detection delay. In this framework, sensing actions (or sensors) are sequentially chosen, and only measurements corresponding to chosen actions are observed. We derive an information-theoretic lower bound on the detection delay for a general class of finitely parameterized probability distributions. We then propose a computationally efficient online sensing scheme, which seamlessly balances the need for exploration of different sensing options with exploitation of querying informative actions. We derive expected delay bounds for the proposed scheme and show that these bounds match our information-theoretic lower bounds at low false alarm rates, establishing optimality of the proposed method. We then perform a number of experiments on synthetic and real datasets demonstrating the efficacy of our proposed method.

1 Introduction

We propose a framework for bandit quickest changepoint detection (BQCD), motivated by real-world applications such as sensor network based surveillance systems [Sun+11], edge computing in IoT devices [AWC19], fault monitoring in large-scale industrial systems [Pur+19a]. Our goal is to identify changes in behavior of the system state, which evolves over a multi-dimensional space. The state can exhibit abrupt temporal changes, which may manifest in only a few components and be difficult to detect. In surveillance applications, due to narrow receptive fields of sensors (such as cameras), a large number of sensors are required to ensure continuous coverage, and intrusions (abrupt changes) manifest only among a few sensors in the field-of-view. Due to resource constraints such as limited battery power on edge-devices, sensing, processing and communication may be expensive [Hal+09], and it is desirable to minimize sensor activity at any time.

Bandit Quickest Changepoint Detection (BQCD). We consider a setting with a finite collection of sensing actions $\mathcal{A}$. At any time, t, a sensing action $A_t \in \mathcal{A}$ at time t can be deployed, and we obtain a measurement X_t in response (the 'bandit' setting[1]). The measurement X_t follows a probability law, $\mathbb{P}_t[\cdot \mid A_t, \nu]$, that depends both on the sensing action chosen as well as the time ν when abrupt change occurs, and is otherwise independent of past history. The probability laws before and after the change belong to a finitely parameterized family of distributions $\mathbb{P}_t[\cdot \mid A_t, \nu] \triangleq \mathbb{P}_{\theta_{\nu,t}}[\cdot|A_t]$. Our objective is

[1]Our usage follows classical multi-armed bandits terminology, where bandits refers to the fact that actions can yield only partial information [LS20].

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

to sequentially choose sensing actions so as to minimize the detection delay from when a change has occurred, while keeping false detection behavior of abrupt change below a pre-specified level[2].

Regret vs. Detection Delay Metric. The performance measure we adopt is not the usual regret or shortfall in cumulative 'reward' but detection delay instead, where each time spent after the real changepoint 'costs' one unit of loss. This is similar to the sample complexity metric in inference-based bandit tasks like best arm identification or bandit hypothesis testing [GK16], but having the added complexity of handling non-stationary distributional change with a false alarm rate constraint.

Information-theoretic Lower Bound. We prove a lower bound on the expected detection delay that any BQCD algorithm must suffer at a fixed false alarm rate. The lower bound exhibits a fundamental tradeoff between early stopping (false-alarm) and detection delay (time to detect abrupt change). It offers the key insight that the quickest way to detect a change, at any false-alarm rate, is by playing the 'most informative action', of an 'oracle' who a priori knows the post-change distribution. This suggests that to quickly identify changepoints, we must direct our effort towards rapidly identifying informative actions. On a technical level, we develop a change-of-measure argument for nonstationary, adaptive change detection, that allows for relating the divergences between random trajectories until stopping to the divergence of probability laws under each action for any two problem instances.

ϵ-Greedy Change Detector (ϵ-GCD). We propose ϵ-GCD, which, at a high level, uses a small amount of forced exploration to identify informative actions. The forced exploration allows for rapid convergence towards informative actions, and playing these actions minimizes detection delay. Our ϵ-GCD is based on the generalized max-likelihood/likelihood ratio principle which is utilized to estimate parametric changes. To prove detection delay bounds we draw upon key insights of ϵ-GCD. We first interpret the scheme in terms of competing parallel 'queues', where each queue corresponds to a candidate post-change parameter, collects 'arrivals' which are log-likelihood ratios of observations, and cannot go negative. The true parameter is the queue which enjoys the highest growth rate after a change, and the detection delay is the time required for it to dominate and become the 'longest queue'. The dynamics of the queues can be related to nonstationary random walks with drifts. Using these insights, we prove that the expected detection delay of ϵ-GCD at low false alarm rates mirrors our information theoretic lower bounds thus establishing optimality of our method.

Experiments. We perform numerical experiments of ϵ-GCD on synthetic and real datasets and show that under variations of changepoints, anomalies, and action sets, we realize gains due to adaptivity.

1.1 Related work

Changepoint detection deals with identifying points in time when probability laws governing a stochastic process changes abruptly. The problem of changepoint detection has been widely studied, dating back to the pioneering works of Page [Pag54] and Lorden [Lor71].

Online change detection focuses on situations where the data is obtained incrementally over time, and one must infer whether a change has occurred at each time. A large part of the online change detection literature, like our paper, adopts a frequentist approach, and, in particular, utilizes parametric models for pre-change and post-change distributions. In this context, the CUSUM algorithm [Pag54] and its variants such as the generalized likelihood ratio statistics have been proposed, in which a change is announced when the likelihood ratio statistic exceeds a threshold. While there are a number of prior works on this topic (see [BN93; CG12; VB14]), specific attention to finite time (i.e., non-asymptotic) guarantees on detection delay is more recent [LX10; Mai19], and as such remains somewhat open ([AC17]). While changepoint detection has also been studied from a Bayesian perspective, much of this literature has focused on the batch setting (see [Gun+21; AM07; FL07]). Though recent works have begun to focus on Bayesian online detection, there has been little work (apart from [AMF20]) on proving finite time guarantees.

Our work is motivated by costs imposed on data collection due to resource constraints. In this context, while adopting a frequentist perspective, we propose methods for adaptive online data selection for multi-stream time-series data. Recent works have begun to focus attention on adaptive online data collection for changepoint detection from both frequentist [ZM20] and Bayesian [OGR10; Gun+21] perspectives. Additionally, while different from our focus, we mention in passing that methods for active change-point detection [HKK19], where the task is to adaptively determine the next input

[2]For a trivial action set ($|\mathcal{A}| = 1$), this is the classical quickest change detection problem where the only adaptive decision is whether to stop or not, see e.g., [TNB14]

have also been proposed. Furthermore, there are a number of works that focus on bandit regret minimization for non-stationary time-series data [GM11; LLS18; MS13]. While these works are related we note that regret minimization is a fundamentally different objective from changepoint identification where the goal is to minimize detection delay.

