2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Quickest Change Point Detection with Measurements over a Lossy Link

Krishna Chaythanya KV, Arpan Chattopadhyay, Anurag Kumar, and Rajesh Sundaresan

2023 62nd IEEE Conference on Decision and Control (CDC) | 979-8-3503-0124-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/CDC49753.2023.10384241

Abstract— **Motivated by Industry 4.0 applications, we consider quickest change point detection (QCD) when process measurements are transmitted by a sensor over a lossy wireless link to a decision maker (DM). The sensor node samples measurements using a Bernoulli sampling process, and places the measurement samples in a transmit queue of the transmitter. The transmitter uses a retransmit-until-success transmission strategy to deliver packets to the DM over the lossy link, which is modeled as an independent Bernoulli process and has different loss probabilities before and after the change. We pose the QCD problem in the non-Bayesian setting under Lorden's framework [1], and derive a CUSUM algorithm. By defining a suitable Markov process, involving the DM measurements and the queue length process, we show that the problem reduces to QCD of a Markov process. Characterizing the information measure I per measurement sample at the DM, our analysis proves the asymptotic optimality of our algorithm when the false alarm rate tends to zero. We discuss extensions of the analysis to periodic sampling and no-retransmission cases. Through numerical analysis, we demonstrate trade-offs that can be used to optimize system design parameters such as the sampling rate of the measurement process in the non-asymptotic regime.**

I. INTRODUCTION

Online condition monitoring of industrial machinery is an important part of the vision of Industry 4.0 [2]. Such monitoring can help detect incipient failure of components, such as bearings, followed by maintenance actions or replacement, thus preventing catastrophic system failures that can lead to expensive repair and downtime. Due to the practical difficulty of laying several wires from each machine to a central data-processing computer, specially when moving parts have to be monitored, Industrial IoT networks are expected to be largely wireless. Indeed, some machine component manufacturers (e.g., bearings, see Schaeffler OPTIME [3]) have already packaged sensors with wireless communication chipsets. In such a situation, the measurements from the sensors are subject to possible data loss on the wireless link. Further, the increased vibrations in a faulty machine can lead to degradation in the quality of the wireless link.

With the above considerations in mind, in this paper we are motivated to study the classical problem of quickest change point detection (QCD) in a stochastic process (e.g., the vibration process in a bearing) with the novel feature that the measurements can experience random loss, and, therefore, delay due to retransmissions. When the change in the measured process coincides with a degradation of the wireless channel, the packet loss process also provides information about the change, in addition to the contents of the delivered measurement packets.

We consider the following model. The wireless channel is slotted (see, e.g., 6TiSCH [4]), with each transmitted packet occupying one slot. Packet delivery success is known to the transmitter by an acknowledgment within the slot. An unacknowledged packet is reattempted until success. The sensor collects at most one sample per channel slot. A packet carrying a measurement sample is placed in the link transmission queue. The problem of the decision maker (DM) at the receiver of the wireless link is to use the measurements, and any other information, to quickly detect a change in the measurement process at the sensor, while controlling the false alarm rate.

We formulate this QCD problem in the framework of Lorden [1], where we attempt to minimize the worst case (in the sense of an essential supremum) average delay to detection (ESADD) under a constraint on the false alarm rate measured by the average run length to false alarm (ARL2FA). In the situation that the wireless packet loss probability changes after the change occurs, and assuming the knowledge of the packet loss probability before and after the change point, we provide a sequential detection algorithm based on Page's CUSUM algorithm [5]. By augmenting the observation space with the queue length process, we show that the QCD problem reduces to a sequential detection problem for Markov processes, and prove analytically, using Lai's [6] framework, that the proposed algorithm is asymptotically the optimal sequential detector as the ARL2FA $\to \infty$. We provide a numerical analysis of the non-asymptotic performance of the algorithm, and show the effect of the delays due to the transmission queue and the sampling rate on the average delay to detection (ADD). We also show, through simulations, that there exists an optimum sampling rate that minimizes the ADD in the non-asymptotic regime.

a) Related literature: The QCD problem has a long history. In [1], Lorden proved the asymptotic optimality of the CUSUM algorithm (originally proposed by Page in

Krishna Chaythanya KV, Anurag Kumar, and Rajesh Sundaresan are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru, India {krishnackv,anurag,rajeshs}@iisc.ac.in

Arpan Chattopadhyay is with the Electrical Engineering Department and the Bharti School of Telecommunication Technology and Management, Indian Institute of Technology Delhi, India arpanc@ee.iitd.ac.in

Krishna Chaythanya KV was supported by a fellowship grant from the Centre for Networked Intelligence (a Cisco CSR initiative) of the Indian Institute of Science, Bangalore.

Arpan Chattopadhyay acknowledges support via the professional development fund and professional development allowance from IIT Delhi, grant no. GP/2021/ISSC/022 from I-Hub Foundation for Cobotics and grant no. CRG/2022/003707 from Science and Engineering Research Board (SERB), India.

Anurag Kumar acknowledges support from the Centre for Networked Intelligence (a CISCO CSR initiative) of the Indian Institute of Science, Bangalore.

