Asymptotically optimal sampling policy for quickest change detection with observation-switching cost

Lau, Tze Siong; Tay, Wee Peng

2021

https://hdl.handle.net/10356/152708

https://doi.org/10.1109/TSP.2021.3057258

Asymptotically Optimal Sampling Policy for Quickest Change Detection with Observation-Switching Cost

Tze Siong Lau and Wee Peng Tay, *Senior Member, IEEE*

Abstract—We consider the problem of quickest change detection (QCD) in a signal where its observations are obtained using a set of actions, and switching from one action to another comes with a cost. The objective is to design a stopping rule consisting of a sampling policy to determine the sequence of actions used to observe the signal and a stopping time to quickly detect for the change, subject to a constraint on the average observation-switching cost. We propose an open-loop sampling policy of finite window size and a generalized likelihood ratio (GLR) Cumulative Sum (CuSum) stopping time for the QCD problem. We show that the GLR CuSum stopping time is asymptotically optimal with a properly designed sampling policy and formulate the design of this sampling policy as a quadratic programming problem. We prove that it is sufficient to consider policies of window size not more than one when designing policies of finite window size and propose several algorithms that solve this optimization problem with theoretical guarantees. Finally, we apply our approach to the problem of QCD of a partially observed graph signal and empirically demonstrate the performance of our proposed stopping times.

Index Terms—Quickest change detection, GLR CuSum, sampling policy, graph sampling

I. INTRODUCTION

Quickest change detection (QCD) is the problem of detecting an abrupt change in a system while keeping the detection delay to a minimum. In a usual scenario, a sequence of independent and identically distributed (i.i.d.) observations $\{x_t : t \in \mathbb{N}\}$, with probability density function (pdf) p_0 up to an unknown change point ν, and i.i.d. with pdf $p_1 \neq p_0$ after ν, is obtained. The objective is to detect for the change at ν as quickly as possible while maintaining false alarm constraints [1]–[3]. QCD has applications across diverse fields, including quality control [4]–[7], fraud detection [8], cognitive radio [9], [10], network surveillance [11], [12], structural health monitoring [13], spam detection [14]–[16], bioinformatics [17], power system line outage detection [18], and sensor networks [19]–[22].

In many applications, the signal of interest x_t may be high dimensional. For example, x_t may consist of observations from many correlated sensors. Due to the large number of sensors in the network, bandwidth and power constraints

T. S. Lau and W. P. Tay are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail: TLAU001@e.ntu.edu.sg, wptay@ntu.edu.sg).

prevent us from observing the entire network at any time instance, and we may only obtain sensor readings from a small subset of sensors at any time instance [23], [24]. While it may seem optimal to observe the maximum number of sensors allowed by the network, this sampling policy may not be feasible due to power and communication bandwidth considerations. Furthermore, the action of switching from one subset to another subset of sensors also incurs power and communication costs. In this paper, we consider both of these costs collectively as the *observation-switching cost*, and we study the problem of QCD while maintaining an average observation-switching cost (AOSC) constraint. To be more general, we consider the case where the signal can only be observed using an action selected from a set of permissible actions with observation-switching costs associated with the sequence of actions chosen. We assume that the pre- and post-change distributions as well as their conditional distributions given the actions are known to the observer. The objective is to design a sampling policy together with a stopping time that satisfies both the QCD false alarm and AOSC requirements. To solve the QCD problem, we propose a sampling policy coupled with a generalized likelihood ratio (GLR) Cumulative Sum (CuSum) stopping time. For open-loop policies with finite window size, we show that the GLR CuSum stopping time is asymptotically optimal with a properly designed sampling policy and formulate the design of the sampling policy as a quadratic programming problem. For observation-dependent policies, we propose a 2-threshold stopping time, prove that it satisfies the AOSC and ARL constraints and demonstrate its performance empirically.

A. Related Work

Existing works in QCD where the signal is not entirely available to the decision maker or the fusion center can be categorized into three main categories. In the first category, the papers [25]–[28] consider the problem of distributed or decentralized QCD where each node observes and processes its signal locally, with some memory of its previous messages, before sending a message to the fusion center. The authors of [26] consider the problem where each sensor only has access to the local information at that node and would process the signal to send a quantized message to the fusion center for further processing.

The second category of papers [29]–[33] consider the QCD problem where the number of observations made during the

pre-change regime is controlled, and a control policy that determines whether a given observation is made. In [33], the authors developed a data-efficient scheme that allows for optional on-off sampling of the observations in the case where either the post-change family of distributions is finite, or both the pre- and post-change distribution belong to a one parameter exponential family.

In the third category, the papers [34]–[36] consider QCD where the observer only has access to compressed or incomplete measurements. The authors of [34] study the problem of sequential change point detection where a randomly generated linear projection is used to reduce the dimensions of a high dimensional signal for the purpose of QCD. In [36], the authors consider QCD with an observation-dependent control of the actions where the current nodes to observe is determined by the maximal likelihood estimate of the post change hypothesis. In [37], we discussed the QCD problem where the observer is only able to obtain a partial observation of the signal through an action with an open-loop control of the actions.

In the fourth category, the papers [38]–[41] considers the QCD problem with observational scheduling considerations. In these papers, it is assumed that there are multiple streams of observations and the cost associated with obtaining observations and the information quality of the streams differ from stream to stream. The QCD problem is to design a stopping time together with a control policy that determines the sequence of observations to perform such that the average cost of observations is controlled.

Unlike the papers mentioned above, in this paper, we provide a general framework by considering randomized decision rules to select the current actions. We also do not give a fixed cost to the sampling of observation. Instead, we consider a more general approach where we use a set of permissible actions to model the practical sampling constraints and a cost is associated to the switching of actions to model observation-switching costs. In this paper, we consider the case where the decision maker is given a finite set of pre-defined actions, and the observed signal is a function of the action and the full signal. We also do not make any assumptions about the pre- and post-change distributions.

B. Our Contributions

To the best of our knowledge, there are no existing work that considers the QCD problem while taking observation-switching costs into account. In this paper, we consider the problem QCD while maintaining an AOSC constraint. The objective is to design a sampling policy together with a stopping time that satisfies both the QCD average run length (ARL) and AOSC requirements. Our main contributions are as follows:

1) We formulate the QCD problem with an AOSC constraint.
2) We propose a condition on an open-loop policy of window size W where the derivation of closed-form expressions for the AOSC and the asymptotic worst-case average detection delay (WADD) of the GLR CuSum stopping time is possible.

3) We prove the existence of an open-loop policy of window size W satisfying our proposed condition for which the GLR CuSum is asymptotically optimal.
4) Using these closed-form expressions, we formulate the design of the open-loop policy of window size W as a quadratic programming problem with an additional combinatorial constraint and prove that for $1 \leq W < \infty$, to obtain an asymptotically optimal policy, it is sufficient to consider only policies with window size 1.

The rest of this paper is organized as follows. In Section II, we present our signal model and problem formulation. In Section III, we present properties of the AOSC for open-loop sampling policies. In Section IV, we present the GLR CuSum stopping time and formulate the design of the open-loop sampling policy as a quadratic programming problem. We present algorithms to solve the open-loop policy design problems in Section V. Numerical results are presented in Section VI. We conclude in Section VII.

Notations: The operator $\mathbb{E}_p$ denotes mathematical expectation with respect to (w.r.t.) pdf p, and $X \sim p$ means that the random variable X has distribution with pdf p. The Kullback-Leibler (KL) divergence between the distributions with pdf P and Q is denoted as $D(P \| Q)$. The pdf of a Gaussian distribution with mean μ and covariance Σ is denoted as $\mathcal{N}(\mu, \Sigma)$. Almost-sure convergence under the probability measure $\mathbb{P}$ is denoted as $\xrightarrow[\text{a.s.}]{\mathbb{P}}$. We use $\mathbf{1}_E$ as the indicator function of the set E, and $\mathbb{N}$ to denote the set of positive integers. We use $\mathbb{R}$ to denote the set of real numbers and $\mathbb{R}^+$ to denote the set of positive real numbers. We also use the notation $a^{k:t}$ to denote the sequence $(a_k, a_{k+1}, \ldots, a_t)$. For $\alpha = (a_1, a_2, \ldots, a_W)$, we use the notation $\alpha[j] = a_j$ to denote its j-th entry. For a probability transition matrix T, we use the notation $T[i, j]$ to denote the probability of moving to a state j given that it is currently at state i. For a probability mass function f, $\operatorname{supp}(f)$ denotes the support of f.

II. PROBLEM FORMULATION: QUICKEST CHANGE DETECTION WITH A COST FOR SWITCHING ACTIONS

Let p_0 and $p_1, \ldots, p_M$ be $M + 1$ distinct pdfs on $\mathbb{R}^N$, and $X_1, X_2, \ldots$ be a sequence of vector-valued random variables satisfying the following:

$$\begin{cases} X_t \sim p_0 \text{ i.i.d. for all } t < \nu, \\ X_t \sim p_m \text{ i.i.d. for all } t \geq \nu. \end{cases} \tag{1}$$

where $\nu \geq 0$ and $m \in \{1, \ldots, M\}$ are unknown but deterministic constants. The QCD problem is to detect the change in distribution as quickly as possible by observing $X_1, X_2, \ldots$, while keeping the false alarm rate low.

In this paper, we assume that the observer is only able to obtain a partial observation (A_t, Y_t) of X_t, where $Y_t \triangleq A_t(X_t)$ is a function of the random variable X_t under the action A_t at each time t. Let $\mathcal{A}$ be the collection of permissible actions. We assume that the set $\mathcal{A} = \{1, 2, \ldots, |\mathcal{A}|\}$ is finite. We also assume that at each time t, under the pdf p_m, the observation Y_t is conditionally independent of $Y_1, \ldots, Y_{t-1}$ and $A_1, \ldots, A_{t-1}$ given the action A_t. We further assume that for each $m \in \{1, \ldots, M\}$ there exists an action $A \in \mathcal{A}$ such

that p_0 and p_m are distinguishable under the action A. Some examples of $\mathcal{A}$ that arise in practical applications include:

1) Network Sampling. The set of rank $n < N$ transformations with

$$\mathcal{A} = \left\{ A_{\mathbf{L}} : \mathbb{R}^N \to \mathbb{R}^n \; \middle| \; \begin{array}{l} A_{\mathbf{L}}(X) = \mathbf{L}X, \\ \mathbf{L} = [e_{i_1}, \ldots, e_{i_n}]^T, \\ i_1 < i_2 < \ldots < i_n \end{array} \right\}$$

where e_i is an $N \times 1$ column vector with all zeros except a 1 at the i-th position.

2) 1-bit quantization. A set of 2^n functions on $\mathbb{R}^+$ with

$$\mathcal{A} = \left\{ A_\phi : \mathbb{R}^+ \to \{0,1\} \; \middle| \; \begin{array}{l} A_\phi(x) = \mathbf{1}_{\{x \geq \phi\}}, \\ \phi \in \{1,2,3,\ldots,2^n\} \end{array} \right\}$$

where n is a positive integer, $\mathbf{1}_A$ is the indicator function of the set A and ϕ is the quantization threshold.

In our sequential change detection problem, we obtain observations $(A_1 = a_1, Y_1 = y_1), (A_2 = a_2, Y_2 = y_2), \ldots$ sequentially and aim to detect the change in pdf from p_0 to p_m for some fixed $m \in \{1,...,M\}$ as quickly as possible. This is determined by a stopping time. A sampling policy is used to determine the action used to obtain the next observation. An observation-dependent sampling policy determines the current action based on both the previous actions and observations while an open-loop sampling policy determines the current action based on only the previous actions. Since an observation-dependent sampling policy has access to more information, an optimized observation-dependent policy is expected to outperform an open-loop policy.