We will outline similarities and differences between our work and closely related prior works. From a practical perspective, [OGR10] is similar to our work in that they too motivate their approach from a sensor network viewpoint, where sensors may undergo faults or changepoints exhibited due to environmental factors. They propose a Bayesian formalism largely based on the well-known Bayesian online change detection (BOCD) method [AM07], and leverage Gaussian processes (GP) for modeling. The GP perspective allows for tractable sequential time-series prediction and sensor selection. Recently, [Gun+21] have proposed to extend the BOCD approach to incorporate costs in making decisions for real-time data acquisition in multi-fidelity sensing scenarios. Different from these perspectives, our approach is frequentist, and does not impose action-specific costs and distributional constraints on the underlying latent parameters or on the changepoints. Furthermore, we derive finite time performance bounds, which is not a focus of any of these works.

Similar to our work, [ZM20] also adopts a frequentist perspective and proposes a method for bandit changepoint detection based on the well-known Shiryaev-Roberts-Pollak scheme. To adaptively choose data streams/sensors, they utilize Thompson sampling [Tho33] to balance exploration of different data streams for acquiring knowledge with exploitation of informative streams. In addition, they present theoretical properties, and show that their method does not trigger false alarm too soon. However, detection delays under false alarm constraints are not explicitly characterized, which is a key challenge.

In summary, the principal difference between these prior works on bandit online changepoint detection and our ϵ-GCD is that we are able to explicitly characterize information theoretic lower bounds on expected detection delay under a false alarm constraint. Furthermore ϵ-GCD is natural variant of CUSUM, and our explicit analysis shows that it exhibits finite-time performance guarantees on expected detection delay and matches our lower bound at low false alarm rates. This leads to optimality that is absent in prior work.

2 Problem model

We consider the following discrete-time model for the bandit quickest changepoint detection problem. At each decision round $t \in \mathbb{N} := \{1, 2, \ldots\}$, the parameter θ_t governs the distribution of the observation taken in that round. Let $\nu \in \mathbb{N}$ denote the round at which the change in behavior occurs, so that $\theta_t = \theta_0$ for $t < \nu$ and $\theta_t = \theta^* \in \Theta^{(1)}$ for $t \geq \nu$. At each round t, a learner decides to either (1) stop and output that a change has taken place, denoted by the random variable $U_t = 1$, or (2) continue (denoted by $U_t = 0$) and play an action or sensing decision A_t from an action set $\mathcal{A}$. Upon playing this action the learner obtains an observation X_t, sampled independent of the past, from a probability distribution $\mathbb{P}_{\theta_t}[\cdot|A_t]$, which depends on both the current system parameter θ_t and the current action A_t. The stopping decision and sensing action (U_t and A_t) are assumed to be chosen in a causal manner, i.e., depending on all past information $U_0, A_1, X_1, U_1, \ldots, A_{t-1}, X_{t-1}$, along with potentially independent internal randomness. The learner is aware, a priori, of the pre-change parameter θ_0, the post-change parameter set $\Theta^{(1)}$ and the observation distribution structure, i.e., $\{\mathbb{P}_\theta[\cdot|a]\}_{\theta,a}$ with $a \in \mathcal{A}$ denotes fixed actions. Crucially, the change time ν and the specific post-change parameter θ^* are not known in advance and must be 'learnt' in order to perform well.

The stopping time of the learner is defined to be $\tau = \min\{t \geq 0 : U_t = 1\}$, and the main performance metric that we are interested in is the detection delay, which is the random variable $(\tau - \nu)^+ = \max\{\tau - \nu, 0\}$. Specifically, we are interested in designing sampling and stopping rules so that (a) When no change occurs ($\nu = \infty$), τ should be at least a specified time delay with high probability, which is a measure of false alarm rate, and (b) when a change occurs ($\nu < \infty$), the expected detection delay should be small implying quick detection of change.

Notation. For a post-change parameter $\theta \in \Theta^{(1)}$, we use $a^*(\theta)$ to denote the most informative action for θ vs. $\theta^{(0)}$, i.e., $a^*(\theta) \in \operatorname{argmax}_{a \in \mathcal{A}} D(\mathbb{P}_\theta[\cdot|a] \,||\, \mathbb{P}_{\theta_0}[\cdot|a])$, where $D(\cdot||\cdot)$ stands for the Kullback-Leibler divergence. To lighten notation, in the sequel we often abbreviate the distribution $\mathbb{P}_\theta[\cdot|a]$ to $\theta(a)$, and thus $D(\mathbb{P}_\theta[\cdot|a] \,||\, \mathbb{P}_{\theta_0}[\cdot|a])$ to $D(\theta(a)||\theta_0(a))$. The simultaneous subscript-superscript notation z_i^j is often used to represent the sequence $(z_i, z_{i+1}, \ldots, z_j)$. For any $\theta \in \Theta^{(1)}$ and $n \in \mathbb{N}$,

we use $\mathbb{P}^{(n,\theta)}$ to denote the probability measure on the process $U_0, A_1, X_1, U_1, A_2, X_2, \ldots$ induced by the learner's decisions, when the pre-change parameter θ_0 changes to θ at time n. We also denote by $\mathbb{P}^{(\infty)}$ the probability measure as above under the no-change situation ($\nu = \infty$).

3 Fundamental limits on bandit change detection performance

Before embarking upon the design of bandit changepoint algorithms, it is worth understanding what the limits of bandit change detection performance are, due to the stochastic and partial nature of the problem's information structure. We expose in this section a universal lower bound on the detection delay of *any* bandit changepoint algorithm with a false alarm rate above a given level.

Theorem 1. *Let* $0 < \alpha \leq \frac{1}{10}$ *and* $m \geq 1$. *For any bandit changepoint algorithm satisfying* $\mathbb{P}^{(\infty)}[\tau < m] \leq \alpha$, *we have* $\mathbb{E}^{(1,\theta^*)}[\tau] \geq \min\left\{\frac{\frac{1}{20}\log\frac{1}{\alpha}}{\max_{a\in\mathcal{A}} D\left(\theta^*(a)||\theta^{(0)}(a)\right)}, \frac{m}{2}\right\} \quad \forall \theta^* \in \Theta^{(1)}$.

The result implies that in the 'small' false alarm rate (α) regime with 'large' $m = \omega(\log(1/\alpha))$, e.g., $m = \log^2(1/\alpha)$, we have that any algorithm meeting the false alarm rate property $\mathbb{P}^{(\infty)}[\tau \geq m] \geq 1 - \alpha$ must suffer a detection delay at least $\Omega\left(\frac{\log\frac{1}{\alpha}}{\max_{a\in\mathcal{A}} D\left(\theta^*(a)||\theta^{(0)}(a)\right)}\right)$ when a change occurs at time 1 (a similar, 'anytime' lower bound holds for detection delay at $\nu \in \mathbb{N}$, see appendix).

