Asymptotic Optimality in Byzantine Distributed Quickest Change Detection

Yu-Chih Huang, Yu-Jui Huang, and Shih-Chun Lin

Abstract

The Byzantine distributed quickest change detection (BDQCD) is studied, where a fusion center monitors the occurrence of an abrupt event through a bunch of distributed sensors that may be compromised. We first consider the binary hypothesis case where there is only one post-change hypothesis and prove a novel converse to the first-order asymptotic detection delay in the large mean time to a false alarm regime. This converse is tight in that it coincides with the currently best achievability shown by Fellouris *et al.*; hence, the optimal asymptotic performance of binary BDQCD is characterized. An important implication of this result is that, even with compromised sensors, a 1-bit link between each sensor and the fusion center suffices to achieve asymptotic optimality. To accommodate multiple post-change hypotheses, we then formulate the multi-hypothesis BDQCD problem and again investigate the optimal first-order performance under different bandwidth constraints. A converse is first obtained by extending our converse from binary to multi-hypothesis BDQCD. Two families of stopping rules, namely the simultaneous d-th alarm and the multi-shot d-th alarm, are then proposed. Under sufficient link bandwidth, the simultaneous d-th alarm, with d being set to the number of honest sensors, can achieve the asymptotic performance that coincides with the derived converse bound; hence, the asymptotically optimal performance of multi-hypothesis BDQCD is again characterized. Moreover, although being shown to be asymptotically optimal only for some special cases, the multi-shot d-th alarm is much more bandwidth-efficient and energy-efficient than the simultaneous d-th alarm. Built upon the above

This paper was presented in part at the 2019 IEEE International Symposium on Information Theory [1], [2].

Y.-C. Huang is with the Department of Communication Engineering, National Taipei University, 237 Sanxia District, New Taipei City, Taiwan (email: ychuang@mail.ntpu.edu.tw).

Y.-J. Huang is with the Department of Applied Mathematics, University of Colorado at Boulder, CO 80309 , USA (email: yujui.huang@colorado.edu)

S.-C. Lin is with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, 106 Daan District, Taipei City, Taiwan (email: sclin@ntust.edu.tw)

The authors are ordered alphabetically.

success in characterizing the asymptotic optimality of the BDQCD, a corresponding leader-follower Stackelberg game is formulated and its solution is found.

I. Introduction

The problem of quickest change detection (QCD), a.k.a. sequential change detection, studies detecting an abnormal event as quickly as possible after its occurrence at a deterministic but unknown time, subject to a certain false alarm rate. It has many applications and has been extensively researched since the early works [3], [4], [5]. In these works, it is assumed that there is only one post-change hypothesis, which we refer to as the binary case. When there are multiple post-change hypotheses, the problem is referred to as multi-hypothesis QCD and has been investigated in [6], [7]. A nice tutorial on QCD can be found in [8]. However, recent applications of cyber-physical systems (CPS) [9] typically involve multiple distributed sensors monitoring the event and reporting their observations to the fusion center via bandwidth-limited links. For example, the abnormal changes of voltage waveforms in smart grids are harmful to delicate electronic devices and recent advances of massive machine-type communications (mMTC) or internet of things (IoT) [10] allow the usage of advanced cyber-physical infrastructures for monitoring voltage quality events distributively [11]. Moreover, some sensors, whose identities are unknown to the fusion center, may be compromised and may try to sabotage the detection task. Motivated by these applications, this paper considers the decentralized version of QCD, where a fusion center monitors the event through distributed sensors, with compromised sensors collaboratively forming attack. This problem has been studied in [12], [13] and is called Byzantine distributed QCD (BDQCD).