[5]) in minimizing the worst-case delay to detection as the false alarm constraint ALR2FA tends to infinity. The optimality of CUSUM in minimizing the ESADD for non-asymptotic ARL2FA was shown in [7]. The QCD problem in the context of a wireless network, that introduces delays and possible losses in the availability of observations, has not received much attention in the literature to the best of our knowledge. In [8], this problem has been studied in a Bayesian setting where multiple sensors, that observe the same process, transmit their observations to a fusion center over a wireless network. While there is no packet loss in this formulation, there is asynchrony in the arrival (at the DM) of the batch of measurement packets from the same sample, leading to the question about how the partial arrivals from a batch should be processed.

b) Outline of the paper: The rest of the paper is organized as follows. In Sec. II, we describe our system model and describe the QCD problem statement. We provide a CUSUM algorithm for QCD in Sec. III and prove the asymptotic optimality of our algorithm in Sec. IV. Sec. V discusses the applicability of our approach to certain extensions of the system model. Discussion about the non-asymptotic aspects of the algorithm and numerical results are provided in Sec. VI. Sec. VII concludes the paper.

II. System Model and Notation

We consider a discrete-time system where a sensor node samples a random process at a sampling rate $0 < r < 1$, which is the probability that a new sample is generated in each discrete time interval. The sample is encapsulated in a packet, and immediately added to the transmit queue of the transmitter of a wireless link connecting the sensor to the DM. The wireless channel is slotted with a known packet loss probability. If its queue is nonempty at the beginning of a slot, the transmitter transmits one packet; if the packet succeeds an acknowledgment is received back in the same slot, else the packet is backlogged for reattempt in the next slot. We assume that packets are attempted till success. The time slots are of unit size (in practice the time taken to transmit one packet and receive its acknowledgment, along with the interpacket gaps and guard times) and are indexed by $k \in \mathbb{Z} = \{\ldots, -1, 0, 1, 2, \ldots\}$, where slot k refers to the time interval $(k-1, k]$. We assume that the nodes in the network are all time synchronized. Further, we assume in-order reception of the transmitted packets, by the sensor node, at the DM. The sensor node samples a measurement X_j at time denoted t_j, where $j = 1, 2, \ldots$. The measurements are independent, and have a probability distribution

$$X_j \sim \begin{cases} f_0 & \text{if } t_j < \nu, \\ f_1 & \text{if } t_j \geq \nu, \end{cases}$$

where $\nu \geq 1$ is an unknown deterministic time, referred to as the *change point* at which the distribution of the observations changes from a known distribution f_0 to a known distribution f_1. We assume that this change point occurs at the end of the slot ν. This change in the distribution of the measurements may occur due to the development of a fault in one of the components of the machinery, whose health is being monitored by the sensor node. Further, the channel over which the sensor node transmits to the DM has a probability of successful transmission p_0 for $k \leq \nu$ and p_1 for $k > \nu$. We assume that the channel is conditionally independent of the sampling process given the change point ν.

The problem that we set out to solve in this work is for the DM to detect the change in the distribution of the samples as quickly as possible, when the DM receives data sequentially, while controlling the false alarms to be below a given threshold. We first define the notation to be used in the rest of the paper before we state our problem formally.

- $P_\nu, \mathbb{E}_\nu$ denote the underlying probability law and the expectation, when the change occurs in the slot ν. In particular, P_0 denotes the probability law when the change has already occurred before the start of the detection procedure.
- $P_\infty, \mathbb{E}_\infty$ denote the probability law, and the expectation, when the change does not occur ($\nu = \infty$).
- S_ν denotes the number of packets that arrived at the transmit queue after the QCD process starts until the change point ν, i.e., it counts the number of arrivals in the time $(0, \nu]$.
- Q_k is the number of samples in the queue at the beginning of the time slot k (see Fig.1). Packet arrivals into the queue during the time $(k-2, k-1]$ are accounted for in Q_k. The sensor node attempts a transmission in slot k if $Q_k > 0$.
- D_k denotes the index of the last measurement packet successfully received at the DM up to the end of slot k.
- Y_k denotes the packet loss/success observed at the DM. Y_k is defined below.
- Z_k is the received measurement at the DM in slot k. To denote that there was no transmission on the channel due to an empty transmit queue, we say that the channel service process and the received data is $\emptyset$. (Y_k, Z_k) are defined as

$$(Y_k, Z_k) = \begin{cases} (1, X_{D_k}), & \text{on successful transmission}, \\ (0, *), & \text{on unsuccessful transmission}, \\ (\emptyset, \emptyset), & \text{on no transmission}. \end{cases}$$

- The probability of successful transmission over the channel is

$$P(Y_k = 1 \mid Q_k > 0) = \begin{cases} p_0 & \text{if } k \leq \nu, \\ p_1 & \text{if } k > \nu. \end{cases}$$

  We assume that p_0, p_1 are known. To ensure stability of the transmit queue, we make the following assumption: **Assumption:** The sampling rate $r < \min\{p_0, p_1\}$.