False alarm-detection Delay Tradeoff. Theorem 1 shows that a basic tradeoff exists between the false alarm rate or early stopping rate in case of no change, on one hand, and the detection delay after a true change occurs, on the other. Specifically, it is impossible to stop 'too early', i.e., before time $\Omega\left(\frac{\log\frac{1}{\alpha}}{\max_{a\in\mathcal{A}} D\left(\theta^*(a)||\theta^{(0)}(a)\right)}\right)$, after a true change if one wishes to stop 'too late' under no change, i.e., $\mathbb{P}^{(\infty)}[\tau \geq m] \geq 1 - \alpha$. We also note that when there is no adaptive sampling of actions (i.e., $|\mathcal{A}| = 1$), then the lower bound reduces to the form of a a standard lower bound for the classical changepoint detection problem [Lai98].

Information Structure. This is perhaps the most valuable implication of the lower bound. The quantity $D\left(\theta^*(a)||\theta^{(0)}(a)\right)$, in the denominator, can be interpreted as the amount of information that playing an action a provides (on average) in order to detect a change from θ_0 to θ^*. Consequently, the quickest way to detect a change is by playing the 'most informative action', $\mathrm{argmax}_{a\in\mathcal{A}} D\left(\theta^*(a)||\theta^{(0)}(a)\right)$. This can be viewed as the sampling strategy of an 'oracle' who knows the post-change parameter θ^* in advance. Unfortunately, this information is not known a priori for a causal algorithm. However, we will see that this can in fact be learnt along the trajectory and the lower bound can be attained order-wise by a suitable learning algorithm that we propose.

Proof Sketch for Theorem 1. The main idea is a measure change argument, adapted to our non-stationary change point setting, from literature on sample complexity bounds for bandit best arm identification [GK16]. On one hand, if the algorithm in consideration is very 'lazy' to begin with, i.e., $\mathbb{P}^{(1,\theta^*)}[\tau \geq m]$ is at least a constant, say $1/2$, then we immediately get $\mathbb{E}^{(1,\theta^*)}[\tau] \geq \frac{1}{2} \cdot m$. On the other hand, if the algorithm is 'active', i.e., $\mathbb{P}^{(1,\theta^*)}[\tau < m] \geq 1/2$, then the KL divergence between the laws of the indicator random variable $\mathbf{1}\{\tau < m\}$ in the instances $(1, \theta^*)$ and $(\infty, \cdot)$ is 'large' (at least a constant times $\log(1/\alpha)$) owing to the hypothesis that $\mathbb{P}^{(\infty)}[\tau < m]$ is 'small' (at most α). But by the data processing inequality, this divergence is bounded above by the divergence between the distributions of the entire trajectory $U_0, A_1, X_1, U_1, A_2, X_2, \ldots$ under the two instances, which can be seen to be equivalent to the quantity $\sum_{a\in\mathcal{A}} D\left(\theta^*(a)||\theta(a)\right) \mathbb{E}^{(1,\theta^*)}[N_\tau(a)]$, and which is further bounded above by $\mathbb{E}^{(1,\theta^*)}[\tau] \max_{a\in\mathcal{A}} D\left(\theta^*(a)||\theta(a)\right)$. See appendix for a complete proof.

4 An adaptive change detection algorithm

We describe the ϵ-GCD adaptive sensing algorithm for the bandit quickest change detection problem. The lower bound in Section 3 suggests that it is beneficial to infer the target post-change parameter, so that playing the most informative action for it can yield the best possible detection delay performance. This is the key principle underlying the design of the ϵ-GCD algorithm (Algorithm 1).

At a high level, ϵ-GCD uses a small amount of forced exploration along with 'greedy' exploitation to play sensing actions. Specifically, it computes, at each round t, a maximum likelihood estimate

(MLE) of the post-change parameter based on the generalized likelihood ratio test (GLRT) principle. It then plays either a randomly chosen action, if the current slot is an exploration slot, or a 'greedy', i.e., most informative, action for the estimated post-change parameter, if it is an exploitation slot.

The MLE of the post-change parameter, in round t, admits an interpretation as the longest 'queue' corresponding to some parameter in $\Theta^{(1)}$. To see this, notice that the MLE for the pair (ν, θ^*), given all previous data in exploration rounds can be written as $\operatorname{argmax}_{\theta\in\Theta^{(1)},1\le v\le t} \prod_{\ell=1}^{v-1} f_{\theta_0}(X_\ell|A_\ell)^{E_\ell} \prod_{\ell=v}^{t-1} f_\theta(X_\ell|A_\ell)^{E_\ell} = \operatorname{argmax}_{\theta\in\Theta^{(1)},1\le v\le t} \sum_{\ell=v}^{t-1} E_\ell \log \frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}$, with $f_\theta(\cdot|a)$ taken to be the probability density or mass function of the distribution $\mathbb{P}_\theta(\cdot|A_\ell)$ (assuming one exists). Observe now that for each candidate post-change parameter θ, the inner maximum over v, $Q_t^{(1)}(\theta) := \operatorname{argmax}_{1\le v\le t} \sum_{\ell=v}^{t-1} E_t \log \frac{f_\theta(X_t|A_t)}{f_{\theta_0}(X_t|A_t)}$, evolves as a 'queue'[3] with arrivals $E_\ell \log \frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}$ at each time slot ℓ, i.e., $Q_{t+1}^{(1)}(\theta) = \left(Q_t^{(1)}(\theta) + E_\ell \log \frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}\right)^+$.

We define the algorithm using general 'processing' functions g_θ in place of the log-likelihood ratios $\log \frac{f_\theta}{f_{\theta_0}}$ above. Broadly speaking, the functions g should ideally be chosen with the hope that (a) $g_\theta(X_\ell, A_\ell)$ is negative in expectation before the change time ($\ell < \nu$), and (b) $g_\theta(X_\ell, A_\ell)$ is large and positive in expectation for $\theta = \theta^*$, the true change parameter, after the change ($\ell \ge \nu$).

The stopping rule that ϵ-GCD uses is based on the generalized likelihood ratio (GLR)-type statistic. It is the largest of an ensemble of evolving exploitation-data queues $Q_t^{(2)}(\theta)$, which mirrors the definition of $Q_t^{(1)}(\theta)$ but with $1 - E_t$ instead of E_t.