In [12], a special case of binary BDQCD, only one compromised sensor is considered. A decision rule called second-alarm rule, where the fusion center declares the occurrence of the event once it receives the second local report from sensors, is proposed and its asymptotic performance is analyzed. In [13], the general binary BDQCD problem with infinite-bandwidth links and that with 1-bit links are investigated. Multiple rules are proposed and their corresponding asymptotic performance are analyzed. In the presence of infinite-bandwidth links, the low-sum CUSUM scheme proposed in [13] achieves the best asymptotic performance among the schemes in [13]. Among the rules with 1-bit links proposed in [13], the voting rule that declares the occurrence of the event after the number of received local reports exceeds a certain threshold has the best first-

order asymptotic performance. When the threshold is set to be the total number of honest sensors, the asymptotic performance of the voting rule, called the consensus rule in this special case, reaches its maximum and also attains the best asymptotic performance in [13]. Two questions naturally arise from this premise:

1) What is the fundamental limit of the first-order asymptotic performance of binary BDQCD?
2) How to handle this problem when there are multiple post-change hypotheses?

For the first question, despite the exciting results in [12], [13], it is thus far unclear what the best first-order asymptotic performance of binary BDQCD is. Although one non-trivial converse can be easily obtained by assuming that a genie reveals to the fusion center the identities of all honest sensors (this simple converse will be presented in Section III), this converse bound and the best achievable asymptotic performance in [13] do not match. This indicates that either the best achievable scheme thus far is not optimal or the converse is not tight, or both.

The first contribution of this work is to prove a new converse to the first-order asymptotic performance for binary BDQCD. In the proof, we first construct an attack strategy for the compromised sensors and then construct a genie who reveals just enough information to the fusion center. After that, inspired by the proof technique in [14], we transform the original problem into a centralized QCD problem. Last, evaluating the corresponding optimal CUSUM procedure establishes the new converse. The converse turns out to coincide with the best achievable first-order asymptotic performance known to date; thereby, the fundamental limit of the first-order asymptotic performance of binary BDQCD is characterized. Specifically, our converse confirms that both the consensus rule (using only 1-bit links) and the low-sum-CUSUM rule (using infinite-bit links) in [13] achieve the optimal first-order scaling. The first optimality unveils an important implication that, at least asymptotically, *1-bit links suffice even with compromised sensors*. As a byproduct of our proof, we explicitly construct an attack strategy, called the reverse attack, where each compromised sensor generates fake i.i.d. observations according to post-change and pre-change distributions before and after the change time (i.e., with pre-change and post-change distributions swapped), and form local reports based on these fake observations. In spite of abandoning potential cooperation among compromised sensors, this reverse attack turns out to be strong enough for us to prove a tight lower bound on the asymptotic performance of

BDQCD[1]. We note that although our converse is inspired from [14], the one-shot hypothesis testing problem studied in [14] is fundamentally different to our sequential change detection that deals with observation sequences having an unknown change time. More detailed comparisons with [14] are provided at the end of Section IV-B.

Our second contribution is to tackle the second question listed above and extend the framework of BDQCD to the multi-hypothesis setting. To this end, we formulate the multi-hypothesis version of the BDQCD problem. We then demonstrate that blindly adopting the existing procedure in [7][13][15] may result in catastrophic events. Two novel stopping rules are proposed and analyzed, which requires $\log(Q)$-bit and Q-bit noiseless links, respectively, for BDQCD with $Q+1$ hypotheses. In our delay analysis, we prove an asymptotic dominance result, which confirms the intuition that although there are multiple hypotheses, for each one being considered, we only have to examine the statistics of another hypothesis that is closest (in the sense of KullbackLeibler (KL) divergence) to the hypothesis being considered. The converse for the binary case is also extended to the multi-hypothesis setting and it is shown that proposed stopping rules can achieve the optimal first-order asymptotic performance under different bandwidth constraints; therefore, the asymptotically optimal performance of multi-hypothesis BDQCD is again characterized.

Last but not least, we formulate a leader-follower Stackelberg game [16] where the fusion center and honest sensors form the leader while the compromised sensors form the follower. The first-order optimality mentioned above yields the game solution, where the leader adopts the aforementioned asymptotically optimal stopping rule and the follower employs the corresponding worst attack.