The DM uses a sequential algorithm to detect the change point using the observations of the packet loss/success process and the measurement samples that it receives. Denote by $\mathcal{F}_t = \sigma(Y_k, Z_k; 1 \leq k \leq t)$, the σ-algebra generated by the observations available at the DM up to the edge of slot

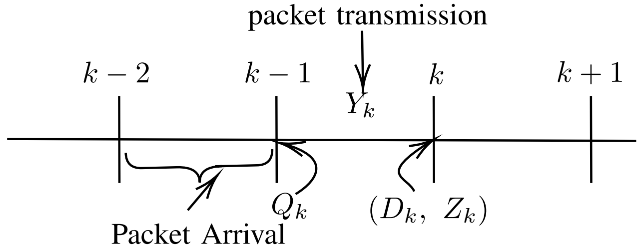

Fig. 1. Slot k corresponds to time $\in [k-1, k)$. The change point occurs at the end of the time slot ν.

t. Then, the DM uses a sequential detection rule, an $\{\mathcal{F}_t\}$-stopping time T to raise an alarm declaring that a change has been detected. We measure the performance of the detection rule, following Lorden's approach [1], in terms of the average delay in the detection of the change point, measured using the ESADD. We aim to minimize ESADD subject to the ARL2FA $\geq \gamma$, where $\gamma > 0$ is a large positive number, i.e., find a T^* that solves

$$\inf_{T \in \mathcal{C}_\gamma} \underbrace{\sup_{\nu \geq 1} \operatorname{esssup} \mathbb{E}_\nu \left[T - \nu + 1 \mid Z_1^{\nu-1}, Y_1^{\nu-1}\right]}_{\doteq \overline{\mathbb{E}}_1[T]}, \quad (1)$$

where $\mathcal{C}_\gamma = \{T : \mathbb{E}_\infty[T] \geq \gamma\}$, $Z_1^{\nu-1} = \{Z_1, \ldots, Z_{\nu-1}\}$, and $Y_1^{\nu-1} = \{Y_1, \ldots, Y_{\nu-1}\}$.

III. Log Likelihood Ratio Analysis

We assume that the transmitter queue length at the beginning of the first slot, Q_1, is known a priori to the DM. The packets already in the transmitter queue, before the QCD process starts, will need to be discarded by the DM. The information Q_1 could be conveyed to the DM by a control packet that initializes the QCD process. As before, we assume that control packets (initialization, acknowledgments) always succeed (due to them being small, and being transmitted with a more robust scheme that is not costed in our current model).

The log-likelihood ratio of P_ν vs. P_∞, based on observations at the DM at time n, is $\log \frac{P_\nu(Y_1^n, Z_1^n | Q_1)}{P_\infty\left(Y_1^n, Z_1^n | Q_1\right)}$. Define

$$L_i = \mathbf{1}_{\{Q_i > 0\}} \log \frac{P_0(Y_i \mid Q_i > 0)}{P_\infty(Y_i \mid Q_i > 0)} + \mathbf{1}_{\{Y_i = 1, D_i > Q_1 + S_\nu\}} \log \frac{f_1(X_{D_i})}{f_0(X_{D_i})}. \quad (2)$$

The following lemma shows that the log-likelihood ratio can be written as a sum of likelihood functions, L_i, with each being a function of the observations at slot i alone.

Lemma 1. *With L_i as in* (2)*, the log-likelihood ratio satisfies*

$$\log \frac{P_\nu(Y_1^n, Z_1^n \mid Q_1)}{P_\infty(Y_1^n, Z_1^n \mid Q_1)} = \sum_{i=\nu+1}^{n} L_i. \quad (3)$$

The proof (provided in appendix) uses the fact that the arrival process of the measurement samples is the same under both the probability laws P_ν and P_∞. □

The DM uses the CUSUM rule [5] with a CUSUM update L_i for detection of the change point. The detection rule is an $\mathcal{F}_t$–stopping time T, defined as

$$T(h) = \min\left\{n \in \mathbb{N} : C_n > h, C_n = (C_{n-1} + L_n)^+\right\}, \quad (4)$$

where we initialize $C_0 = 0$. The decision threshold h, is tuned to achieve the target false alarm performance.

IV. Asymptotic Analysis

In this section, we analyze the performance of CUSUM (4), in the asymptotic regime as $\gamma \to \infty$. Since we assume $r < \min\{p_0, p_1\}$, the transmit queue is stable. Under the probability law P_ν, we have a Geom/Geom/1 queue [9] with arrival rate r and service rate p_1.