Algorithm 1 ϵ-GCD

Input: $\epsilon \in [0,1]$, θ_0, $\Theta^{(1)}$, $\beta \ge 1$, π, Observation function $g_\theta(x,a)\ \forall(\theta,x,a) \in \Theta^{(1)} \times \mathcal{X} \times \mathcal{A}$.
Init: $Q_1^{(1)}(\theta) \leftarrow 0$, $Q_1^{(2)}(\theta) \leftarrow 0\ \forall\theta \in \Theta^{(1)}$ {CUSUM statistics for exploration and exploitation per θ}
for round $t = 1, 2, 3, \ldots$ **do**
  if $\max_{\theta\in\Theta^{(1)}} Q_t^{(2)}(\theta) \ge \beta$ **then**
    break {Stop sampling and exit}
  end if
  Sample $E_t \sim \text{Ber}(\epsilon)$ independently
  if $E_t == 1$ **then**
    Play action $A_t \sim \pi$ independently {Explore}
    Get observation X_t
    Set $\forall\theta \in \Theta^{(1)} : Q_{t+1}^{(1)}(\theta) \leftarrow \left(Q_t^{(1)}(\theta) + g_\theta(X_t, A_t)\right)^+$ {Update exploration CUSUM statistics}
  else
    Compute $\hat{\theta}_t = \operatorname{argmax}_{\theta\in\Theta^{(1)}} Q_t^{(1)}(\theta)$ {Most likely post-change distribution based on exploration data}
    Play action $A_t = \operatorname{argmax}_{a\in\mathcal{A}} D\left(\theta^{(0)}(a)||\hat{\theta}_t(a)\right)$
    Get observation X_t
    Set $\forall\theta \in \Theta^{(1)} : Q_{t+1}^{(2)}(\theta) \leftarrow \left(Q_t^{(2)}(\theta) + g_\theta(X_t, A_t)\right)^+$ {Update exploitation CUSUM statistics}
  end if
end for

5 Theoretical guarantees for the ϵ-GCD algorithm

In this section, we present and discuss theoretical guarantees on the false alarm rate and detection delay performance of the ϵ-GCD algorithm of Section 4.

We make the following assumptions on the parameter space and observation distributions in order to derive performance bounds for the algorithm.

Assumption 1. *The post-change parameter set $\Theta^{(1)}$ is finite, i.e., $|\Theta^{(1)}| < \infty$.*

This assumption is made primarily for ease of analysis, whereas the algorithm is defined even for arbitrary parameter spaces. Specifically, it allows for easy control of the fluctuations of an ensemble of (drifting) random walks via a union bound over the parameter set. While we believe that this can be relaxed to handle general parameter spaces via appropriate netting or chaining arguments, this chiefly technical task is left to future investigation.

Assumption 2. *Every post-change parameter is detectable by some action, i.e., $\forall\theta \in \Theta^{(1)}\ \exists a \in \mathcal{A} : D\left(\theta(a)||\theta^{(0)}(a)\right) > 0$.*

[3] This is also known as the Lindley recursion equation in queueing theory.

This is a basic identifiability requirement of the setting, without which some parameter changes could be completely undetectable. Put differently, one can only hope to detect changes that the sensing set can tease apart.

Assumption 3 (Bounded marginal KL divergences)**.** *There is a constant $D_{\max}$ satisfying $D\left(\theta_1(a)||\theta_2(a)\right) \leq D_{\max}$ for each $a \in \mathcal{A}$, $\theta_1, \theta_2, \in \Theta^{(1)}$.*

This assumption is easily met, for instance, if all log likelihood ratios are bounded, or if the observations are modelled as Gaussian distributions; in this case log likelihood ratios simply become linear functions of the observation.

Assumption 4. *Every observation probability distribution $\mathbb{P}_\theta\left[\cdot|a\right]$, for $a \in \mathcal{A}$ and $\theta \in \Theta^{(1)} \cup \{\theta_0\}$, has either a density[4] or a mass function, denoted by $f_\theta(\cdot|a)$. Moreover, there exists $r > 0$ such that for any $\theta, \theta', \theta'' \in \Theta^{(1)} \cup \{\theta_0\}$, and $a \in \mathcal{A}$, the log likelihood ratio $\frac{f_{\theta'}(X|a)}{f_{\theta''}(X|a)}$ is r-subgaussian [5] under $X \sim \mathbb{P}_\theta\left[\cdot|a\right]$.*

This assumption is common in the statistical inference literature; we use it to be able to control the fluctuations of the exploration and exploitation CUSUM statistics in the algorithm via standard subgaussian concentration tools.

A concrete example of a setting that meets Assumptions 1-4 is the linear observation model with additive Gaussian noise (the standard 'linear bandit' model [APS11]). Here, the observation distribution for action $a \in \mathbb{R}^d$ when the system parameter is $\theta \in \{\theta_0\} \cup \Theta^{(1)} \subset \mathbb{R}^d$ is $\mathcal{N}(\langle a, \theta\rangle, \sigma^2_{a,\theta})$, where $\sigma^2_{a,\theta} \leq r$ and $\left|\Theta^{(1)}\right| < \infty$, and where $\langle \cdot, \cdot \rangle$ represents the standard inner product in $\mathbb{R}^d$.

We are now in a position to state our key theoretical result.

Theorem 2 (False alarm and detection delay for general change point)**.** *Under assumptions 1-4, the following conclusions hold for ϵ-GCD (Algorithm 1) run with the log-likelihood observation function $g_\theta(x, a) \equiv \log \frac{f_\theta(x|a)}{f_{\theta_0}(x|a)}$.*

1. (Time to false alarm) Let $\alpha \in (0, 1)$ and $m \in \mathbb{N}$. If the stopping threshold β is set as $\beta \geq \frac{m\left|\Theta^{(1)}\right|}{\alpha}$, then the stopping time satisfies $\mathbb{P}^{(\infty)}\left[\tau < m\right] \leq \alpha$.
2. (Detection delay) For a change from θ_0 to $\theta^* \in \Theta^{(1)}$ occurring at time $\nu \in \mathbb{N}$,

$$\mathbb{E}^{(\nu,\theta^*)}\left[\tau\right] \leq \frac{8\log\beta}{(1-\epsilon)\max_a D\left(\theta^*(a)||\theta^{(0)}(a)\right)} + O\left(poly\left(\frac{1}{\epsilon}\right), \mathbb{E}^{(\infty)}\left[Q^{(1)}_\nu\right], \mathcal{P}\right), \qquad (1)$$

provided $\pi(a^*_{\theta^*}) > 0$, with $Q^{(1)}_\nu := \left(Q^{(1)}_\nu(\theta)\right)_{\theta\in\Theta^{(1)}}$ denoting the explore queue statistics for all parameters and $\mathcal{P} \equiv \left(\mathbb{P}_\theta\left[\cdot|a\right]\right)_{\theta\in\{\theta_0\}\cup\Theta^{(1)}}$ all the observation distributions.