However, an observation-dependent policy may not be suitable for some applications due to physical constraints. One such class of applications involves time-synchronized sensor networks with no feedback mechanism from the fusion center to the sensors. In this case, each sensor in a time-synchronized network may have several modes (temperature, humidity, pressure, etc.) of obtaining observations. Without a feedback mechanism from the fusion center, the only way for the sensor to select a mode of observation is to use an open-loop policy. Another class of applications is when the sensor network has a transmission delay that is significantly larger the sampling rate of the sensors. When an observation-dependent policy is applied, the sampling rate of this network is greatly reduced due to the time taken to communicate actions to the sensors. An open-loop policy does not suffer from the same problem as the sequence of actions does not depend on the observations, and thus can be pre-generated and made available to the sensors. Geo-stationary satellites usually have a communication round-trip time of between $600 - 800$ms while carrying sensors capable of sampling at rates larger than 1kHz. An observation-dependent policy would only be able to obtain a sample from the geo-stationary satellites at sampling frequency of less than 2Hz since it has to wait for the observations to be transmitted from the satellites to the ground station before deciding on the next action. An open-loop policy does not need to wait for the observations to decide the next action as the policy has been pre-determined. Hence, an open-loop policy would be able to sample the network of satellite at rates larger than 1kHz with a delay of between

$300 - 400$ms from the time the observation is obtained to the time the observation is processed by the stopping time. In this example, it is possible that a stopping time using an open-loop policy out-performs one that uses an observation-dependent policy due to large differences in the amount of observations available. In both classes, a long sequence of actions may be pre-generated and programmed into the sensors and the fusion center so that there is no need to communicate the sequence of actions during test time. In this paper, we consider only open loop-sampling policies. The interested reader may refer to the extended version of this paper [42] for a discussion on observation-dependent sampling policies.

Definition 1. *A policy π is a sequence of functions $(\rho_t)_{t \in \mathbb{N}}$, where ρ_t is a randomized function that determines the action A_{t+1} using observations $((A_1, Y_1), (A_2, Y_2), \ldots, (A_t, Y_t))$ up to time t.*
An open-loop policy π of window size W is a policy $(\rho_t)_{t \in \mathbb{N}, t \geq W}$, where $\rho_t = \rho$ for a fixed randomized function ρ that determines the action A_{t+1} based on W past actions $(A_{t-W+1}, A_{t-W+2}, \ldots, A_t)$.

It can be shown that an open-loop policy π of window size W is equivalent to a Markov chain of order W on $\mathcal{A}$ with initial distribution q and probability transition matrix T. Thus, we use the notation $\pi = (q, T)$ to represent an open-loop policy of window size W for the rest of this paper. If the change point is at ν and post-change distribution has pdf p_m, we let $\mathbb{P}_{\nu,m}$ and $\mathbb{E}_{\nu,m}$ be the probability measure and mathematical expectation, respectively. We let $\mathbb{P}_\infty$ and $\mathbb{E}_\infty$ denote the probability measure and mathematical expectation when there is no change.

For a stopping time τ and a policy π, we quantify its detection delay using the worst case average detection delay (WADD) as proposed by Lorden [43]:

$$\text{WADD}(\tau, \pi) = \max_{1 \leq m \leq M} \text{WADD}_m(\tau, \pi), \qquad (2)$$

where

$$\begin{aligned} &\text{WADD}_m(\tau, \pi) \\ &= \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathbb{E}_{\nu,m} \left[(\tau - \nu + 1)^+ \; \middle| \; A^{1:(\nu-1)}, Y^{1:(\nu-1)} \right], \end{aligned}$$
$$(3)$$

and its ARL to false alarm as $\text{ARL}(\tau, \pi) = \mathbb{E}_\infty[\tau]$.

In order to take the observation-switching costs of a policy π into consideration, we let $\mathcal{C}$ be a $|\mathcal{A}| \times |\mathcal{A}|$ matrix where its (i,j)-th entry $\mathcal{C}[i,j]$ denotes the cost of switching from action i to action j. Inspired by a similar cost first proposed in [30], we define the AOSC of the policy π as

$$\text{AOSC}(\pi) = \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}_\infty \left[\sum_{t=2}^{n+1} \mathcal{C}[A_{t-1}, A_t] \right]. \qquad (4)$$

Formally, our quickest change detection with AOSC constraint can be formulated as a minimax problem: find a sampling

policy π and a stopping time τ to

$$
\begin{aligned}
\min_{\tau,\pi} \quad & \text{WADD}(\tau,\pi) \\
\text{s.t.} \quad & \text{ARL}(\tau,\pi) \geq \gamma, \\
& \text{AOSC}(\pi) \leq \alpha_{\text{AOSC}},
\end{aligned} \tag{5}
$$

for some given thresholds α_{AOSC} and γ.

For a fixed policy π, the GLR CuSum [33] stopping time τ_π w.r.t. the observed sequence $(A_t = a_t, Y_t = y_t)_{t \in \mathbb{N}}$ is defined as:

$$
\tau_\pi = \inf \{ t \mid S(t,\pi) > \log(M\gamma) \}, \tag{6}
$$

$$
S(t,\pi) = \max_{1 \leq m \leq M} \max_{1 \leq i \leq t+1} \sum_{j=i}^{t} \log \frac{p_m(y_j|a_j)}{p_0(y_j|a_j)}, \tag{7}
$$

where $\gamma \geq 0$ is a pre-selected threshold and $p_m(y_j|a_j)$ is the conditional pdf of $Y_j = y_j$ given the action $A_j = a_j$ under the distribution with pdf p_m. The GLR CuSum stopping time τ_π can be re-written as

$$
\tau_\pi = \min_{1 \leq m \leq M} \tau_{m,\pi}, \quad S(t,\pi) = \max_{1 \leq m \leq M} S_m(t,\pi), \tag{8}
$$

$$
\tau_{m,\pi} = \inf \{ t \mid S_m(t,\pi) > \log(M\gamma) \}, \tag{9}
$$

$$
S_m(t,\pi) = \left(S_m(t-1,\pi) + \log \frac{p_m(y_t|a_t)}{p_0(y_t|a_t)} \right)^+ \text{ for } t > 0, \tag{10}
$$

$$
S_m(0,\pi) = 0, \quad \text{for } m \in \{1,\dots,M\} \tag{11}
$$

where $x^+ \triangleq \max(x,0)$. We note that for $m \in \{1,\dots,M\}$, $S_m(t,\pi)$ is the CuSum statistic corresponding to the post-change pdf p_m and policy π. Thus, the GLR CuSum statistic $S(t,\pi)$ is the maximum of the CuSum statistics for each of the post-change pdf p_m.

We discuss results pertaining to open-loop policies of window size W in Sections III to V.

III. PROPERTIES OF THE AOSC FOR OPEN-LOOP POLICIES OF WINDOW SIZE W

In this section, we present results regarding the AOSC of an open-loop policy of window size W. When the window size W of the open loop policy is zero, the actions used to observe the signal are generated i.i.d. with respect to the distribution q. In this case, the observations $\{(A_t, Y_t) = (a_t, y_t)\}_{t \in \mathbb{N}}$ are also generated i.i.d. However, when the window size W of the open-loop policy is positive, unlike the former case, it is possible that the actions $\{A_t\}_{t \in \mathbb{N}}$ and observations $\{(A_t, Y_t)\}_{t \in \mathbb{N}}$ are not generated i.i.d. We denote the joint probability density function of $(A^{1:t}, Y^{1:t})$ under the distribution with pdf p_m as

$$
\begin{aligned}
& p_m(a^{1:t}, y^{1:t}) \\
& = q(a^{1:W}) \prod_{j=1}^{t} p_m(y_j|a_j) \prod_{k=W+1}^{t} p_T(a_k \mid a^{k-W:k-1}), \quad (12)
\end{aligned}
$$

where p_T is the conditional probability mass function of A_k given $A^{k-W:k-1}$ induced by the probability transition matrix T, W is the window size of the policy.

An open-loop policy $\pi = (q,T)$ with window size W can also be written as a Markov chain $\pi' = (q',T')$ of order 1

where T' satisfies $T'[\alpha,\beta] = 0$ whenever $\beta[i] \neq \alpha[i-1]$ for some $i \in \{2,\dots,W\}$ and $\alpha, \beta \in \mathcal{A}^W$. For the rest of this paper, we switch between either representation of an open-loop policy π to simplify the computations in the proofs. We denote the observation-switching costs associated with the latest two actions $\mathcal{C}[\alpha[W-1], \alpha[W]]$, from $\alpha[W-1]$ to $\alpha[W]$, as $\mathcal{C}_\alpha$.

For the rest of this section, we present results that relate the AOSC of open-loop polices with different initial distributions but equal probability transition matrices. First, we recall a relation between the average number of visits and the stationary distributions of a Markov chain. Let $N_t(\alpha;\beta)$ denote the number of times, up to time t, that the state α is visited given that the initial state is β. Since $\mathcal{A}^W$ is finite, the Markov chain defined by the transition matrix T has at least one recurrence class. Let R be the number of recurrence classes and U be the number of transient states. By the Ergodic Theorem for finite state Markov chains [44], for a finite state Markov Chain with R recurrent classes $\{\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_R\}$, there exists R stationary distributions $\xi_1, \dots \xi_R$ where $\xi_r[\alpha] = 0$ if the state $\alpha \notin \mathcal{R}_r$, and for recurrent states $\beta \in \mathcal{R}_r$, we have

$$
\lim_{t \to \infty} \frac{N_t(\alpha;\beta)}{t} = \xi_r[\alpha] \text{ a.s.}, \tag{13}
$$

$$
\lim_{t \to \infty} \frac{\mathbb{E}\left[N_t(\alpha;\beta)\right]}{t} = \xi_r[\alpha], \tag{14}
$$

for any state $\alpha \in \mathcal{A}^W$ and $r = 1, \dots, R$. For transient states β,

$$
\lim_{t \to \infty} \frac{\mathbb{E}\left[N_t(\alpha;\beta)\right]}{t} = \sum_{r=1}^{R} f_{\beta,r} \xi_r[\alpha], \tag{15}
$$

where $f_{\beta,r}$ is the first-passage probability of initializing at state β and entering the recurrence class $\mathcal{R}_r$ before any other recurrence classes.

For any state β, denoting ξ as the vector of expected proportion of visits to each of the states initializing at state β such that $\xi[\alpha] = \lim_{t \to \infty} \frac{\mathbb{E}[N_t(\alpha;\beta)]}{t}$, we can see that ξ is a stationary distribution of the probability transition matrix T as it is a convex linear combination of stationary distributions.

Definition 2. *For any initial distribution q, the expected proportion of visits to each of the state, $\overline{q}$, is defined as*

$$
\overline{q}(\alpha) = \lim_{t \to \infty} \mathbb{E}\left[\frac{1}{t} \sum_{j=1}^{t} \mathbf{1}_{\{A_j = \alpha\}} \right].
$$

Thus, for any initial distribution q, $\overline{q}$ is a stationary distribution of T. In the next lemmas, we see that the AOSC and asymptotic log likelihood ratios depend only on T and $\overline{q}$.