Define $\zeta_k = (Q_k, Y_k, Z_k)$; see Fig. 1 for the embedding of the component processes. It is clear from the evolution of queue dynamics defined in Sec. III that ζ_k is a Markov process, given the change point ν. Under the probability law P_0, where the channel success probability is p_1, the stationary distribution of the Markov process ζ_k is given by $\Pi_\zeta^{(p_1)}$. The log-likelihood ratio of ζ_1^n, given Q_1, under P_ν versus P_∞ is equal to

$$\ell_n = \log \frac{P_\nu(\zeta_1^n \mid Q_1)}{P_\infty(\zeta_1^n \mid Q_1)} = \log \frac{P_\nu(Z_1^n, Y_1^n \mid Q_1)}{P_\infty(Z_1^n, Y_1^n \mid Q_1)}$$

since, given $\left(Q_1, Y_1^k,\right)$, Q_k is equal under both the probability laws, due to the arrival process not depending on the hypotheses. By Lemma 1, we then have $\ell_n = \sum_{i=1}^n L_i$. Define the stationary expected value of the log-likelihood ratio update L_i under the distribution $\Pi_\zeta^{(p_1)}$ as

$$I = \mathbb{E}_{\Pi_\zeta^{(p_1)}}[L_1] = P_0(Q_1 > 0)\left(\mathcal{I}(p_1, p_0) + p_1 \mathcal{I}(f_1, f_0)\right), \quad (5)$$

where $\mathcal{I}(p_1, p_0) = \mathbb{E}_{\Pi_\zeta^{(p_1)}} \log \frac{P_0(Y_k | Q_k > 0)}{P_\infty(Y_k | Q_k > 0)}$ and is the Kullback-Leibler divergence between the Bernoulli distributions with parameters p_1 and p_0, i.e., $p_1 \log \frac{p_1}{p_0} + (1 - p_1) \log \frac{1-p_1}{1-p_0}$. More generally, the quantity $\mathcal{I}(f_1, f_0) = \mathbb{E}_{\Pi_\zeta^{(p_1)}} \log \frac{f_1(X_j)}{f_0(X_j)}$ is the Kullback-Leibler divergence between the distributions of the samples f_1 and f_0. The quantity $P_0(Q_1 > 0)$ can be computed using the Little's theorem [9] for queues. For a Geom/Geom/1 queue with arrival probability r and service probability p_1, $P_0(Q_1 > 0) = r/p_1$. The quantity I can be rewritten as $I = r\left(\frac{1}{p_1}\mathcal{I}(p_1, p_0) + \mathcal{I}(f_1, f_0)\right)$.

Since $\{\zeta_k\}$ is an aperiodic and recurrent Markov process, by the ergodic theorem for Markov processes [10], we have

$$\lim_{n \to \infty} \frac{1}{n} \ell_n = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} L_i = \mathbb{E}_{\Pi_\zeta^{(p_1)}}[L_1] = I. \quad (6)$$

In the sequel, we prove the asymptotic optimality of the CUSUM rule (4) by showing that a lower bound exists on the ESADD defined in (1) when the ARL2FA $\mathbb{E}_\infty[T] \geq \gamma$, and then by proving that the CUSUM rule defined in (4) achieves the lower bound asymptotically. The proofs for the theorems are provided in the Appendix, where we use Lai's

[6] generalization of Lorden's asymptotic theory to general processes to prove the claims.

Theorem 2 (Lower bound on ESADD)**.** *For the Markov process ζ_k, as the ARL2FA $\gamma \to \infty$, we have*

$$\inf \left\{\overline{\mathbb{E}}_1 [T] : \mathbb{E}_\infty [T] \geq \gamma\right\} \geq \left(I^{-1} + o(1)\right) \log \gamma.$$

Next, we show that there exists an upper bound on the ESADD of the CUSUM detector (4).

Theorem 3 (Upper bound on CUSUM ESADD)**.** *For the CUSUM (4) with a threshold h, we have $\mathbb{E}_\infty [T] < \infty$ and*

$$\overline{\mathbb{E}}_1 [T] \leq \left(I^{-1} + o(1)\right) h, \text{ as } h \to \infty.$$

By the above two theorems, we claim that, as $\gamma \to \infty$, the CUSUM algorithm (4) achieves the lower bound on the ESADD with a threshold $h = \log \gamma$, and is hence an asymptotically optimal sequential detection algorithm. For a threshold h, and I defined as in (5), the detection rule thus has ESADD $\overline{\mathbb{E}}_1 [T] \approx \frac{h}{I} (1 + o(1))$.

V. Extensions

In this section, we show that our analysis holds for two simple extensions of the system model.

A. Periodic sampling

In Sec. II, we assumed that the measurements are sampled such that the probability of a packet arrival at the transmit queue at each slot is given by the Bernoulli distribution with parameter r. Suppose instead that we consider that the sample measurements are produced by a periodic sampler, with sampling interval $s = 1/r$. Then, there is a packet arrival at the transmit queue once every s time instants. We first note that since the packet arrivals are independent of the change, Lemma 1 holds for this case, and the DM uses the same CUSUM algorithm (4) in this case too. The analysis in Sec. IV crucially uses the Markov property of ζ_k. To preserve the Markovian property in this case, we need to augment the state space of observations by defining $\zeta_k = (Q_k, V_k, Y_k, Z_k)$, where $V_k, k \in \mathbb{N}$ counts the number of slots to the next packet arrival. Note that given V_1, and the sampling interval, V_k can be computed for all $k > 1$. The analysis in Sec. IV too holds for this case after augmenting the state space of observations.