Interpretation of the Delay Bound. The first term in the detection delay bound (1) can be interpreted as the time for an *oracle* fixed-action strategy (e.g., CUSUM), that always plays the most informative action for θ^*, to stop. The second term, on the other hand, is a bound on the time taken by the algorithm to *learn* to play the most informative action for θ^* (details follow in proof sketch). We omit the precise dependence of the second term on the problem structure here for brevity, but detail it explicitly in the appendix. Specifically, for small forced exploration rates ϵ, the second term depends on the information geometry of the problem approximately as $\frac{1}{\epsilon^4}\sum_{\theta\in\Theta^{(1)}:a^*_\theta\neq a^*_{\theta^*}}\frac{1}{\Delta^4_\theta}$. Here, for each candidate post-change parameter $\theta \in \Theta^{(1)}$, its 'gap' Δ_θ is a measure of how difficult it is for the algorithm to eliminate θ as an estimate for the true post-change parameter θ^* during this parameter-learning phase. We formally define it as $\Delta_\theta := \bar{D}_{\theta^*,\theta_0} - \frac{1}{2}(\bar{D}_{\theta^*,\theta_0} + \left(\bar{D}_{\theta^*,\theta_0} - \bar{D}_{\theta^*,\theta}\right)^+)$, where $\bar{D}_{\theta^*,\theta} := \sum_{a\in\mathcal{A}}\pi(a)D\left(\theta^*(a)||\theta(a)\right)$ is the average information divergence between parameters θ^*, θ offered by playing from the exploration distribution π. With finer analysis, the dependence on gaps and ϵ could be improved. We omit it for simplicity.

We now turn to the (additive & linear; see the appendix) dependence on $\mathbb{E}^{(\infty)}\left[Q^{(1)}_\nu\right]$ of the second term. The insight as to why this arises is as follows. The 'learning' phase for θ^*, after a change takes

[4] with respect to a standard reference measure, e.g., Lebesgue measure

[5] A random variable X is r-subgaussian under $X \sim \mathbb{P}$ if $\forall\lambda \in \mathbb{R}$: $\mathbb{E}[e^{\lambda(X-\mathbb{E}[X])}] \leq e^{\lambda^2 r/2}$.

place at time ν, can be seen as a 'race' between competing log-likelihood queues of all the parameters vying to become the MLE $\hat{\theta}_t$. At the change time, each non-ground truth queue $Q_\nu^{(1)}(\theta)$ is, on average, at level $\mathbb{E}^{(\infty)}\left[Q_\nu^{(1)}(\theta)\right]$, representing its initial 'advantage' over the ground-truth queue $Q_\nu^{(1)}(\theta^*)$ which in the worst case could be at level 0. Thus, the difference $\mathbb{E}^{(\infty)}\left[Q_\nu^{(1)}(\theta)\right] - Q_\nu^{(1)}(\theta^*) \leq Q_\nu^{(1)}(\theta)$ is the extra amount of 'time work' that the θ^* queue must do to overcome the θ queue. A special case of the result is for $\nu = 1$ in which case all queues start out at 0. For general ν, one can establish that $\mathbb{E}^{(\infty)}\left[Q_\nu^{(1)}(\theta)\right]$ must be finite, by noticing that each queue is non-negative and accumulates arrivals with negative mean before the change, and applying standard queueing stability arguments. This also indicates that the detection delay bound is roughly invariant given the location of the change point ν.

Optimality at low false-alarms. The result implies that in the 'small' time to false alarm regime, i.e., $\alpha \to 0$, $m = \omega(\log(1/\alpha))$ and $\log(m) = o(\log(\alpha))$, the detection delay when β is set as above is dominated by the first term: $\mathbb{E}^{(\nu,\theta^*)}[\tau] = O\left(\frac{\log\beta}{(1-\epsilon)\max_a D\left(\theta^*(a)||\theta^{(0)}(a)\right)}\right)$. This matches, order-wise, the universal lower bound of Theorem 1 up to the algorithm-dependent multiplicative factor $1/(1-\epsilon)$, which may be interpreted as a 'penalty' for the forced exploration mechanism.

Benefits of Adaptivity. Let us consider a concrete setting to elucidate the detection delay bound. In this example, there are d sensing actions, represented by the canonical basis vectors in $\mathbb{R}^d$. The pre-change parameter is $0 \in \mathbb{R}^d$, and there are d candidate post-change parameters, each of which is a canonical basis vector e_i, $i \in [d]$, multiplied by a constant δ. The observation from applying action a at system parameter θ is Gaussian distributed with mean $\langle a, \theta\rangle$ and variance 1. Thus, the aim is to detect a (sparse) change of magnitude δ in one of the d coordinates in $\theta_0 = 0$, when at any time only a single coordinate can be noisily sensed. Suppose $\theta^* = \delta e_1$ without loss of generality, for which the most informative action (in fact, the only informative one) is $a^*_{\theta^*} = e_1$. The first term in the detection delay given by Theorem 2 is then $O\left(\frac{\log(1/\alpha)}{(1-\epsilon)\delta^2}\right)$. This is a factor of d smaller than $\Omega\left(d \cdot \frac{\log(1/\alpha)}{(1-\epsilon)\delta^2}\right)$ that a standard CUSUM rule with non-adaptive uniform sampling over coordinates would achieve – this is seen by applying the lower bound (Theorem 1) to the case of a single (trivial action) where the divergence in the denominator reduces to the *average* divergence across all coordinates: $O(\delta^2/d)$. For estimating the second term, we calculate that for each $\theta = \delta e_i$ with $i \neq 1$, $\Delta_\theta = \bar{D}_{\theta^*,\theta_0} - \frac{(\bar{D}_{\theta^*,\theta_0} + \left(\bar{D}_{\theta^*,\theta_0} - \bar{D}_{\theta^*,\theta}\right)^+)}{2} = \frac{\delta^2}{2d} - \frac{\frac{\delta^2}{2d} - \left(\frac{\delta^2}{2d} - \frac{\delta^2}{d}\right)^+}{2} = \frac{\delta^2}{2d} - \frac{\frac{\delta^2}{2d} - 0}{2} = \frac{\delta^2}{4d}$. Thus, the second term representing the time to learn the optimal sensing action scales as $O\left(\frac{(d-1)d^4}{\epsilon^4\delta^8}\right) = O\left(\frac{d^5}{\epsilon^4\delta^8}\right)$. Although we have not optimized dependence on d, ϵ and δ, the overall bound still gives an idea of how the active sensing problem gets harder as the dimension d grows or as the minimum change amount δ changes, making it akin to finding a 'needle in a haystack'.