Lemma 1. *Let $\pi_1 = (q,T)$ be an open-loop policy of finite window size W and $\pi_2 = (\overline{q}, T)$, then we have*

$$
\text{AOSC}(\pi_1) = \text{AOSC}(\pi_2) \tag{16}
$$

Proof: The proof is based on standard Markov chain theory and is provided for completeness in the supplementary material [42]. ∎

IV. ASYMPTOTIC PROPERTIES OF GLR CUSUM FOR OPEN-LOOP POLICIES OF WINDOW SIZE W

Next, we present some asymptotic properties of $S(t, \pi)$ and τ_π for an open-loop policy π. In this paper, we use $\asymp$ to denote the notion of asymptotic equivalence [45]:

$$f \asymp g \quad \text{if and only if} \quad \lim_{x \to \infty} \frac{f(x)}{g(x)} = 1. \quad (17)$$

Definition 3. *The* asymptotic ARL-WADD trade-off rate I of the GLR CuSum stopping time τ_π is defined as

$$I = \liminf_{\gamma \to \infty} \frac{\log \mathrm{ARL}(\tau_\pi(\gamma))}{\mathrm{WADD}(\tau_\pi(\gamma), \pi)}.$$

When the signal is generated i.i.d. before and after the change point, the asymptotic trade-off rate for the GLR CuSum stopping time is well studied and has a nice closed-form expression in terms of the KL divergence between the pre-change and post-change distributions [33]. In our case, even though the signal $X_1, X_2, \ldots$ is originally i.i.d. before and after the change-point, any sampling procedure that switches action would inevitably result in non-i.i.d. observations. We let

$$\Lambda_m(k, t) = \log \frac{p_m(A^{k:t}, Y^{k:t})}{p_0(A^{k:t}, Y^{k:t})} = \sum_{i=k}^{t} \log \frac{p_m(Y_i|A_i)}{p_0(Y_i|A_i)}. \quad (18)$$

In the next two Lemmas, we present results regarding Λ_m when the open-loop policy $\pi = (q, T)$ satisfies the property that $\overline{q}$ has support in only one recurrence class of T. The property that $\overline{q}$ has support in only one recurrence class of T plays an important role in solving Problem (5). It allows us to quantify the performance of the GLR CuSum using a closed form expression. This is key to obtaining asymptotically optimal policies using numerical optimization tools.

Lemma 2. *For any open-loop policy $\pi = (q, T)$ of finite window size W where $\overline{q}$ has support in only one recurrence class $\mathcal{R}$, and any $m \in \{1, \ldots, M\}$ and change-point $\nu < \infty$, we have*

$$\mathbb{P}_{\nu,m} \left(\lim_{t \to \infty} \frac{1}{t} \Lambda_m(\nu, \nu + t - 1) = I_{m,\pi} \right) = 1, \quad (19)$$

with

$$I_{m,\pi} = \sum_{\alpha \in \mathcal{R}} \overline{q}(\alpha) D(p_m(\,\cdot\,|\,\alpha[W]) \,\|\, p_0(\,\cdot\,|\,\alpha[W])). \quad (20)$$

Proof: See Appendix A. ∎

Lemma 3. *Let $\pi = (q, T)$ be an open-loop policy of finite window size W where $\overline{q}$ has support in only one recurrence class. For any $\epsilon > 0$ and $m \in \{1, \ldots, M\}$, we have*

$$\lim_{t \to \infty} \mathbb{P}_{\nu,m} \left(\left| \frac{1}{t} \Lambda_m(\nu, \nu + t - 1) - I_{m,\pi} \right| > \epsilon \right) = 0, \quad (21)$$

for $0 \le \nu < \infty$, and

$$\sup_{0 \le \nu < \infty} \operatorname{ess\,sup} \mathbb{P}_{\nu,m} \left(\frac{1}{t} \max_{0 \le j < t} \Lambda_m(\nu, \nu + j) \right.$$
$$\left. > (1 + \epsilon) I_{m,\pi} \,\middle|\, A^{1:\nu-1}, Y^{1:\nu-1} \right) \to 0 \quad \text{as } t \to \infty. \quad (22)$$

Proof: See Appendix B. ∎

For a fixed open-loop policy $\pi = (q, T)$ such that the expected proportion of visits to each of the states, $\overline{q}$, has support in one recurrence class, we apply Lemma 3 together with [1, Theorem 8.2.3] to obtain the following proposition.

Proposition 1. *For a fixed open-loop policy $\pi = (q, T)$ such that $\overline{q}$ has support in one recurrence class, we have the following asymptotic ARL-WADD trade-off for any $1 \le m \le M$:*

$$\mathrm{ARL}(\tau_{m,\pi}, \pi) \ge M\gamma, \ \mathrm{WADD}_m(\tau_{m,\pi}, \pi) \le \frac{\log \gamma}{I_{m,\pi}} (1 + o(1)), \quad (23)$$

as $\gamma \to \infty$ where a function $g(\gamma) = o(1)$ if and only if $\lim_{\gamma \to \infty} g(\gamma) = 0$.

Thus, when the signal $\{X_t \ : \ t \in \mathbb{N}\}$ is sampled using the open-loop policy π such that $\overline{q}$ has support in one recurrence class, using similar techniques from [2, Theorem 6.16] together with Proposition 1, we know that the GLR CuSum stopping time τ_π gives us a stopping time satisfying $\mathrm{ARL}(\tau_\pi, \pi) \ge \gamma$ and τ_π is asymptotically optimal for the following problem as $\gamma \to \infty$:

$$\min_\tau \quad \mathrm{WADD}(\tau, \pi) \quad \text{s.t.} \quad \mathrm{ARL}(\tau, \pi) \ge \gamma. \quad (24)$$

It should be noted that, for open-loop policies $\pi = (q, T)$ with $\overline{q}$ having support in more than one recurrence class, Proposition 1 does not hold, making it difficult to use [2, Theorem 6.16] to derive the asymptotic ARL-WADD trade-off rate and show the asymptotic optimality of the GLR CuSum. Furthermore, in the next proposition, we show that an open-loop policy π with $\overline{q}$ having support in multiple recurrence classes is suboptimal in terms of AOSC and WADD.

Proposition 2. *Let the open-loop policy $\pi = (q, T)$ be such that $\overline{q}$ has support in multiple recurrence classes. Then, there exists an open-loop policy $\pi' = (q', T)$ where $\overline{q'}$ has support in only one recurrence class such that for any stopping time τ,*

$$\mathrm{AOSC}(\pi') \le \mathrm{AOSC}(\pi) \text{ and } \mathrm{WADD}(\tau, \pi') \le \mathrm{WADD}(\tau, \pi).$$

Proof: See Appendix C. ∎

Using this proposition, we obtain a result regarding asymptotically optimal solutions of Problem (5).

Theorem 1. *When the signal is sampled using an open-loop policy $\pi = (q, T)$ with q having support in one recurrence class, satisfying $\mathrm{AOSC}(\pi) \le \alpha_{AOSC}$, the GLR CuSum τ_π is asymptotically optimal with the asymptotic WADD-ARL trade-off given as $\min_m I_{m,\pi}$.*

There exists an open-loop policy $\pi = (q', T')$ with $\overline{q'}$ having support in one recurrence class such that (τ_π, π) is asymptotically optimal for Problem (5) as $\gamma \to \infty$, where τ_π is the GLR CuSum stopping time.

Proof: See Appendix D. ∎

Thus, the additional constraint that the open-loop policy $\pi = (q, T)$ satisfies the condition that $\overline{q}$ has support in one recurrence class does not affect the asymptotic optimality of the GLR CuSum for Problem (5). Furthermore, for policies satisfying this constraint, we are able to obtain a closed-form

expression for the ARL-WADD trade-off rate. The closed-form expressions are important as numerical optimization tools are used to optimize the ARL-WADD trade-off for the GLR CuSum.

Using Theorem 1, the minimization in Problem (5), over the open-loop policy π and stopping time τ, can be decoupled.

Let π^* be an optimal solution to the following problem:

$$
\begin{aligned}
\min_{\pi} \quad & \max_{1 \leq m \leq M} I_{m,\pi}^{-1} \\
\text{s.t.} \quad & \text{AOSC}(\pi) \leq \alpha_{\text{AOSC}}, \\
& \text{supp}(\overline{q}) \subseteq \text{one recurrence class.}
\end{aligned}
\tag{25}
$$

By the argument above, (τ_{π^*}, π^*) is asymptotically optimal for Problem (5). We call Problem (25) the *open-loop policy design* problem.

V. OPTIMAL OPEN-LOOP POLICY OF WINDOW SIZE W

In this section, we investigate the open-loop policy design Problem (25) under the cases where the switching costs from one action to another are all equal or not.

A. Equal Switching Costs

In this subsection, we propose a method to solve the open-loop policy design problem in which the switching costs are constant, i.e., $\mathcal{C}[a,b] = c$, for a fixed $c \in \mathbb{R}$ and any $a, b \in \mathcal{A}$. First, we note that Problem (25) is feasible if and only if $\alpha_{\text{AOSC}} \geq c$. Furthermore, if $\alpha_{\text{AOSC}} \geq c$ then Problem (5) reduces to

$$
\begin{aligned}
\min_{\tau,\pi} \quad & \text{WADD}(\tau,\pi) \\
\text{s.t.} \quad & \text{ARL}(\tau,\pi) \geq \gamma,
\end{aligned}
\tag{26}
$$

where the AOSC constraint is automatically satisfied. Next, we show that for the case when all action-switching costs are equal, there exists a memoryless open-loop policy π (i.e., $W = 0$) for which the GLR CuSum τ_π achieves asymptotic optimality.

Proposition 3. *Suppose Problem (25) is feasible and (τ_π, π) is an asymptotically optimal solution for Problem (5). Then, there exists an open-loop policy π_0 with window size $W = 0$ such that*

$$
\text{WADD}(\tau_{\pi_0}, \pi_0) \asymp \text{WADD}(\tau_\pi, \pi) \quad \text{as } \gamma \to \infty.
\tag{27}
$$

Proof: See Appendix E. ∎

From Proposition 3, to solve the open-loop policy design problem Problem (25) for some $W \in \mathbb{N}$, it suffices to solve Problem (25) for the case where $W = 0$.

When $W = 0$, Problem (25) becomes

$$
\begin{aligned}
\min_{q} \quad & \max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1} \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \\
& q(a) \geq 0 \quad \text{for all } a \in \mathcal{A}.
\end{aligned}
\tag{28}
$$

Proposition 4. *For the optimization problem*

$$
\begin{aligned}
\min_{q,z} \quad & z \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \\
& q(a) > 0 \quad \text{for all } a \in \mathcal{A}, \\
& \sum_{a \in \mathcal{A}} q(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) + z \geq 0 \\
& \text{for all } m \in \{1, \ldots, M\},
\end{aligned}
\tag{29}
$$

the following holds

(i) *If (q^*, z^*) is an optimal solution to Problem (29) then q^* is an optimal solution to Problem (28).*

(ii) *If q^* is an optimal solution to Problem (28) then there exists x^* such that (q^*, x^*) is an optimal solution to Problem (29).*

Proof: See Appendix F. ∎

By Proposition 4, we can solve Problem (28) by solving the linear optimization problem, Problem (29). Let q_0 be the solution for Problem (29) and T_0 be the probability transition matrix with rows equal to q_0. Using similar techniques from [2, Theorem 6.16] together with Proposition 1, we know that the GLR CuSum algorithm with $\pi_0 = (q_0, T_0)$ as the open-loop policy gives us a stopping time satisfying $\text{ARL}(\tau_{q^*}) \geq \gamma$ with asymptotically optimal ARL-WADD trade-off as γ tends to infinity.

B. Unequal Switching Costs

In this subsection, we propose a method to solve the policy design problem in which the switching costs are not all equal. First, we present a proposition regarding the structure of asymptotically optimal solutions of Problem (5).

Proposition 5. *Suppose (τ_π, π) is an asymptotically optimal solution for Problem (5) with finite window size of at least 1. There exists an open-loop policy π_1 with window size $W = 1$ such that $\text{AOSC}(\pi_1) = \text{AOSC}(\pi)$ and*

$$
\text{WADD}(\tau_{\pi_1}, \pi_1) \asymp \text{WADD}(\tau_\pi, \pi) \quad \text{as } \gamma \to \infty.
\tag{30}
$$

Proof: See Appendix G. ∎

From Proposition 5, the open-loop policy design problem for window size $W \in \mathbb{N}$ can be reduced to a problem of window size $\min(W, 1)$. Thus, we only need to study the cases where $W = 0$ or $W = 1$. It should be noted that Proposition 5 does not hold when $W = \infty$. An example is provided in Appendix H. Proposition 5 holds primarily because of the Markovian structure of the sampling policy together with the form of AOSC function.