B. QCD over a lossy link with no retransmissions

In Sec. II, we assumed that the packet transmissions are attempted till success, and that the DM receives every measurement sample that the sensor node samples. Suppose now that the transmitter uses a best-effort service, i.e., it attempts no retransmissions upon packet transmission failure. The sensor, on sampling a new measurement, immediately places it in a new packet and adds it to transmit queue. The transmitter transmits this packet in the next slot and removes the packet from the queue. In this case, the DM does not receive all the measurement samples that the sensor node samples. The DM must make a decision on whether a change has occurred, based on the samples that it has received.

In this case too, since the packet arrival process at the transmitter is the same under both the probability laws, Lemma 1 holds, and the DM makes a decision using the CUSUM detector (4). Note that $Q_k \in \{0, 1\}$, when there are no retransmissions, and Q_k is i.i.d. whenever the packet arrivals are i.i.d. and $P_0 (Q_1 > 0) = r$, where r is the packet arrival rate, as before. The expected value of the CUSUM update L_k under P_0 is

$$I = \mathbb{E}_1 [L_1] = r \left(\mathcal{I} (p_1, p_0) + p_1 \mathcal{I} (f_1, f_0)\right). \qquad (7)$$

The process $\{\zeta_k : k \geq 1\}$ is i.i.d. before and after the change, and the usual CUSUM calculations [11] hold. The ESADD in this case for the CUSUM detector (4) with a threshold h is also $h(I^{-1} + o(1))$, where I is given by (7).

VI. Discussion and Numerical Analysis

In Sec. IV, the asymptotic performance of the CUSUM algorithm, when the ARL2FA $\to \infty$, was analyzed. Moustakides et al. [12] state that when the observation process is Markov, the optimum threshold in the non-asymptotic regime is a function of the initial state. However, in our analysis in Sec. IV, we find that the asymptotic performance, as ARL2FA $\to \infty$, of our CUSUM algorithm is independent of the initial state of the Markov process ζ_k. In the asymptotic regime, as the target ARL2FA goes to ∞, the CUSUM threshold also goes to ∞ [1]. In this regime, the Markov process $\{\zeta\}$ approaches the stationary state under the post-change distribution. Then, the initial state makes little difference to the overall performance of the detection algorithm.

In Fig. 2, we plot the the ratio of the simulated ADD to the threshold (h) against the average sampling interval ($s = 1/r$). To generate the plot, in the system described in Sec.II, we fix the channel parameters as $p_0 = 0.61$ and $p_1 = 0.60$. We take the sensor measurement distribution to be Normal with a mean $\mu_0 = 0$ before the change and a mean $\mu_1 = 10$ after the change; the variance is held constant at $\sigma^2 = 1/2$. For each run of the simulation, we sample Q_1 from the stationary distribution of the queue, with arrival r and service probability p_0. We fix the change point to have occurred before the start of the detection procedure and run each simulation until the CUSUM statistic crosses the threshold h. The ADD plotted in Fig. 2 is averaged over 10^6 repetitions. We compare the ratio ADD$/h$ for various values of h to the asymptotic value $\frac{1}{I}$, computed in Theorem 3. We note that, ADD$/h$ approaches $1/I$ as $h \to \infty$

In practice, the transmit queue adds a queuing delay component (denoted by $\overline{d}_Q$) to the overall detection delay. The queuing delay consists of two parts: a) the time required for the packets already present in the transmit queue at the first slot to be flushed out, b) the delay that the last packet transmitted by the sensor node to the DM before a decision about the change is made at the DM. If the sampling rate r is close to p_0, the former delay is high; if r is close to p_1, the latter delay is high. It is thus desirable to keep the r small to reduce $\overline{d}_Q$. The sampling process, too, adds a sampling delay component (denoted by $\overline{d}_S$) to the overall delay to detection.

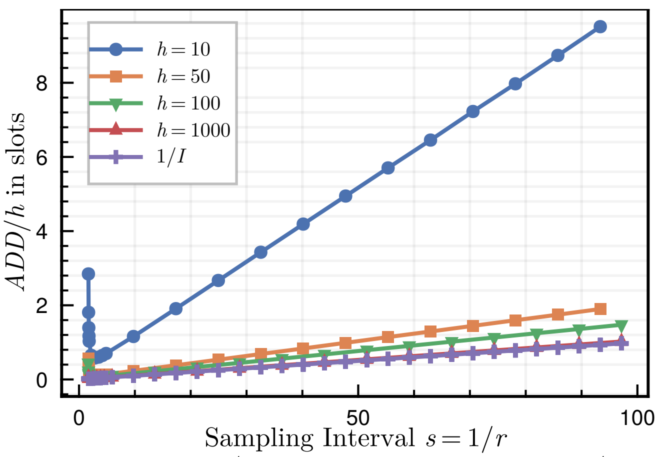

Fig. 2. Simulated 'ADD/h' vs Average Sampling Interval ($s = 1/r$) for various values of threshold h. Notice how the simulated ADD/h approaches the asymptotic ESADD/$h = 1/I$, computed in Theorem 3 as h increases.