Sketch of the Proof of Theorem 2. This section lays down the key arguments involved in the proof of Theorem 2 for bounding the false alarm time and detection delay of the ϵ-GCD algorithm.

1. Bounding the probability of early stopping under no change. The stopping time τ for ϵ-GCD, by line **??** in Algorithm 1, is equivalent to the first instant t when the worst-case (over $\Theta^{(1)}$) CUSUM statistic computed over exploitation data, i.e., $\max_{\theta\in\Theta^{(1)}} Q_t^{(2)}(\theta) \equiv \max_{\theta\in\Theta^{(1)}} \max_{1\leq s\leq t} \prod_{\ell=s}^{t-1}\left(\frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}\right)^{(1-E_\ell)}$ exceeds the level $\beta \geq 1$. Under no change, each observation X_ℓ is distributed as $\mathbb{P}_{\theta_0}[\cdot|A_\ell]$ given A_ℓ, and the product $\prod_{\ell=s}^{t-1}\left(\frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}\right)^{(1-E_\ell)}$, over $t = s, s+1, s+2, \ldots$, behaves as a (standard) likelihood ratio martingale with unit expectation under the appropriate filtration. Hence, the chance that the largest among this ensemble of martingales, one for each $\theta \in \Theta^{(1)}$ and starting time $s \in [m]$, rises above β before time m is bounded using a union bound and Doob's inequality for each individual martingale, yielding the probability bound $\frac{m\left|\Theta^{(1)}\right|}{\beta}$.

2. Control of the detection delay. Suppose that the change takes place at time $\nu = 1$ to the post-change parameter $\theta^* \in \Theta^{(1)}$. The proof strategy is to show, with high probability, that τ is no more than $t_1 + t_2$, where

- t_1 is an upper bound for the time taken for the plug-in estimate $\hat{\theta}_t$, of the post-change parameter, to 'settle' to θ^*. In other words, we show that after time t_1 it is very unlikely that an action other than $a^*_{\theta^*} = \text{argmax}_{a\in\mathcal{A}}\, D\left(\theta^{(0)}(a)||\theta^*(a)\right)$ is played in every exploitation round.
- t_2 is an upper bound for the time taken for the worst-case CUSUM statistic $\max_{\theta\in\Theta^{(1)}} Q_t^{(2)}(\theta)$ to grow to the level β *assuming that the optimal (i.e., most informative) action $a^*_{\theta^*}$ is always played at all exploitation rounds*.

For clarity of exposition, we will assume that we are in the setting of *linear measurements* in $\mathbb{R}^d$ with *additive standard Gaussian noise* – this makes KL divergences easy to interpret as Euclidean distances – and that the changepoint is at $\nu = 1$.

Step 1: Finding t_1 (time to learn θ^*). We observe that the MLE $\hat{\theta}_t$ at time t can be written as the parameter θ associated with the largest stochastic process $\max_{v=1}^{t-1} \log J_{v,t-1}(\theta)$, one for each θ: $\hat{\theta}_t = \text{argmax}_{\theta\in\Theta^{(1)}} \max_{v=1}^{t-1} \log J_{v,t-1}(\theta)$, where $J_{v,t-1}(\theta) := \prod_{\ell=v}^{t-1} \left(\frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}\right)^{E_\ell}$.

Under the post-change distribution $\mathbb{P}^{(1,\theta^*)}$, $\log J_{v,t-1}(\theta)$ can be seen to evolve (as a function of t) as a random walk with drift $\epsilon(\|\theta^* - \theta\|_H^2 - \|\theta\|_H^2)$. Here, $\|\cdot\|_H$ is the usual matrix-weighted Euclidean norm in $\mathbb{R}^d$, governed by how much different directions are explored in expectation[6] by the exploration distribution π: $H = \sum_{a\in\mathcal{A}} \pi(a) aa^T$. Note that $\|\theta^* - \theta\|_H^2$ here is the KL divergence between the distributions of the observation that results when an action sampled from π is played, under parameters θ^* and θ.

The preceding discussion implies that the drift is the largest (and positive) for the random walk corresponding to the true post-change parameter $\theta = \theta^*$. The remainder of this part of the proof uses martingale concentration tools (subgaussian Chernoff and maximal Hoeffding bounds) to find the time t_1 at which the fastest-growing random walk, $\log J_{v,t}(\theta^*)$ dominates all other 'competing' random walks $\log J_{v,t}(\theta)$, $\theta \neq \theta^*$.

Step 2: Finding t_2 (time to stop under optimal action plays). In a manner similar to that of Step 1, we observe that the logarithm of the log-likelihood ratio for θ^* w.r.t. θ_0 computed over only *exploitation* rounds, i.e., $\log \prod_{\ell=v}^{t-1} \left(\frac{f_\theta(X_\ell|A_\ell)}{f_{\theta_0}(X_\ell|A_\ell)}\right)^{1-E_\ell}$, evolves as a random walk with drift rate $(1-\epsilon)\,\|\theta^*\|_H$. A Chernoff bound can then be used to show the bound (**??**) on the probability of the 'bad' event that this growing random walk has *not* crossed the level $\log\beta$ in a certain time duration t_2 (t_2 does not appear explicitly in the main derivation).

6 Experiments

Our goal in this section is to illustrate various aspects of our theory through experiments on synthetic datasets and explore performance of the proposed ϵ-GCD algorithm on a setting based on a real-world dataset. We compare the performance of three change detection strategies: ϵ-GCD, uniformly random selection (URS) and the 'oracle' benchmark that plays the most informative arm for the post-change parameter right from the start. Note that all these strategies share the same stopping rule. For ϵ-GCD, we found it advantageous to update the MLE estimate $\hat{\theta}_t$ both during exploration and exploitation phases (the appendix contains details of the algorithm used). All experiments were performed on laptop Intel Core i5 CPUs with 8GB of RAM, and take under an hour to execute.

Size	Oracle	ϵ-GCD	URS
10	66 ± 32	100 ± 21	300 ± 30
15	64 ± 32	120 ± 21	450 ± 76
20	62 ± 32	170 ± 45	630 ± 85
25	65 ± 33	190 ± 85	790 ± 100

Table 1: Observed mean and standard deviation for the simulated stopped time for varying graph-sizes with pointy action set $\mathcal{A}_1$, for change occurring at $\nu = 40$.

Synthetic Experiments. We explore how average detection delay varies under controlled variation of various parameters such as changepoint location, dimensionality and structure of the ambient space, and the type of action sets. Our objective is twofold: (a) Illustrate gains due to adaptivity of proposed ϵ-GCD over the non-adaptive method, where actions are chosen uniformly at random; (b) Demonstrate "near" optimality by baselining against Oracle.