In the following, we present algorithms to solve Problem (25) for each of these cases.

1) Window size $W = 0$: Using a similar argument from Section V-A, we can see that for any optimal open-loop policy

$\pi = (q, T)$ of window size W, we have $\overline{q} = q$ and T has only one recurrence class. Thus, we have

$$I_{m,\pi} = \sum_{a \in \mathcal{A}} q(a) D(p_m(\cdot \mid a) \| p_0(\cdot \mid a)), \tag{31}$$

$$\text{AOSC}(\pi) = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{A}} q(a)q(b)\mathcal{C}[a, b], \tag{32}$$

and Problem (25) becomes

$$
\begin{aligned}
\min_{q} \quad & \max_{1 \le m \le M} \left(\sum_{a \in \mathcal{A}} q(a) D(p_m(\cdot \mid a) \| p_0(\cdot \mid a)) \right)^{-1} \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \\
& q(a) \ge 0 \quad \text{for all } a \in \mathcal{A}, \\
& \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{A}} q(a)q(b)\mathcal{C}[a, b] \le \alpha_{\text{AOSC}}.
\end{aligned}
\tag{33}
$$

Using the same argument from Problem (28), we can introduce a new variable z to obtain a linear cost function

$$
\begin{aligned}
\min_{q,z} \quad & z \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \\
& q(a) > 0 \quad \text{for all } a \in \mathcal{A}, \\
& \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{A}} q(a)q(b)\mathcal{C}[a, b] \le \alpha_{\text{AOSC}}, \\
& \sum_{a \in \mathcal{A}} q(a) D\left(p_0(\cdot \mid a) \| p_m(\cdot \mid a)\right) + z \ge 0 \\
& \quad \text{for all } m \in \{1, \dots, M\}.
\end{aligned}
\tag{34}
$$

This is a quadratically constrained quadratic program (QCQP), and we may assume that $\mathcal{C}$ is symmetric without loss of generality. However, without additional assumptions, the problem is NP-hard.

First, we discuss some special cases where the global optimum can be obtained easily. When $\mathcal{C}$ is positive semi-definite, Problem (34) is a convex programming problem. A convex program solver [46], [47] can be used to obtain globally optimal solutions. For the case where there are only cost of observations rather than cost of switching (i.e., $\mathcal{C}[a, b] = h(b)$ for some function $h : \mathcal{A} \to \mathbb{R}$), the quadratic constraint in Problem (34) reduces to

$$\sum_{b \in \mathcal{A}} q(b)h(b) \le \alpha_{\text{AOSC}}. \tag{35}$$

In this case, Problem (34) becomes a linear programming problem, which can be solved by a linear program solver [46].

For the general case, we use the IRM algorithm [48] to obtain a locally optimal solution. In order to apply the IRM algorithm, we have to first convert Problem (34) into a rank-constrained convex optimization problem. We first rewrite the quadratic constraint as

$$\sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{A}} q(a)q(b)\mathcal{C}[a, b] = q^T \mathcal{C} q = \text{Tr}(\mathcal{C}qq^T) \le \alpha_{\text{AOSC}}.$$

By introducing a new $|\mathcal{A}| \times |\mathcal{A}|$ variable Q such that $Q = qq^T$, the quadratic constraint becomes a linear constraint $\text{Tr}(\mathcal{C}Q) \le$

α_{AOSC}. To ensure that $Q = qq^T$ holds, we require that $Q \succeq 0$, $Q\mathbb{1} = q$ and $\text{rank}(Q) = 1$ where $\mathbb{1}$ is a $|\mathcal{A}| \times 1$ vector of ones. Hence, Problem (34) is equivalent to

$$
\begin{aligned}
\min_{Q,q,z} \quad & z \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \ q(a) > 0 \quad \text{for all } a \in \mathcal{A}, \\
& \text{Tr}(\mathcal{C}Q) \le \alpha_{\text{AOSC}}, \\
& Q\mathbb{1} = q, \ Q \succeq 0 \quad \text{and} \quad \text{rank}(Q) = 1, \\
& \sum_{a \in \mathcal{A}} q(a) D\left(p_0(\cdot \mid a) \| p_m(\cdot \mid a)\right) + z \ge 0 \\
& \quad \text{for all } m \in \{1, \dots, M\}.
\end{aligned}
\tag{36}
$$

We note that Problem (36) becomes a convex programming problem when the constraint $\text{rank}(Q) = 1$ is ignored. We are now ready to present the IRM algorithm [48]. Fix $\omega > 1$.

First, we solve the convex problem

$$
\begin{aligned}
\min_{Q,q,z} \quad & z \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \\
& q(a) > 0 \quad \text{for all } a \in \mathcal{A}, \\
& \text{Tr}(\mathcal{C}Q) \le \alpha_{\text{AOSC}}, \ Q\mathbb{1} = q, \ Q \succeq 0, \\
& \sum_{a \in \mathcal{A}} q(a) D\left(p_0(\cdot \mid a) \| p_m(\cdot \mid a)\right) + z \ge 0 \\
& \quad \text{for all } m \in \{1, \dots, M\},
\end{aligned}
\tag{37}
$$

to obtain a solution (Q_0, q_0, z_0) and let $Q_0 = VDV^T$ be the eigen-decomposition of Q_0. Let V_0 be the eigenvectors corresponding to the $|\mathcal{A}| - 1$ smallest eigenvalues of Q_0.

At each step k, we solve the following convex problem:

$$
\begin{aligned}
\min_{Q,q,z,r} \quad & z + \omega^k r \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}} q(a) = 1, \\
& q(a) > 0 \quad \text{for all } a \in \mathcal{A}, \\
& \text{Tr}(\mathcal{C}Q) \le \alpha_{\text{AOSC}}, \ Q\mathbb{1} = q, \ Q \succeq 0, \\
& rI_{n-1} - V_{k-1}^T Q V_{k-1} \succeq 0, \\
& \sum_{a \in \mathcal{A}} q(a) D\left(p_0(\cdot \mid a) \| p_m(\cdot \mid a)\right) + z \ge 0 \\
& \quad \text{for all } m \in \{1, \dots, M\},
\end{aligned}
\tag{38}
$$

to obtain a solution (Q_k, q_k, z_k, r_k) and let V_k be the eigenvectors corresponding to the $|\mathcal{A}| - 1$ smallest eigenvalues of Q_k. We iterate until $r_k < \epsilon$, where ϵ is a small threshold chosen as a stopping criterion. Following similar methods from [48], it can be shown that $r_k \to 0$ at a linear rate and that q_k converges to a locally optimal solution for Problem (36) if Problem (38) is feasible for all k.

2) Window size $W = 1$: Unlike the case when $W = 0$, not every distribution q is a stationary distribution of T.

Furthermore, when $W > 0$, it is possible that more than one recurrence class exists. In this case, Problem (25) becomes

$$\min_{T,q} \quad \max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q(a) D\left(p_0(\cdot \mid a) \parallel p_m(\cdot \mid a)\right)\right)^{-1}$$

$$\text{s.t.} \quad \sum_{a \in \mathcal{A}} q(a) = 1, \; q^T T = q^T,$$

$$q(a) > 0 \quad \text{for all } a \in \mathcal{A},$$

$$\sum_{b \in \mathcal{A}} T[a,b] = 1 \quad \text{for all } a \in \mathcal{A}, \tag{39}$$

$$0 \leq T[a,b] \leq 1 \quad \text{for all } a, b \in \mathcal{A},$$

$$\sum_{b \in \mathcal{A}} \sum_{a \in \mathcal{A}} T[a,b]\mathcal{C}[a,b]q(a) \leq \alpha_{\text{AOSC}},$$

$$\text{supp}(q) \subseteq \text{ one recurrence class.}$$

Using the same argument from Problem (28), Problem (39) is equivalent to

$$\min_{T,q,z} \quad z$$

$$\text{s.t.} \quad \sum_{a \in \mathcal{A}} q(a) = 1,$$

$$q(a) > 0 \quad \text{for all } a \in \mathcal{A}$$

$$\sum_{b \in \mathcal{A}} T[a,b] = 1 \quad \text{for all } a \in \mathcal{A}$$

$$0 \leq T[a,b] \leq 1 \quad \text{for all } a, b \in \mathcal{A} \tag{40}$$

$$q^T T = q^T,$$

$$\sum_{b \in \mathcal{A}} \sum_{a \in \mathcal{A}} T[a,b]\mathcal{C}[a,b]q(a) \leq \alpha_{\text{AOSC}},$$

$$\text{supp}(q) \subseteq \text{ one recurrence class,}$$

$$\sum_{a \in \mathcal{A}} q(a) D\left(p_m(\cdot \mid a) \parallel p_0(\cdot \mid a)\right) + z \geq 0,$$

$$\text{for all } m \in \{1, \ldots, M\}.$$

Problem (40) has two quadratic constraints $Tq = q$ and $\sum_{b \in \mathcal{A}} \sum_{a \in \mathcal{A}} T[a,b]\mathcal{C}[a,b]q(b) \leq \alpha_{\text{AOSC}}$. Thus, it is a QCQP with an additional combinatorial constraint that $\text{supp}(q)$ is contained in a single recurrence class of T. Even without the combinatorial constraint, finding the global optimal for Problem (40) would be difficult without additional assumptions on $\mathcal{C}$.

By considering a similar construction used in the proof of Proposition 5, we show that any open-loop policy $\pi_1 = (q, T)$ with window size 1 can be expressed as an open-loop policy $\pi_2 = (q_2, T_2)$ of window size 2 satisfying the following:

$$\sum_{b \in \mathcal{A}} q_2(a,b) = \sum_{c \in \mathcal{A}} q_2(c,a) \quad \text{for any } a \in \mathcal{A}, \tag{41}$$

$$T_2[(a,b),(c,d)] = \begin{cases} \frac{q_2(c,d)}{\sum_{d \in \mathcal{A}} q_2(c,d)} & \text{if } b = c, \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

Furthermore, any open-loop policy $\pi_2 = (q_2, T_2)$ with window size 2 that satisfies (41) and (42) can be expressed as an open-loop policy π_1 of window size 1. When we consider an open-loop policy $\pi_2 = (q_2, T_2)$ with window size 2 that satisfies (41) and (42), the constraint $T_2 q_2 = q_2$ becomes automatically satisfied and the AOSC constraint is linearized

to $\sum_{(a,b) \in \mathcal{A}^2} \mathcal{C}[a,b]q(a,b) \leq \alpha_{\text{AOSC}}$. Hence, Problem (40) is equivalent to the following problem

$$\min_{q,z} \quad z$$

$$\text{s.t.} \quad \sum_{(a,b) \in \mathcal{A}^2} q(a,b) = 1,$$

$$q(a,b) > 0 \quad \text{for all } a, b \in \mathcal{A},$$

$$\sum_{(a,b) \in \mathcal{A}^2} \mathcal{C}[a,b]q(a,b) \leq \alpha_{\text{AOSC}},$$

$$\sum_{b \in \mathcal{A}} q(a,b) = \sum_{c \in \mathcal{A}} q(c,a) \quad \text{for all } a \in \mathcal{A}, \tag{43}$$

$$\text{supp}(q) \subseteq \text{ one recurrence class,}$$

$$\sum_{(a,b) \in \mathcal{A}^2} q(a,b) D\left(p_m(\cdot \mid b) \parallel p_0(\cdot \mid b)\right) \geq -z,$$

$$\text{for all } m \in \{1, \ldots, M\}, \, a, b \in \mathcal{A},$$

which is a linear programming problem with an additional combinatorial constraint. In order to handle the combinatorial constraint, we can solve Problem (43) without the combinatorial constraint $\text{supp}(q) \subseteq$ one recurrence class. The solution obtained by solving Problem (43) without the combinatorial constraint will not have support in one recurrence class in general. However, if the solution has support in one recurrence class, it is an optimal solution to Problem (43). Alternatively, we can select a sufficiently small $\epsilon > 0$, and require that $q(a,b) > \epsilon$ for all $a, b \in \mathcal{A}$. The new constraint ensures that there is only one recurrence class for any feasible open-loop policy and thus, the recurrence class constraint is satisfied. The relaxed problem becomes

$$\min_{q,z} \quad z$$

$$\text{s.t.} \quad \sum_{(a,b) \in \mathcal{A}^2} q(a,b) = 1,$$

$$q(a,b) > \epsilon \quad \text{for all } a, b \in \mathcal{A},$$

$$\sum_{(a,b) \in \mathcal{A}^2} \mathcal{C}[a,b]q(a,b) \leq \alpha_{\text{AOSC}},$$

$$\sum_{a \in \mathcal{A}} q(a,b) = \sum_{c \in \mathcal{A}} q(c,b) \text{ for } b \in \mathcal{A}, \tag{44}$$

$$\sum_{(a,b) \in \mathcal{A}^2} q(a,b) D\left(p_m(\cdot \mid b) \parallel p_0(\cdot \mid b)\right) \geq -z,$$

$$\text{for all } m \in \{1, \ldots, M\},$$

which can be solved by a linear program solver [46].