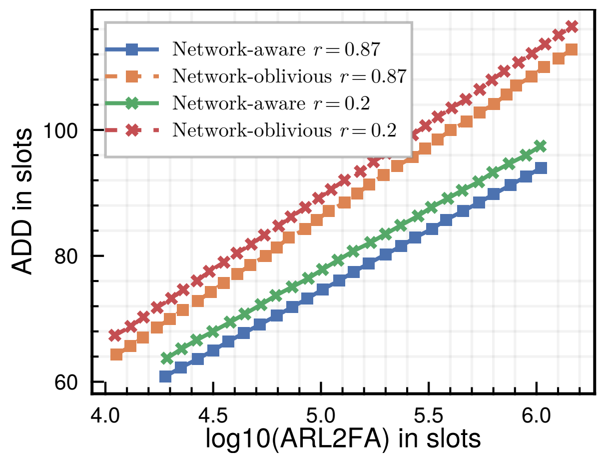

Fig. 4. Simulated ADD vs ARL2FA for various Sampling Rate r, where $p_0 = 0.95, p_1 = 0.9$. Plot compares the simulated performance of network-oblivious detection with that of network-aware detection.

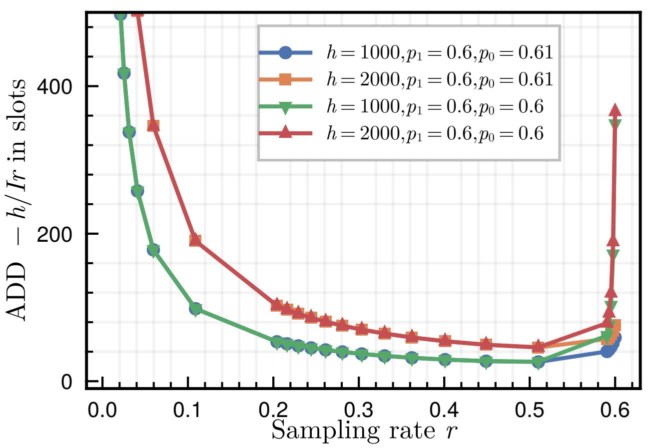

Fig. 3. Simulated ADD vs Sampling Rate r for various values of threshold h. The queuing delay $\overline{d}_Q$ increases as $r \uparrow p_1$. The sampling delay $\overline{d}_S$ increases with decreasing r.

The average sampling delay with Bernoulli arrivals at a rate r is $1/r$. Decreasing the sampling rate increases the sampling delay.

Fig. 3 shows the trade-off that exists between increasing the sampling rate to decrease $\overline{d}_S$ and decreasing the sampling rate to keep the $\overline{d}_Q$ low. We plot the ADD for two values of the threshold h, and sweep the sampling rate r. The simulation parameters to produce Fig. 3 are the same as those in Fig. 2. In Fig. 3, we show the increase in the sampling delay with a decrease in the sampling rate. When r is swept close to $p_1 = 0.61$, we see that the ADD shoots up – this corresponds to the increase in $\overline{d}_Q$ as the sampling rate becomes close to the queue service rate p_1. In Fig. 2, the gap between the ESADD$/h = 1/I$ curve and the simulated ADD$/h$ curves correspond to $(\overline{d}_S + \overline{d}_Q)/h$. Hence, as the h increases, the difference between the two curves reduces.

In Fig. 4, we plot the simulated ADD against the simulated ARL2FA, for different values of sampling rate r. To generate this plot, we fix $p_0 = 0.95, p_1 = 0.90$, and take the sensor measurements distribution to be Normal with mean $\mu_0 = 0$ prior to the change, and $\mu_1 = 1$ after change; the variance is held constant at $\sigma^2 = 1/2$. The figure confirms the well known [11] linear growth of ADD with increase in the $\log$(ARL2FA). Fig. 4 also shows the advantage of the network-aware detection procedure that we describe in this work over a detector that is network-oblivious. The network-oblivious DM too uses a CUSUM detector; but, being network oblivious, it updates the CUSUM statistic only on the successful reception of a packet, and does not use the channel service process observations $\{Y_k\}$. We see from the plot that the network-aware CUSUM detector has a much lower ADD compared to the network-oblivious detector for the same ARL2FA. In Fig. 3, we show the importance of choosing the right sampling rate r for a given channel loss probability. The sampling delay $\overline{d}_S$ and $\overline{d}_Q$ are added to the detection delay even when the channel success probabilities do not change with the change point. A network-oblivious detector, on account of not having information about the channel state, may choose a mismatched sampling rate, resulting in a large delay penalty.

To conclude, we study the performance of the CUSUM algorithm, defined in (4), using numerical simulations, and assert the accuracy of our analysis by comparing against simulated runs of the CUSUM. Further, we show that a network-aware DM that uses the additional information provided by the channel performs much better than a network-oblivious DM that only uses information in the received packets.