[6]This matrix also arises commonly in linear design of experiments as the Kiefer-Wolfowitz matrix.

We report results for the case when the ambient space is a line graph. Nodes are interpreted as physical locations, and take values in $[N]$. Nodes j, k are connected if $|j-k|=1$. Each node $n \in [N]$ offers a Gaussian-distributed observation depending on the changepoint $\nu \in \mathbb{N}$. In particular, the signal at time, t, is a random vector, $S(t) = [S_n(t)]_{n\in[N]} \in \mathbb{R}^N$ where $S_n(t) = \theta_n \mathbf{1}\{[t \geq \nu]\} + W_n(t),\ t = 0,\ 1,\ \ldots$; $W_n(t) \in \mathcal{N}(0, \sigma^2)$ is IID Gaussian.

Isolated and Structured Anomalies. The vector of change parameters $\theta = [\theta_n]$ are of two types: (a) Isolated singleton change, namely, $\theta_n \in \{0,\ 1\}$ and $\sum_{n\in[N]} \theta_n = 1$; (b) Structural changes, i.e., $\text{Supp}(\theta) = \{n \in [N] : \theta_n \neq 0\}$ is k-connected set with $\theta \in \{0,1\}^N$.

Diffuse and 'Pointy' Action sets. In a parallel fashion we allow actions, $a_n \in \mathcal{A} \subset \{0,\ 1\}^N$, and actions sets, $\mathcal{A}_1$, to be either pointy, namely, $|\text{Supp}(\mathcal{A}_1)| = 1$, which allows probing only single nodes, or diffuse, $\mathcal{A}_2$ where only a connected subset of nodes can be queried. In either case, the observation received on an action, $a \in \mathcal{A}$ is given by $X_a = \langle \frac{a}{\|a\|_2}, S\rangle$, where the normalization arises because we want to maintain the same SNR across different probes.

Results. Here we report pointy actions and isolated anomalies, and other results in Appendix. We choose $\sigma^2 = 0.5$, β such that the false alarms are about 1% for the Oracle, and a forced exploration rate $\epsilon = 0.2$. Our results are averaged over 5000 Monte-Carlo runs. With this choice for β we did not observe false alarms in any of the algorithms.
Gains from Adaptivity. As pointed out in Sec. 5 (see Lines 292-309), ϵ-GCD exhibits improved detection performance relative to non-adaptive URS. This is seen in Table. 1, where as the graph size increases, non-adaptive methods scale with ambient dimension, while ϵ-GCD is observed to be relatively stable.
Optimality. We next compare ϵ-GCD against Oracle with varying change-points. As seen, ϵ-GCD somewhat saturates with ambient dimension. The difference between Oracle and ϵ-GCD arises from forced exploration. URS has significantly larger delay.
Variation with Changepoint ν. For a fixed graph size, we observed that the average expected delay is relatively constant for all methods, which is consistent with Theorem 2.

Audio based recognition of machine anomalies.

We experiment using the MIMII audio dataset [Pur+19b]. Detailed specifics are in the Appendix. The dataset has four machines (ID00,ID02,ID04,ID06), each equipped with audio sensors recording the health of the machine. There are three types of anomaly – rail damage, loose belt, no grease – which can occur at any time, in any machine. Corresponding to each anomaly there is an audio stream, and the anomaly occurs in one of the four machines at an arbitrary time. For each machine, the dataset contains audio-streams of 1000 normal and 300 abnormal files, and each audio-stream is about 10 seconds long.

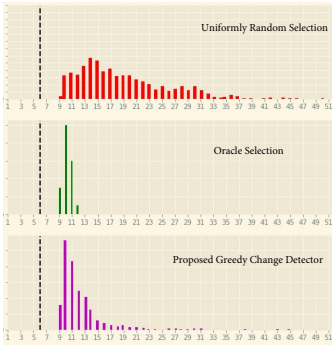

Figure 1: Audio based change detection of machine anomaly: Histogram of stopping times by URS, Oracle, and ϵ-GCD. Histograms for Oracle and ϵ-GCD demonstrates clear adaptivity gains.

Audio Processing. For each audio-stream we train auto-encoders on normal data using mel-spectrogram features, and fit Gaussians to the reconstruction errors. This results in pre- and post change parameters for each machine's autoencoded reconstruction error score, corresponding to normal and abnormal operation.

Experiment. To simulate BQCD, we introduced anomalies in machine bearing ID00 as follows. We concatenated 6 normal files and 54 abnormal files chosen uniformly at random from machine ID00. For the other machines we concatenated 60 normal files at random. The 60 files correspond to 600 seconds. Our changepoint corresponds to 6th file, which we denote as $\nu = 6$ and our task is to detect this change. Note that both the machine ID and the changepoint is not known to the learner. Our results are depicted as histograms for changepoints of anomaly detection in Fig. 1. As observed, we notice that ϵ-GCD's performance is close to Oracle both in mean and distribution, while URS exhibits larger delay and significant variance. The appendix simulates histograms for a larger changepoint.

References

[AC17] Samaneh Aminikhanghahi and Diane J. Cook. "A Survey of Methods for Time Series Change Point Detection". In: *Knowl. Inf. Syst.* 51.2 (May 2017), pp. 339–367.

[AM07] Ryan Prescott Adams and David J. C. MacKay. *Bayesian Online Changepoint Detection*. 2007. arXiv: 0710.3742 [stat.ML].

[AMF20] Reda Alami, Odalric Maillard, and Raphael Feraud. "Restarted Bayesian Online Changepoint Detector achieves Optimal Detection Delay". In: *Proceedings of the 37th International Conference on Machine Learning*. Ed. by Hal Daumé III and Aarti Singh. Vol. 119. Proceedings of Machine Learning Research. PMLR, 13–18 Jul 2020, pp. 211–221. URL: http://proceedings.mlr.press/v119/alami20a.html.

[APS11] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. "Improved Algorithms for Linear Stochastic Bandits." In: *NIPS*. Vol. 11. 2011, pp. 2312–2320.

[AWC19] Samaneh Aminikhanghahi, Tinghui Wang, and Diane J. Cook. "Real-Time Change Point Detection with Application to Smart Home Time Series Data". In: *IEEE Transactions on Knowledge and Data Engineering* 31.5 (2019), pp. 1010–1023.

[BN93] Michèle Basseville and Igor V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. USA: Prentice-Hall, Inc., 1993. ISBN: 0131267809.

[CG12] Jie Chen and Arjun K Gupta. *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance; 2nd ed.* Boston: Springer, 2012. DOI: 10.1007/978-0-8176-4801-5. URL: https://cds.cern.ch/record/1498647.