VI. NUMERICAL RESULTS

In this section, we present numerical results to illustrate the relationship between the asymptotic ARL-WADD trade-off rate and the AOSC of the open-loop policy. We also compare our proposed sampling policy which takes observation-switching costs constraints into account, against observation-dependent polices which only take sampling cost constraints into account.

A. Applications to graph signals

In this subsection, we consider the QCD problem with AOSC based on partially observed graph signals under the various conditions discussed in this paper. We consider the problem of quickest detection of a rogue node in a graph. We assume that our graph $\mathcal{G}$ is a connected graph with N nodes. We model the graph signal [49] in the pre-change regime with a zero-mean Gaussian distribution with covariance $\Sigma = L^\dagger + \eta I$, where $L^\dagger$ is the psuedo-inverse of the graph Laplacian matrix L, I is an $N \times N$ identity matrix and η^2 is the noise power in the graph signal. For all the simulations in this section, the noise power η^2 is set at 0.01.Thus, in the pre-change regime, we have

$$X_t \sim p_0 = \mathcal{N}(\mathbf{0}, \Sigma). \tag{45}$$

For the post-change regime, we assume that the signal obtained at the rogue node follows the same distribution as the pre-change distribution. However, this signal becomes independent of signals obtained from the rest of the graph. Thus, in the post-change regime, we have

$$X_t \sim p_m = \mathcal{N}(\mathbf{0}, \Sigma_m) \tag{46}$$

for $m \in \{1, \dots, M\}$, where the covariance matrix Σ_m is given as

$$\Sigma_m[i,j] = \begin{cases} \Sigma[i,j] & \text{for } i \neq m \text{ and } j \neq m, \\ \Sigma[m,m] & \text{for } i = j = m, \\ 0 & \text{otherwise.} \end{cases} \tag{47}$$

The set of actions $\mathcal{A}$ is the set of partial observations where

$$\mathcal{A} = \bigcup_{n=2}^{N-2} \{\mathbf{M} \mid \mathbf{M} = [e_{i_1}, e_{i_2}, \dots, e_{i_n}]^T, i_1 < i_2 < \dots < i_n\}. \tag{48}$$

In our experiments, we consider the graph $\mathcal{G}$, as shown in Fig. 1. In this case, there is a total of 50 actions that can be used to observe the signal on the graph $\mathcal{G}$.
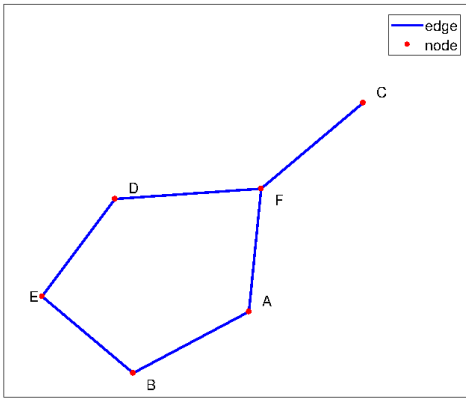

Fig. 1. Graph $\mathcal{G}$ generated using the Erdős-Rényi (ER) random graph model with $N = 6$ nodes and probability of an edge $p = 0.2$.

We consider three possible cases of observation-switching costs $\mathcal{C}$:

1) no costs are involved with $\mathcal{C}^1 = \mathbf{0}$;
2) there is only activation cost where $\mathcal{C}^2[i,j]$ is the number of nodes observed using action i;
3) there are both switching and activation costs with $\mathcal{C}^3 = \mathcal{C}^2 + \mathcal{C}'$ where $\mathcal{C}'[i,j]$ is the number of elements in the symmetric difference between the set of nodes observed by action i and set of nodes observed by action j.

Note that each $\mathcal{C}^1, \mathcal{C}^2, \mathcal{C}^3 \in \mathbb{R}^{50 \times 50}$.

When $\mathcal{C} = \mathcal{C}^1$, using results from Section V-A, we only need to design an open-loop policy of window size 0 if we are considering only open-loop policies. By considering the problem for $W = 0$, we obtain the optimal open-loop policy $\pi_{1,W=0}$ by solving Problem (29). Similarly, for the cases where $\mathcal{C} = \mathcal{C}^2, \mathcal{C}^3$, using results from Section V-B, we only need to consider the open-loop policies with $W = 0, 1$ if we are only considering open-loop policies. When $\mathcal{C} = \mathcal{C}^2$ and $W = 0$, we obtain the optimal open-loop policy $\pi_{2,W=0}$ by solving Problem (34) with a linear AOSC constraint (35). For the case $\mathcal{C} = \mathcal{C}^3$ with $W = 0$, we obtain a locally optimal open-loop policy $\pi_{3,W=0}$ by performing the IRM algorithm described in Section V-B. In the case where $\mathcal{C} = \mathcal{C}^2$ or $\mathcal{C}^3$ with $W = 1$, we obtain an approximately optimal open-loop policy $\pi_{2,W=1,\epsilon}$ or $\pi_{3,W=1,\epsilon}$ respectively by solving Problem (44) with $\epsilon \in \{10^{-4}, 10^{-5}, 10^{-6}\}$.

When no costs are involved (i.e., $\mathcal{C} = \mathbf{0}$), the optimal asymptotic ARL-WADD trade-off rate (cf. Definition 3) achieved by the stopping rule and policy $(\tau_{\pi_{1,W=0}}, \pi_{1,W=0})$ is $\bar{I} = 1.5586$. This is an upper bound for the asymptotic ARL-WADD trade-off rates for all other choices of $\mathcal{C}$.

In Fig. 2, we compare the asymptotic ARL-WADD trade-off rate of $\pi_{2,W=0}$ and $\pi_{2,W=1,\epsilon}$. First, we observe that the asymptotic ARL-WADD trade-off is below the upper bound $\bar{I}$ across the range of achievable AOSC. In the case where $\mathcal{C} = \mathcal{C}^2$, the AOSC constraint in Problem (34) reduces to

$$\text{AOSC}(\pi) = \sum_{b \in \mathcal{A}} \left(\sum_{a \in \mathcal{A}} q(a,b) \right) h(b) \tag{49}$$

where $h(b)$ is the number of nodes observed using action b. Using similar arguments from the proof of Proposition 3, we can see that there is an asymptotically optimal trade-off for a policy window of size $W = 0$ is equal to the asymptotically optimal trade-off for a policy of window size $W = 1$. Hence, the performance between $\pi_{2,W=0}$ and $\pi_{2,W=1,\epsilon}$ becomes more similar as ϵ tends to zero.

In Figs. 3 and 4, we compare the asymptotic ARL-WADD trade-off rate of $\pi_{3,W=0}$ and $\pi_{3,W=1,\epsilon}$. Similarly, we observe that the asymptotic ARL-WADD trade-off is below the upper bound $\bar{I}$ across the range of achievable AOSC. In this case, we observe that the optimal open-loop policies of window size $W = 1$ significantly outperforms the optimal open-loop policies of window size $W = 0$. When we are using a policy of window size $W = 0$, the lowest AOSC achievable while Problem (5) remains feasible is about 4. However, using a policy of window size $W = 1$, we are able to reduce AOSC to about 2 while Problem (5) remains feasible. These can be seen be comparing the lowest AOSC achieved by the respective curves in Fig. 3.
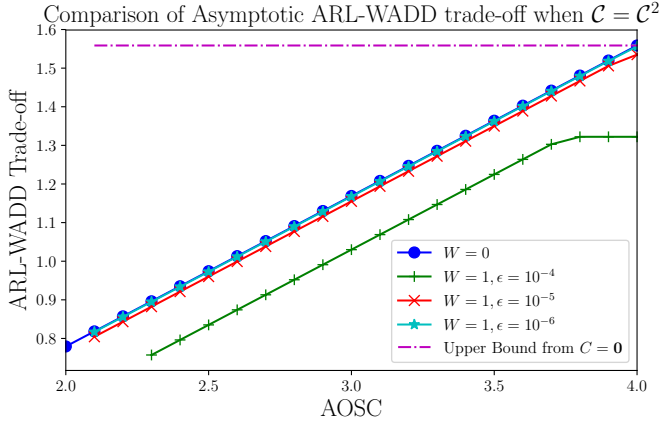
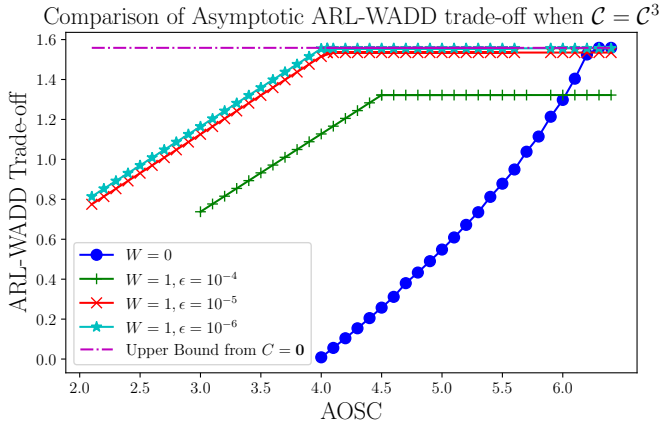Fig. 2. Comparison of the ARL-WADD trade-off rate for different stopping times when $\mathcal{C} = \mathcal{C}^2$.

Fig. 3. Comparison of the ARL-WADD trade-off rate for different stopping times when $\mathcal{C} = \mathcal{C}^3$.
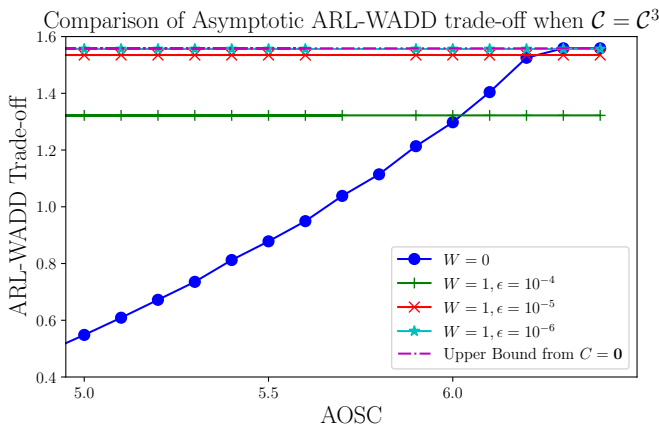

Fig. 4. Zoomed version of Fig. 3 into AOSC ≥ 5.