VII. Conclusion

We have studied the problem of QCD when a remote DM receives delayed measurement samples from a sensor node that transmits its measurements over a lossy link. For the setting when the channel loss probability changes after the change point, we design a CUSUM algorithm and show its optimality, by modeling the observation process as Markovian. We study the non-asymptotic detection delay of our algorithm numerically and show a trade-off between the sampling delay and the queuing delay. We also show the gains made by a detector that is network-aware. We plan to study the case of multiple sensors transmitting their measurements over a network in future work.

REFERENCES

[1] G. Lorden, "Procedures for Reacting to a Change in Distribution," *The Annals of Mathematical Statistics*, vol. 42, pp. 1897–1908, Dec. 1971.

[2] A. Rai and S. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, vol. 96, pp. 289–306, Apr. 2016.

[3] "Schaeffler OPTIME Condition Monitoring." https://medias.schaeffler.co.in/en/monitor/optime-condition-monitoring.

[4] "IPv6 over the TSCH mode of IEEE 802.15.4e (6TiSCH)." https://datatracker.ietf.org/wg/6tisch/about/.

[5] E. S. Page, "Continuous Inspection Schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.

[6] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, pp. 2917–2929, Nov. 1998.

[7] G. V. Moustakides, "Optimal Stopping Times for Detecting Changes in Distributions," *The Annals of Statistics*, vol. 14, pp. 1379–1387, Dec. 1986.

[8] P. Karumbu, V. K. Prasanthi, and A. Kumar, "Delay optimal event detection on ad hoc wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 8, pp. 12:1–12:35, Mar. 2012.

[9] L. Kleinrock, *Theory, Volume 1, Queueing Systems*. USA: Wiley-Interscience, 1975.

[10] R. Durrett, *Probability: Theory and Examples*. No. 49 in Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge ; New York, NY: Cambridge University Press, fifth edition ed., 2019.

[11] A. Tartakovsky, I. V. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. No. 136 in Monographs on Statistics and Applied Probability, Boca Raton, Fla.: CRC Press, 2015.

[12] G. V. Moustakides and A. G. Tartakovsky, "A note on "The Optimal Stopping Time for Detecting Changes in Discrete Time Markov Processes" by Han and Tsung," *Sequential Analysis*, vol. 29, no. 4, pp. 483–486, 2010.

APPENDIX I
PROOF OF LEMMA 1

We write

$$\log \frac{P_\nu\left(Y_1^n, Z_1^n \mid Q_1\right)}{P_\infty\left(Y_1^n, Z_1^n \mid Q_1\right)} = \sum_{i=1}^{n} \log \frac{P_{i,\nu}\left(Y_i \mid Q_1, Y_1^{i-1}\right)}{P_{i,\infty}\left(Y_i \mid Q_1, Y_1^{i-1}\right)} + \sum_{i=1}^{n} \log \frac{P_{i,\nu}\left(Z_i \mid Q_1, Y_1^{i}, Z_1^{i-1}\right)}{P_{i,\infty}\left(Z_i \mid Q_1, Y_1^{i}, Z_1^{i-1}\right)}, \quad (8)$$

where $P_{i,\nu}$ and $P_{i,\infty}$ are the probability distributions at time i, given that the change occurs at a finite ν and $\nu = \infty$ (no change), respectively. We drop the dependence of Y_i on Z_1^i since $Y_i \perp\!\!\!\perp Z_i$ given $\left(Q_1, Y_1^{i-1}\right)$. To simplify the first term in (8), we note that $\log \frac{P_{i,\nu}\left(Y_i|Q_1,Y_1^{i-1}\right)}{P_{i,\infty}\left(Y_i|Q_1,Y_1^{i-1}\right)} = \mathbf{1}_{\{Q_i>0\}} \log \frac{P_{i,\nu}\left(Y_i|Q_1,Y_1^{i-1}\right)}{P_{i,\infty}\left(Y_i|Q_1,Y_1^{i-1}\right)}$, since $Y_i = \emptyset$ w.p. 1 on the set where $\{Q_i = 0\}$ under both the probability laws. Further, $\mathbf{1}_{\{Q_i>0\}} \log \frac{P_{i,\nu}\left(Y_i|Q_1,Y_1^{i-1}\right)}{P_{i,\infty}\left(Y_i|Q_1,Y_1^{i-1}\right)} = \mathbf{1}_{\{Q_i>0\}} \log \frac{P_{i,\nu}(Y_i|Q_i>0)}{P_{i,\infty}(Y_i|Q_i>0)}$, because Y_i depends only on Q_i, $P\left(Q_i \mid Q_1, Y_1^i\right)$ depends only on the arrival process, and $P\left(Y_i \mid Q_i = 0\right) = \mathbf{1}_{\{Y_i=\emptyset\}}$ under both the probability laws. Then, the first summand in (8) can be simplified to $\sum_{i=\nu+1}^{n} \mathbf{1}_{\{Q_i>0\}} \log \frac{P_0(Y_i|Q_i>0)}{P_\infty(Y_i|Q_i>0)}$.