[FL07] Paul Fearnhead and Zhen Liu. "On-line inference for multiple changepoint problems". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69.4 (Sept. 2007), pp. 589–605. DOI: 10.1111/j.1467-9868.2007.00601.x. URL: https://doi.org/10.1111%2Fj.1467-9868.2007.00601.x.

[GK16] Aurelien Garivier and Emilie Kaufmann. "Optimal Best Arm Identification with Fixed Confidence". In: *Conference On Learning Theory*. 2016, pp. 998–1027.

[GM11] Aurélien Garivier and Eric Moulines. "On upper-confidence bound policies for switching bandit problems". In: *International Conference on Algorithmic Learning Theory*. Springer. 2011, pp. 174–188.

[Gun+21] Gregory W Gundersen, Diana Cai, Chuteng Zhou, Barbara E Engelhardt, and Ryan P Adams. "Active multi-fidelity Bayesian online changepoint detection". In: *arXiv preprint arXiv:2103.14224* (2021).

[Hal+09] Malka N Halgamuge, Moshe Zukerman, Kotagiri Ramamohanarao, and Hai Le Vu. "An estimation of sensor energy consumption". In: *Progress In Electromagnetics Research B* (2009).

[HKK19] Shogo Hayashi, Yoshinobu Kawahara, and Hisashi Kashima. "Active Change-Point Detection". In: *Proceedings of The Eleventh Asian Conference on Machine Learning*. Ed. by Wee Sun Lee and Taiji Suzuki. Vol. 101. Proceedings of Machine Learning Research. Nagoya, Japan: PMLR, 17–19 Nov 2019, pp. 1017–1032.

[Lai98] Tze Leung Lai. "Information bounds and quick detection of parameter changes in stochastic systems". In: *IEEE Transactions on Information Theory* 44.7 (1998), pp. 2917–2929.

[LLS18] Fang Liu, Joohyun Lee, and Ness Shroff. "A change-detection based framework for piecewise-stationary multi-armed bandit problem". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. 1. 2018.

[Lor71] Gary Lorden. "Procedures for reacting to a change in distribution". In: *The Annals of Mathematical Statistics* 42.6 (1971), pp. 1897–1908.

[LS20] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

[LX10] Tze Leung Lai and Haipeng Xing. "Sequential Change-Point Detection When the Pre- and Post-Change Parameters are Unknown". In: *Sequential Analysis* 29.2 (2010), pp. 162–175.

[Mai19] Odalric-Ambrym Maillard. "Sequential change-point detection: Laplace concentration of scan statistics and non-asymptotic delay bounds". In: *Proceedings of the 30th International Conference on Algorithmic Learning Theory*. Ed. by Aurélien Garivier and Satyen Kale. Vol. 98. Proceedings of Machine Learning Research. Chicago, Illinois: PMLR, 22–24 Mar 2019, pp. 610–632. URL: http://proceedings.mlr.press/v98/maillard19a.html.

[MS13] Joseph Mellor and Jonathan Shapiro. "Thompson sampling in switching environments with bayesian online change detection". In: *Artificial Intelligence and Statistics*. PMLR. 2013, pp. 442–450.

[OGR10] Michael A. Osborne, Roman Garnett, and Stephen J. Roberts. "Active Data Selection for Sensor Networks with Faults and Changepoints". In: *24th IEEE International Conference on Advanced Information Networking and Applications, AINA 2010, Perth, Australia, 20-13 April 2010*. IEEE Computer Society, 2010, pp. 533–540. DOI: 10.1109/AINA.2010.36. URL: https://doi.org/10.1109/AINA.2010.36.

[Pag54] E. S. Page. "Continuous Inspection Schemes". In: *Biometrika* 41.1/2 (1954), pp. 100–115. ISSN: 00063444. URL: http://www.jstor.org/stable/2333009.

[Pur+19a] Harsh Purohit, Ryo Tanabe, Kenji Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi. *MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection*. 2019. arXiv: 1909.09347.

[Pur+19b] Harsh Purohit, Ryo Tanabe, Kenji Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi. *MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection*. Version public 1.0. Zenodo, Sept. 2019. DOI: 10.5281/zenodo.3384388. URL: https://doi.org/10.5281/zenodo.3384388.

[Sun+11] Zhi Sun, Pu Wang, Mehmet C. Vuran, Mznah A. Al-Rodhaan, Abdullah M. Al-Dhelaan, and Ian F. Akyildiz. "BorderSense: Border patrol through advanced wireless sensor networks". In: *Ad Hoc Networks* 9.3 (2011), pp. 468–477. ISSN: 1570-8705. DOI: https://doi.org/10.1016/j.adhoc.2010.09.008. URL: https://www.sciencedirect.com/science/article/pii/S1570870510001484.

[Tho33] William R Thompson. "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples". In: *Biometrika* 25.3/4 (1933), pp. 285–294.

[TNB14] Alexander Tartakovsky, Igor Nikiforov, and Michele Basseville. *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.

[VB14] Venugopal V Veeravalli and Taposh Banerjee. "Quickest change detection". In: *Academic Press Library in Signal Processing*. Vol. 3. Elsevier, 2014, pp. 209–255.

[ZM20] Wanrong Zhang and Yajun Mei. *Bandit Change-Point Detection for Real-Time Monitoring High-Dimensional Data Under Sampling Control*. 2020. arXiv: 2009.11891 [stat.ME].

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] See Section 1.
   (b) Did you describe the limitations of your work? [Yes] See Section 5 for assumptions under which the results hold.
   (c) Did you discuss any potential negative societal impacts of your work? [No] As a primarily theoretical paper concerned with efficiency, our work does not have any meaningful negative impacts on society beyond the larger issues of how it is employed.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] See Section 5.
   (b) Did you include complete proofs of all theoretical results? [Yes] See the supplementary material for all proofs.
3. If you ran experiments...

(a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] We have included all the code and instructions to run it as part of the supplementary material. The dataset (MIMII audio sounds) used is large, and is available publicly.

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] We have specified the details of the experiments and parameter choices in the supplementary material.

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] We report either standard deviations or empirical distributions in our experiments.

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] See Section 6.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

(a) If your work uses existing assets, did you cite the creators? [Yes] See subsection on "Audio based recognition of machine anomalies" in Section 6.

(b) Did you mention the license of the assets? [Yes] See the citation to the MIMII audio dataset in Section 6.

(c) Did you include any new assets either in the supplemental material or as a URL? [N/A] No new assets were created.

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A] We have not used data from people.

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A] The only measured data used in the paper are those of machine sounds.

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]