B. Comparison with observation-dependent policies having only sampling cost constraints

While there are no existing work on QCD with action switching-costs, for some special cases, existing work in the

literature may be relevant. In this subsection, we compare our proposed GLR CuSum with the GDECuSum proposed in [33] and discuss their strengths and weaknesses. The GDECuSum is proposed in which an on-off observation control is used to take into account the cost of observations for the purpose of QCD. The main difference between the GDECuSum and our proposed method is that our proposed method uses action switching costs while the GDECuSum only takes observation costs into account. In the next two simulations, we demonstrate the differences between the two methods by using different switching cost matrix settings. We use the same distributions and parameters for the GDECuSum as [33]. The signal is generated by a pre-change distribution with pdfs $p_0 = \mathcal{N}(0, 1)$ and 4 possible post-change distributions with pdf $p_1 = \mathcal{N}(0.4, 1)$, $p_2 = \mathcal{N}(0.6, 1)$, $p_3 = \mathcal{N}(0.8, 1)$ and $p_4 = \mathcal{N}(1, 1)$. The parameters for the GDECuSum are $\mu = 0.08$ and $h = \infty$. Using these parameters, the GDECuSum achieves $\text{ARL} - \text{WADD}$ trade-off of 0.08 and a Pre-change Duty Cycle (PDC) of 0.5 which means that only 50% of the samples are observed under the pre-change regime. At any time instance t, we let $A_t = 1$ denote the action where the t-th sample is observed and $A_t = 2$ denote the action where the t-th sample is skipped.

In the first set of simulations, we use the observation-switching cost matrix $\mathcal{C}^4 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. In this case, the AOSC is equal to the PDC as the AOSC measures the average proportion of samples observed in the pre-change regime. We solve Problem (44) with $\text{AOSC} = \text{PDC} = 0.5$ and $\epsilon = 10^{-6}$ to obtain the $\text{ARL} - \text{WADD}$ trade-off of 0.04 for our proposed GLR CuSum stopping time with an open-loop policy of window size 1. Thus, the GDECuSum which uses an observation-dependent policy out-performs our proposed GLR CuSum as expected when we use an open-loop policy of window size 1.

In the second set of simulations, we use another switching cost matrix to illustrate a key difference between our proposed method and the GDECuSum. For gas sensors, the recovery time [50], [51] is the time taken for the sensor to reset after taking a measurement. The recovery time for some sensors can be rather long and in some cases, the recovery time may exceed one duty cycle of the sensor. Thus, the sensor is unable to make a reliable observation if it has already made an observation in the previous duty cycle. We can model this phenomenon as having a high action-switching cost when the sensor is switched on consecutively. For the next set of simulations, we use the observation-switching cost matrix $\mathcal{C}^5 = \begin{pmatrix} 10^6 & 0 \\ 1 & 0 \end{pmatrix}$. To estimate the AOSC of the GDECuSum stopping time, we generate 100 test sequences of size 10^5 under the pre-change regime. Using these test sequences, the empirical AOSC of the GDECuSum stopping time is 1349. This means that the GDECuSum frequently makes consecutive observations which is undesirable. We solve Problem (44) with $\text{AOSC} = 0.5$ and $\epsilon = 10^{-10}$ to obtain the $\text{ARL} - \text{WADD}$ trade-off of 0.04 for our proposed GLR CuSum stopping time with an open-loop policy of window size 1. Here, it can be seen that our proposed GLR CuSum yields a feasible solution with low AOSC while the GDECuSum is unable to keep the

AOSC low as the GDECuSum does not take action-switching costs into account.

VII. Conclusion and Future Work

In this paper, we discussed the problem of QCD while taking sampling and switching costs into consideration. We formulated the QCD problem with an additional AOSC constraint. Asymptotically optimal stopping times were proposed and the design of optimal open-loop policies were formulated as quadratic optimization problems. We showed that open-loop policies of window size $W > 1$ can be reduced to an open-loop policy of window size $W = 1$ while maintaining the same asymptotic ARL-WADD trade-off and AOSC. Thus, it is sufficient to solve the policy design problem for $W = 0, 1$. We applied the IRM algorithm to the policy design problem to obtain locally optimal solutions. For cases with additional assumptions on the observation-switching cost matrix $\mathcal{C}$, globally optimal solutions can be obtained. The methods developed are for the case when the window W is finite. The results regarding the structure of asymptotically optimal stopping times such as Proposition 5, do not hold in general when $W = \infty$ and would be an interesting direction for future work.

Appendix A
Proof of Lemma 2

Since $\mathrm{supp}(q)$ lies in the recurrence class $\mathcal{R}$, for any $t \geq \nu$ and $\beta \in \mathcal{A}^W$ such that $q(\beta) > 0$, we have $\beta \in \mathcal{R}$ and

$$\frac{1}{t}\Lambda_m(\nu, \nu + t - 1)$$
$$= \sum_{\alpha \in \mathcal{A}^W} \frac{1}{t} \sum_{\left\{i \text{ s.t. } \substack{\nu \leq i \leq \nu+t-1, \\ A^{i-W+1:i}=\alpha}\right\}} \log \frac{p_m(Y_i|\alpha[W])}{p_0(Y_i|\alpha[W]))}$$
$$= \sum_{\alpha \in \mathcal{R}} \frac{N_t(\alpha;\beta)}{t} \frac{1}{N_t(\alpha;\beta)} \sum_{\left\{i \text{ s.t. } \substack{\nu \leq i \leq \nu+t-1, \\ A^{i-W+1:i}=\alpha}\right\}} \log \frac{p_m(Y_t|\alpha[W])}{p_0(Y_t|\alpha[W]))},$$

where the last equality follows because for any $\alpha \notin \mathcal{R}$, $\left\{i \text{ s.t. } \substack{\nu \leq i \leq \nu+t-1, \\ A^{i-W+1:i}=\alpha}\right\} = \emptyset$. We have

$$\frac{N_t(\alpha;\beta)}{t} \xrightarrow[\text{a.s.}]{\mathbb{P}_{\nu,m}} \bar{q}(\alpha) \quad \text{as } t \to \infty, \tag{50}$$

and

$$\frac{1}{N_t(\alpha;\beta)} \sum_{\left\{i \text{ s.t. } \substack{\nu \leq i \leq \nu+t-1, \\ A^{i-W+1:i}=\alpha}\right\}} \log \frac{p_m(Y_t|\alpha[W])}{p_0(Y_t|\alpha[W]))}$$
$$\xrightarrow[\text{a.s.}]{\mathbb{P}_{\nu,m}} D(p_m(\cdot \mid \alpha[W]) \| p_0(\cdot \mid \alpha[W])) \quad \text{as } t \to \infty. \tag{51}$$

Thus from (50), we obtain

$$\frac{1}{t}\Lambda_m(\nu, \nu + t - 1)$$
$$\xrightarrow[\text{a.s.}]{\mathbb{P}_{\nu,m}} \sum_{\alpha \in \mathcal{R}} \bar{q}(\alpha) D(p_m(\cdot \mid \alpha[W]) \| p_0(\cdot \mid \alpha[W])) = I_{m,\pi}. \tag{52}$$

The proof is now complete.

Appendix B
Proof of Lemma 3

Suppose $\mathrm{supp}(\bar{q}) \subseteq \mathcal{R}$ for a recurrence class $\mathcal{R}$, then $\mathrm{supp}(q) \subseteq \mathcal{R} \cup \mathcal{U}$, where $\mathcal{U}$ is a set of transient states such that the first-passage probability of entering $\mathcal{R}$ from each $\beta \in \mathcal{U}$ is one. Then, from Lemma 2, (21) follows. Let $0 \leq j < t$ and $E_j(\nu)$ be the event $\{t^{-1} \max_{0 \leq j < t} \Lambda_m(\nu, \nu + j) > (1 + \epsilon)I_{m,\pi}\}$ and we have

$$\sup_{0 \leq \nu < \infty} \mathrm{ess\,sup}\, \mathbb{P}_{\nu,m}\left(E_j(\nu) \mid A^{1:\nu-1}, Y^{1:\nu-1}\right)$$
$$= \sup_{0 \leq \nu < \infty} \mathrm{ess\,sup}\, \mathbb{P}_{\nu,m}\left(E_j(\nu) \mid A^{\nu-W:\nu-1}\right) \tag{53}$$
$$= \sup_{0 \leq \nu < \infty} \max_{\substack{\alpha \in \mathcal{A}^W \\ q(\alpha)>0}} \mathbb{P}_{\nu,m}\left(E_j(\nu) \mid A^{\nu-W:\nu-1} = \alpha\right)$$
$$= \max_{\substack{\alpha \in \mathcal{A}^W \\ q(\alpha)>0}} \mathbb{P}_{W+1,m}\left(E_j(W+1) \mid A^{1:W} = \alpha\right)$$
$$\to 0 \quad \text{as } t \to \infty, \tag{54}$$

where (53) is because $\Lambda_m(\nu, \nu + j)$ is independent of $A^{1:\nu-W-1}$ and $Y^{1:\nu-1}$ given $A^{\nu-W:\nu-1}$ for each $0 \leq j < t$, and (54) is because each of the terms within the set $\{\alpha \in \mathcal{A}^W \mid q(\alpha) > 0\}$ that we take maximum over converges to zero due to Lemma 2.

Appendix C
Proof of Proposition 2

Without loss of generality, we consider the case where $\bar{q}$ has support in the recurrence classes $\mathcal{R}_1$ and $\mathcal{R}_2$. Let q_1, q_2 be stationary distributions on $\mathcal{A}^W$ such that q_i has support only in $\mathcal{R}_i$ for $i \in \{1, 2\}$ and that there exists $0 < \lambda < 1$ such that $\bar{q} = \lambda q_1 + (1 - \lambda)q_2$. Let $\pi_1 = (q_1, T)$ and $\pi_2 = (q_2, T)$. From Lemma 1, we obtain

$$\mathrm{AOSC}(q, T) = \mathrm{AOSC}(\bar{q}, T)$$
$$= \lambda \mathrm{AOSC}(q_1, T) + (1 - \lambda)\mathrm{AOSC}(q_2, T)$$
$$\geq \min\{\mathrm{AOSC}(\pi_1), \mathrm{AOSC}(\pi_2)\}. \tag{55}$$

Let E_1 be the event that $\mathcal{R}_1$ is visited before $\mathcal{R}_2$ and E_2 be the event that $\mathcal{R}_2$ is visited before $\mathcal{R}_1$. We have

$$\mathrm{WADD}(\tau, \pi)$$
$$= \sup_{\substack{\nu \geq 1 \\ 1 \leq m \leq M}} \mathrm{ess\,sup}\, \mathbb{E}_{\nu,m}\left[(\tau - \nu + 1)^+ \mid A^{1:\nu-1}, Y^{1:\nu-1}\right]$$
$$= \sup_{\substack{\nu \geq 1 \\ 1 \leq m \leq M}} \max_{i=1,2} \mathrm{ess\,sup}\, \mathbb{E}_{\nu,m}\left[(\tau - \nu + 1)^+ \mid A^{1:\nu-1}, Y^{1:\nu-1}, E_i\right]$$
$$= \max_{i=1,2} \sup_{\substack{\nu \geq 1 \\ 1 \leq m \leq M}} \mathrm{ess\,sup}\, \mathbb{E}_{\nu,m}\left[(\tau - \nu + 1)^+ \mid A^{1:\nu-1}, Y^{1:\nu-1}, E_i\right]$$
$$= \max_{i=1,2} \sup_{\substack{\nu \geq 1 \\ 1 \leq m \leq M}} \mathrm{ess\,sup}\, \mathbb{E}_{\nu,m}\left[(\tau - \nu + 1)^+ \mid A^{1:\nu-1}, Y^{1:\nu-1}, E_i\right]$$
$$= \max_{i=1,2} \mathrm{WADD}(\tau, \pi_i).$$

Setting $\pi' = \pi_j$ where $j = \arg\min_{i=1,2} \mathrm{AOSC}(\pi_i)$ completes the proof.