To simplify the second term in (8), note that D_i can be determined using $\left(Q_1, Y_1^i\right)$, and that Z_i depends only on (Y_i, D_i). Hence, $P\left(Z_i \mid Q_1, Y_1^i, Z_1^{i-1}\right) = P\left(Z_i \mid D_i, Y_i\right)$ under both the probability laws. Further, on the sample paths where $\{Y_i = 0\}$, we have $Z_i = \emptyset$. Also, on the sample paths where $\{Y_i = 1\}$, we have $Z_i = X_{D_i}$. Hence, we write $\log \frac{P_{i,\nu}\left(Z_i|Q_1,Y_1^i,Z_1^{i-1}\right)}{P_{i,\infty}\left(Z_i|Q_1,Y_1^i,Z_1^{i-1}\right)} = \mathbf{1}_{\{Y_i=1\}} \log \frac{P_{i,\nu}\left(X_{D_i}|D_i,Y_i=1\right)}{P_{i,\infty}\left(X_{D_i}|D_i,Y_i=1\right)}$. Thus, the second term in (8) can be written as $\sum_{i=\nu+1}^{n} \mathbf{1}_{\{Y_i=1, D_i \geq s_\nu + Q_1\}} \log \frac{f_1\left(X_{D_i}\right)}{f_0\left(X_{D_i}\right)}$.

APPENDIX II
PROOF OF THEOREM 2

We will need the following Lemma to prove Theorem 2.

Lemma 4. *Given $\{S_k : k \geq 1\}$, a sequence of i.i.d. random variables with $\mathbb{E}[S_1] < \infty$, if $\frac{1}{n}\sum_{i=1}^n S_i \xrightarrow{a.s.} 0$, then*

$$\lim_{n\to\infty} P_0\left(\max_{k\leq n} \frac{1}{n}\sum_{i=1}^{k} S_i \geq \delta\right) = 0, \forall \delta > 0.$$

Proof. Define $K_n = \operatorname{argmax}_{k\leq n} \frac{1}{n}\sum_{i=1}^k S_i$, a random variable. For each ω in the set $\left\{\omega : \max_{k\leq n} \frac{1}{n}\sum_{i=1}^k S_i \geq \delta\right\}$, $K_n(\omega) < n$, and $K_n(\omega) \to \infty$ as $n \to \infty$. Thus, $\forall \delta > 0$, $P_0\left(\frac{1}{n}\sum_{i=1}^{K_n} S_i \geq \delta\right) \leq P_0\left(\frac{1}{K_n}\sum_{i=1}^{K_n} S_i \geq \delta\right) \xrightarrow{n\to\infty} 0$ since $\frac{1}{n}\sum_{i=1}^n S_i \xrightarrow{a.s.} 0$ implies in probability convergence. This concludes the proof. □

Proof for Theorem 2. To prove this, we make use of Lai's [6, Thm 1] lower bound on the asymptotic ESADD. Following the discussion in [6, Sec. IV], we only need to show

$$\lim_{n\to\infty} \sup_x P_0\left\{\max_{t\leq n} \sum_{i=1}^{t} L_i \geq I\left(1+\delta\right)n \mid \zeta_0 = x\right\} = 0, \quad (9)$$

for each $\delta > 0$. From (6), we obtain $\frac{1}{n}\sum_{i=1}^n L_i \xrightarrow{a.s.} I$. Let us define $\lambda_i = L_i - I$ and apply Lemma 4 to show that

$$\forall x, \forall \delta > 0, P_0\left\{\max_{t\leq n} \frac{1}{n}\sum_{i=1}^{t} \lambda_i > \delta \mid \zeta_0 = x\right\} \to 0.$$

To show that (9) is true, it is sufficient to show that $\sup_{\zeta_0=x} \frac{1}{n}\sum_{i=1}^n L_i \xrightarrow{a.s.} I$. We note that since $\{\zeta_k\}$ is a Markov process, it is sufficient to only show that $\sup_{\zeta_0=x} L_1 < \infty$. This can be easily observed by noting that the first term in L_1 (see (2)) is bounded whenever $0 < p_0, p_1 < \infty$, and the second term in L_1 depends on ζ_0 only through indicator functions. This proves the assertion that $\sup_{\zeta_0=x} \sum_{i=1}^n L_i \xrightarrow{a.s.} I$, and hence the theorem. □

APPENDIX III
PROOF OF THEOREM 3

To prove this claim, we make use of Lai's [6, Thm. 4] upper bound on the asymptotic ESADD, for a CUSUM detector, with threshold h. As before, following the discussion in [6, Sec. IV], we need to show that

$$\lim_{n\to\infty} \sup_x P_0\left\{\sum_{i=1}^{n} L_i < \left(I-\delta\right)n \mid \zeta_0 = x\right\} = 0, \quad (10)$$

for each $\delta > 0$. For each x, and $\forall \delta > 0$, the probability $P_0\left\{\sum_{i=1}^n L_i < \left(I-\delta\right)n \mid \zeta_0 = x\right\}$ approaches zero since almost sure convergence implies convergence in probability of $\frac{1}{n}\sum_{i=1}^n L_i$. To prove (10), we follow a similar approach as in the proof of the previous theorem.