APPENDIX D
PROOF OF THEOREM 1

We first note that Problem (5) is feasible and an optimal solution exists. Let the open-loop policy $\pi^* = (q, T)$ and stopping-time τ^* such that (τ^*, π^*) is optimal for Problem (5). From Proposition 2, we can find an open-loop policy $\pi = (q', T)$ such that $\overline{q'}$ has support in only one recurrence class and

$$\text{AOSC}(\pi) \leq \text{AOSC}(\pi^*),$$
$$\text{WADD}(\tau^*, \pi) \leq \text{WADD}(\tau^*, \pi^*).$$

From the discussion before (24), the GLR CuSum stopping time τ_π is asymptotically optimal for the following problem:

$$\min_\tau \quad \text{WADD}(\tau, \pi) \quad \text{s.t.} \quad \text{ARL}(\tau, \pi) \geq \gamma. \quad (56)$$

Thus, the GLR CuSum stopping time τ_π satisfies

$$\text{WADD}(\tau_\pi, \pi) \asymp \text{WADD}(\tau^*, \pi) \leq \text{WADD}(\tau^*, \pi^*),$$

as $\gamma \to \infty$. Hence, (τ_π, π) is also an asymptotically optimal solution to Problem (5).

APPENDIX E
PROOF OF PROPOSITION 3

If π is equivalent to an open-loop policy of window size 0, then $\pi_0 = \pi$ proves the proposition. If $\pi = (q, T)$ is an open-loop policy of window size $W > 0$, by Theorem 1, it suffices to consider the case where the support of $\overline{q}$ is a subset of a recurrence class of T. Let $\pi_0 = (q_0, T_0)$ such that $q_0 = \overline{q}^0$ and T_0 be the probability transition matrix with rows equal to q_0 where

$$\overline{q}^0(a) = \sum_{\alpha \in \mathcal{A}^W : \alpha[W] = a} \overline{q}(\alpha)$$

for any $a \in \mathcal{A}$. For any $m \in \{1, \ldots, M\}$, we have

$$
\begin{aligned}
I_{m,\pi} &= \sum_{\alpha \in \mathcal{A}^W} \overline{q}(\alpha) D(p_m(\cdot \mid \alpha[W]) \,\|\, p_0(\cdot \mid \alpha[W])) \\
&= \sum_{a \in \mathcal{A}} \sum_{\left\{ \substack{\alpha \in \mathcal{A}^W : \\ \alpha[W] = a} \right\}} \overline{q}(\alpha) D(p_m(\cdot \mid \alpha[W]) \,\|\, p_0(\cdot \mid \alpha[W])) \\
&= \sum_{a \in \mathcal{A}} D(p_m(\cdot \mid a) \,\|\, p_0(\cdot \mid a)) \sum_{\left\{ \substack{\alpha \in \mathcal{A}^W : \\ \alpha[W] = a} \right\}} \overline{q}(\alpha) \\
&= \sum_{a \in \mathcal{A}} q_0(a) D(p_m(\cdot \mid a) \,\|\, p_0(\cdot \mid a)) = I_{m,\pi_0}, \quad (57)
\end{aligned}
$$

where (57) is due to the fact that $\overline{q_0} = q_0$ as π_0 is an open-loop policy of window size 0. By Proposition 1, we obtain $\text{WADD}(\tau_\pi, \pi) \asymp \text{WADD}(\tau_{\pi_0}, \pi_0)$ as $\gamma \to \infty$. The proof is now complete.

APPENDIX F
PROOF OF PROPOSITION 4

We first proof claim (i). Let (q^*, z^*) be an optimal solution to Problem (29). It can be easily checked that q^* is a feasible solution for Problem (28). We prove the statement by contradiction. Suppose q^* is not optimal for Problem (28). Then there exists a feasible q' such that

$$
\max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1}
$$
$$
< \max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q^*(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1}.
$$

Let

$$
z' = -1 / \max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1}.
$$

This gives

$$
\min_{1 \leq m \leq M} \sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) + z' = 0.
$$

Thus, z' satisfies $\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) + z' \geq 0$ for all $m \in \{1, ..., M\}$ and (q', z') is a feasible solution with $z' < z^*$. This contradicts the optimality of (q^*, z^*). Hence, q^* is an optimal solution to Problem (28).

We now proof claim (ii). Let q^* be an optimal solution to Problem (28). Let

$$
z^* = -1 / \max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1}.
$$

Following the manipulation above, z^* satisfies $\sum_{a \in \mathcal{A}} z^*(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) + z^* \geq 0$ for all $m \in \{1, ..., M\}$. Thus, (q^*, z^*) is a feasible solution. We prove the statement by contradiction. Suppose (q^*, z^*) is not optimal for Problem (29). Then there exists a feasible (q', z') such that $z' < z^*$. Since (q', z') is feasible, $\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) + z' \geq 0$ holds for all $m \in \{1, \ldots, M\}$. Thus, with some algebraic manipulation,

$$
\max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1} \leq \frac{-1}{z'}.
$$

Since $z' < z^*$, we have $\frac{-1}{z'} < \frac{-1}{z^*}$. By the choice of z^*, we have

$$
\max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q'(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1}
$$
$$
\leq \frac{-1}{z'}
$$
$$
< \frac{-1}{z^*} = \max_{1 \leq m \leq M} \left(\sum_{a \in \mathcal{A}} q^*(a) D(p_0(\cdot \mid a) \,\|\, p_m(\cdot \mid a)) \right)^{-1}.
$$

This contradicts the optimality of q^*. Hence, (q^*, z^*) is an optimal solution to Problem (29). The proof is now complete.

APPENDIX G
PROOF OF PROPOSITION 5

If π is equivalent to an open-loop policy of window size 1, we let $\pi_1 = \pi$ and we have proved the proposition. If π is an open-loop policy of window size $W > 1$, by Theorem 1, it suffices to consider the case where the support of $\overline{q}$ is a subset of a recurrence class $\mathcal{R}$. The AOSC of π can be expressed as

$$
\begin{aligned}
& \text{AOSC}(\pi) \\
& = \text{AOSC}(\overline{q}, T) \\
& = \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}_\infty \left[\sum_{t=2}^{n+1} \mathcal{C}[A_{i-1}, A_i] \right] \\
& = \limsup_{n \to \infty} \frac{1}{n} \sum_{\beta \in \mathcal{A}^W} \overline{q}(\beta) \mathbb{E}_\infty \left[\sum_{t=2}^{n+1} \mathcal{C}[A_{i-1}, A_i] \,\middle|\, A^{1:W} = \beta \right] \\
& = \limsup_{n \to \infty} \frac{1}{n} \sum_{\beta \in \mathcal{A}^W} \overline{q}(\beta) \mathbb{E}_\infty \left[\sum_{\alpha \in \mathcal{A}^W} \mathcal{C}_\alpha N_{n+1}(\alpha; \beta) \,\middle|\, A^{1:W} = \beta \right] \\
& = \sum_{\beta \in \mathcal{A}^W} \sum_{\alpha \in \mathcal{A}^W} \overline{q}(\beta) \mathcal{C}_\alpha \overline{q}(\alpha) = \sum_{\alpha \in \mathcal{A}^W} \mathcal{C}_\alpha \overline{q}(\alpha).
\end{aligned}
$$

Let $\pi_2 = (q_2, T_2)$ be an open-loop policy of window size 2 and it can be similarly shown that $\text{AOSC}(\pi_2) = \sum_{\alpha \in \mathcal{A}^2} \mathcal{C}_\alpha \overline{q_2}(\alpha)$. Thus, it can be seen that the $\text{AOSC}(\pi_2)$ depends only on $\overline{q_2}$. Let $q_2 = \overline{q}^1$ be the projection of $\overline{q}$ onto $\mathcal{A}^2$ where

$$
\overline{q}^1(a, b) = \sum_{\substack{\alpha \in \mathcal{A}^W: \\ \alpha[W-1]=a, \alpha[W]=b}} \overline{q}(\alpha),
$$

for any $a, b \in \mathcal{A}$ and T_2 be defined as

$$
T_2[(a,b),(c,d)] = \begin{cases} \frac{\overline{q}^1(c,d)}{\sum_{d \in \mathcal{A}} \overline{q}^1(c,d)} & \text{if } c = b, \\ 0 & \text{otherwise.} \end{cases}
$$

For any $(a, b) \in \mathcal{A}^2$, we have

$$
\begin{aligned}
\sum_{(c,d) \in \mathcal{A}^2} T_2[(a,b),(c,d)] & = \sum_{d \in \mathcal{A}} \frac{\overline{q}^1(b,d)}{\sum_{d_1 \in \mathcal{A}} \overline{q}^1(b,d_1)} \\
& = \frac{\sum_{d \in \mathcal{A}} \overline{q}^1(b,d)}{\sum_{d_1 \in \mathcal{A}} \overline{q}^1(b,d_1)} = 1,
\end{aligned}
$$

and hence, T_2 is a probability transition matrix. Since the support of $\overline{q}$ is a subset of a single recurrence class of T, the support of q_2 also lies in a single recurrence class of T_2. Next, we show that q_2 is a stationary distribution of T_2:

$$
\begin{aligned}
T_2 q_2[(c,d)] & = \sum_{(a,b) \in \mathcal{A}^2} T_2[(a,b),(c,d)] \overline{q}^1(a,b) \\
& = \sum_{a \in \mathcal{A}} T_2[(a,c),(c,d)] \overline{q}^1(a,c) \\
& = \sum_{a \in \mathcal{A}} \frac{\overline{q}^1(c,d)}{\sum_{d \in \mathcal{A}} \overline{q}^1(c,d)} \overline{q}^1(a,c) \\
& = \frac{\overline{q}^1(c,d)}{\sum_{d \in \mathcal{A}} \overline{q}^1(c,d)} \sum_{a \in \mathcal{A}} \overline{q}^1(a,c) \\
& = \overline{q}^1(c,d) = q_2(c,d).
\end{aligned}
$$

Therefore, q_2 is a stationary distribution of T_2 which has support contained in one recurrence class. Hence, $\overline{q_2} = q_2$ and $\text{AOSC}(\pi) = \text{AOSC}(\pi_2)$. Using similar arguments from the proof of Proposition 3, we have $I_{m,\pi} = I_{m,\pi_1}$ for $m = \{1, \ldots, M\}$ and $\text{WADD}(\tau_\pi, \pi) \asymp \text{WADD}(\tau_{\pi_2}, \pi_2)$ as $\gamma \to \infty$.

Furthermore, a quick computation shows that $\pi_2 = (q_2, T_2)$ is equivalent to the open-loop policy $\pi_1 = (q_1, T_1)$ of window size 1 with $q_1(a) = \sum_{b \in \mathcal{A}} q_2(a,b)$ and $T_1[a,b] = \frac{q_2(a,b)}{\sum_{b \in \mathcal{A}} q_2(a,b)}$.

Thus, we have $\text{AOSC}(\pi) = \text{AOSC}(\pi_1)$ and $\text{WADD}(\tau_\pi, \pi) \asymp \text{WADD}(\tau_{\pi_1}, \pi_1)$ as $\gamma \to \infty$. The proof is complete.

APPENDIX H
COUNTEREXAMPLE TO PROPOSITION 5 FOR $W = \infty$

We consider the problem with two actions $\mathcal{A} = \{1, 2\}$ and a 2×2 observation-switching costs matrix $\mathcal{C} = \mathcal{C}^*$ as follows:

$$
\mathcal{C}^*[i,j] = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \neq j, \end{cases}
$$

Let $M = 2$. The pre- and post-change distributions are chosen with the following pmfs:

$$
p_0 = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right], \quad p_1 = \left[\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\right], \quad p_2 = \left[\frac{1}{2}, \frac{1}{6}, \frac{1}{3}\right].
$$

We define the actions so that given the actions, pre- and post-change distributions are discrete distributions with the following pmfs:

$$
\begin{aligned}
& \mathbb{P}_{\nu,m}\left(Y_t = 1 \,|\, A = 1\right) = p_m[1] + p_m[2], \\
& \mathbb{P}_{\nu,m}\left(Y_t = 2 \,|\, A = 1\right) = p_m[3], \\
& \mathbb{P}_{\nu,m}\left(Y_t = 1 \,|\, A = 2\right) = p_m[1], \\
& \mathbb{P}_{\nu,m}\left(Y_t = 2 \,|\, A = 2\right) = p_m[2] + p_m[3],
\end{aligned}
$$

for $t \geq \nu$ and $m \in \{1, 2\}$,

$$
\begin{aligned}
& \mathbb{P}_{\nu,m}\left(Y_t = 1 \,|\, A = 1\right) = p_0[1] + p_0[2], \\
& \mathbb{P}_{\nu,m}\left(Y_t = 2 \,|\, A = 1\right) = p_0[3], \\
& \mathbb{P}_{\nu,m}\left(Y_t = 1 \,|\, A = 2\right) = p_0[1], \\
& \mathbb{P}_{\nu,m}\left(Y_t = 2 \,|\, A = 2\right) = p_0[2] + p_0[3],
\end{aligned}
$$

for $t < \nu$, and

$$
\begin{aligned}
& \mathbb{P}_\infty\left(Y_t = 1 \,|\, A = 1\right) = p_0[1] + p_0[2], \\
& \mathbb{P}_\infty\left(Y_t = 2 \,|\, A = 1\right) = p_0[3], \\
& \mathbb{P}_\infty\left(Y_t = 1 \,|\, A = 2\right) = p_0[1], \\
& \mathbb{P}_\infty\left(Y_t = 2 \,|\, A = 2\right) = p_0[2] + p_0[3],
\end{aligned}
$$

for all $t \in \mathbb{N}$. To compute the upper-bound of the ARL-WADD trade-off, we consider the QCD with observation-switching costs problem with $\mathcal{C} = \mathcal{C}^0 = \mathbf{0}$ where $\mathbf{0}$ is the 2×2 zero matrix. The ARL-WADD trade-off when $\mathcal{C} = \mathcal{C}^0$ is $\frac{1}{4}\left(\log \frac{3}{4} + \log \frac{3}{2}\right)$.

We consider a deterministic policy π_∞(i.e., Markov chain with $W = \infty$) and sample the signal using the following sequence of actions:

$$
\{A_t\} = \{1, 2, 1, 1, 2, 2, 1, 1, 1, 2, 2, 2, 1, 1, 1, 1, 2, 2, 2, 2, \ldots\}
$$

where the number of times each action is used to sample the signal increases by 1 before switching. When $\mathcal{C} = \mathcal{C}^*$, the average number of switches tends to zero as $t \to \infty$ and we have $\mathrm{ASOC}(\pi_\infty) = 0$ and

$$\mathbb{P}_{\nu,m}\left(\lim_{t\to\infty} \frac{1}{t}\Lambda_m(\nu, \nu+t) = \frac{1}{4}\left(\log\frac{3}{4} + \log\frac{3}{2}\right)\right) = 1,$$

for any $\nu \in \mathbb{N}$ and $m \in \{1, 2\}$. Thus, the policy π_∞ achieves asymptotically optimal WADD-ARL trade-off with $\mathrm{ASOC}(\pi_\infty) = 0$.

However, $\mathrm{ASOC}(\pi_1) > 0$ for any policy π_1 of window size $W = 1$ where both actions 1 and 2 are used. On the other hand, if only one of the actions is used, then either I_{1,π_1} or I_{2,π_1} is zero. Hence, π_∞ cannot be reduced to a policy of window size $W = 1$.

REFERENCES

[1] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.

[2] H. V. Poor and O. Hadjiliadis, *Quickest detection*. Cambridge University Press Cambridge, 2009, vol. 40.

[3] V. V. Veeravalli and T. Banerjee, "Quickest change detection," *Academic Press Library in Signal Processing*, vol. 3, pp. 209–255, 2014.

[4] W. H. Woodall, D. J. Spitzner, D. C. Montgomery, and S. Gupta, "Using control charts to monitor process and product quality profiles," *J. of Quality Technology*, vol. 36, no. 3, p. 309, 2004.

[5] T. L. Lai, "Sequential changepoint detection in quality control and dynamical systems," *J. of the Roy. Statistical Soc.*, pp. 613–658, 1995.

[6] T. S. Lau, W. P. Tay, and V. V. Veeravalli, "Quickest change detection with unknown post-change distribution," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.* IEEE, 2017, pp. 3924–3928.

[7] T. S. Lau and W. P. Tay, "Quickest change detection in the presence of a nuisance change," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5281 – 5296, Oct. 2019.

[8] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," *Statistical Sci.*, pp. 235–249, 2002.

[9] L. Lai, Y. Fan, and H. V. Poor, "Quickest detection in cognitive radio: A sequential change detection framework," in *IEEE Conf. Global Telecommun.* IEEE, 2008, pp. 1–5.

[10] Y. Zhang, W. P. Tay, K. H. Li, M. Esseghir, and D. Gaiti, "Learning temporal-spatial spectrum reuse," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 3092 – 3103, Jul. 2016.

[11] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blazek, and H. Kim, "A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3372–3382, 2006.

[12] A. G. Tartakovsky, "Rapid detection of attacks in computer networks by quickest change-point detection methods," *Data anal. for network cyber-security*, pp. 33–70, 2014.

[13] H. Sohn, J. A. Czarnecki, and C. R. Farrar, "Structural health monitoring using statistical process control," *J. Structural Eng.*, vol. 126, no. 11, pp. 1356–1363, 2000.

[14] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection via temporal pattern discovery," in *Proc. Conf. Knowl. Discovery and Data Mining*. ACM, 2012, pp. 823–831.

[15] F. Ji, W. P. Tay, and L. Varshney, "An algorithmic framework for estimating rumor sources with different start times," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2517 – 2530, May 2017.

[16] W. Tang, F. Ji, and W. P. Tay, "Estimating infection sources in networks using partial timestamps," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 2, pp. 3035 – 3049, Dec. 2018.

[17] V. M. R. Muggeo and G. Adelfio, "Efficient change point detection for genomic sequences of continuous measurements," *Bioinformatics*, vol. 27, no. 2, pp. 161–166, 2011.

[18] T. Banerjee, Y. C. Chen, A. D. Dominguez-Garcia, and V. V. Veeravalli, "Power system line outage detection and identification-A quickest change detection approach," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.* IEEE, 2014, pp. 3450–3454.

[19] P. R. Coppin and M. E. Bauer, "Digital change detection in forest ecosystems with remote sensing imagery," *Remote sensing reviews*, vol. 13, no. 3-4, pp. 207–234, 1996.

[20] Y.-W. Hong and A. Scaglione, "Distributed change detection in large scale sensor networks through the synchronization of pulse-coupled oscillators," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, May 2004, pp. iii–869.

[21] J. Yang, X. Zhong, and W. P. Tay, "A dynamic Bayesian nonparametric model for blind calibration of sensor networks," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3942 – 3953, Oct. 2018.

[22] T. S. Lau, W. P. Tay, and V. V. Veeravalli, "A binning approach to quickest change detection with unknown post-change distribution," *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 609–621, Feb 2019.

[23] X. Liu, W. P. Tay, Z.-W. Liu, and G. Xiao, "Quasi-synchronization of heterogeneous networks with a generalized Markovian topology and event-triggered communication," *IEEE Trans. Cybern.*, 2020, in press.

[24] F. Ji and W. P. Tay, "A Hilbert space theory of generalized graph signal processing," *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6188 – 6203, Dec. 2019.

[25] V. V. Veeravalli, T. BaSar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 433–442, 1993.

[26] V. V. Veeravalli, "Decentralized quickest change detection," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1657–1665, 2001.

[27] O. Hadjiliadis, H. Zhang, and H. V. Poor, "One shot schemes for decentralized quickest change detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3346–3359, 2009.

[28] B. Jiang and L. Lai, "Distributed quickest detection with optional observations at the fusion center," in *IEEE Conf. Signals, Syst. and Comput.,*. IEEE, 2016, pp. 112–116.

[29] T. Banerjee and V. V. Veeravalli, "Data-efficient quickest change detection with on–off observation control," *Sequential Analysis*, vol. 31, no. 1, pp. 40–77, 2012.

[30] ——, "Data-efficient quickest change detection in minimax settings," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6917–6931, 2013.

[31] T. Banerjee, V. V. Veeravalli, and A. Tartakovsky, "Decentralized data-efficient quickest change detection," in *Proc. IEEE Int. Symp. Inform. Theory*. IEEE, 2013, pp. 2587–2591.

[32] J. Geng and L. Lai, "Non-bayesian quickest change detection with stochastic sample right constraints," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 5090–5102, 2013.

[33] T. Banerjee and V. V. Veeravalli, "Data-efficient minimax quickest change detection with composite post-change distribution," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 5172–5184, 2015.

[34] Y. Xie, M. Wang, and A. Thompson, "Sketching for sequential change-point detection," in *Proc. IEEE Global Conf. Signal and Inform. Process.* IEEE, 2015, pp. 78–82.

[35] G. K. Atia, "Change detection with compressive measurements," *IEEE Signal Process. Lett.*, vol. 22, no. 2, pp. 182–186, 2015.

[36] J. Heydari and A. Tajer, "Quickest change detection in structured data with incomplete information," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.*, March 2017, pp. 6434–6438.

[37] T. S. Lau and W. P. Tay, "Optimal sampling policy for quickest change detection," in *Proc. IEEE Global Conf. Signal and Inform. Process.* IEEE, 2017, pp. 78–82.

[38] K. Premkumar and A. Kumar, "Optimal sleep-wake scheduling for quickest intrusion detection using wireless sensor networks," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 2008, pp. 1400–1408.

[39] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, 2011.

[40] Y. Mei, "Quickest detection in censoring sensor networks," in *2011 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2011, pp. 2148–2152.

[41] X. Ren, K. H. Johansson, and L. Shi, "Quickest change detection with observation scheduling," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2635–2647, 2016.

[42] T. S. Lau and W. P. Tay, "Asymptotically optimal sampling policy for quickest change detection with observation-switching cost," ArXiV, Tech. Rep. arXiv:1912.11693 [math.OC], January 2021. [Online]. Available: http://arxiv.org/abs/1912.11693

[43] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Stat.*, pp. 1897–1908, 1971.

[44] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to probability*. Athena Scientific Belmont, MA, 2002, vol. 1.

[45] D. H. Greene and D. E. Knuth, *Mathematics for the Analysis of Algorithms*. Springer Science & Business Media, 2007.

[46] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[47] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.

[48] C. Sun and R. Dai, "Rank-constrained optimization and its applications," *Automatica*, vol. 82, pp. 128–136, 2017.

[49] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning laplacian matrix in smooth graph signal representations," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6160–6173, 2016.

[50] A. Lahlalia, O. Le Neel, R. Shankar, S. Selberherr, and L. Filipovic, "Improved sensing capability of integrated semiconducting metal oxide gas sensor devices," *Sensors*, vol. 19, no. 2, p. 374, 2019.

[51] M. G. Masteghin, D. R. Godoi, and M. O. Orlandi, "Heating method effect on sno micro-disks as no 2 gas sensor," *Nano-Hetero-Structures for Chemical Sensing: Opportunities and Challenges*, 2020.

[52] H. Chernoff, "Sequential design of experiments," *The Ann. of Math. Stat.*, vol. 30, no. 3, pp. 755–770, 1959.

[53] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2451–2464, 2013.