A. Organization

The rest of the paper is organized as follows. We will separately introduce the problem of binary BDQCD and multi-hypothesis BDQCD in Sections II-A and II-B, respectively, and review the current state-of-the-art in Section III. We will then split our discussion into two parts, namely the binary BDQCD in Section IV and multi-hypothesis BDQCD in Section V, even though the former is a special case of the latter. The main reasons are: 1) binary BDQCD is of substantial interest in its own right and has been a subject of research in the literature [12], [13], 2) due

[1]Throughout the paper, such an attack strategy is said to be an asymptotically worst (to the fusion center) attack, or simply a worst attack.

to the nature of having multiple post-change hypotheses, one has to consider a sequence of stopping times, which is in sharp contrast to binary BDQCD where only one stopping time is considered, 3) our proofs for tight converse bounds in Sections IV and V are quite involved with heavy notations and it is better to start with the binary case in Section IV and mention only notable differences later in Section V, and 4) for achievability, we borrow existing results in [13] for the binary BDQCD, while we devise new efficient stopping rules for the multi-hypothesis BDQCD whose first-order asymptotic performance coincides with our converse bound. Finally, a leader-follower Stackelberg game and its solution are then presented in Section VI.

B. Notational conventions

For a positive integer K, define $[K] := \{1, \ldots, K\}$ and $[K]^+ = \{0\} \cup [K]$. Function $(x)^+$ outputs x if $x \geq 0$ and zero otherwise. For two real functions $f_1(x)$ and $f_2(x)$, as $x \to \infty$, we write $f_1(x) \sim f_2(x)$ when $f_1(x)/f_2(x) \to 1$ and $f_1(x) \gtrsim f_2(x)$ when $\liminf(f_1(x)/f_2(x)) \geq 1$. The $o(.)$ and $w(.)$ follow the asymptotic notations in [17].

II. Problem formulation

In this section, we formally state the problem of BDQCD. We will first describe binary BDQCD and then formulate the generalization to the multi-hypothesis case.

A. Binary BDQCD

The binary BDQCD problem consists of a fusion center and K sensors indexed by $[K]$. Among these sensors, there is an *unknown* subset $\mathcal{N} \subset [K]$ of honest sensors, with the remaining $M := K - |\mathcal{N}|$ sensors being potentially compromised. The goal of the honest sensors is to monitor an event and help the fusion center decide whether the event has occurred, while the goal of compromised sensors is to collaboratively confuse the fusion center. Although the exact information about which sensors are honest and which sensors are compromised is unknown, we assume that M, the maximum number of sensors the attacker can compromise, is known by the fusion center. Moreover, it is assumed that there are more honest sensors than compromised sensors, i.e., $|\mathcal{N}| > M$. The observations of all K sensors are sequences of independent random variables with known distributions, subject to the same distribution change at an unknown but deterministic time ν. Before the change time ν, sensor k's observations $X_1^k, X_2^k, \ldots, X_\nu^k$ are

independent and identically distributed (i.i.d.) with the density P_0, while $X_{\nu+1}^k, X_{\nu+2}^k, \ldots$ are i.i.d. with the density P_1. If the change never happens, i.e., $\nu = \infty$, X_t^k are i.i.d. with P_0 for all t. We denote by $\mathbf{X}_t = [X_t^1, X_t^2, \ldots, X_t^K]$ the collection of observations at time t and we use the notation $\mathbf{X}_{t_1}^{t_2}$ for $t_1 < t_2$ to denote the collection $[\mathbf{X}_{t_1}, \mathbf{X}_{t_1+1}, \ldots, \mathbf{X}_{t_2}]$. Also, we define the KL divergence from P_0 to P_1 as [18] to be $I := \int \log\left(\frac{P_1(x)}{P_0(x)}\right) P_1(x)dx$. Throughout the paper, we assume that I is finite and strictly positive and

$$\int \log\left(P_1(x)/P_0(x)\right)^2 P_1(x)dx < \infty. \tag{1}$$

All the local reports from honest or compromised sensors belong to the set $\mathcal{X}$, which satisfies the underlying bandwidth constraint on the noiseless link between each sensor and the fusion center. It is worth emphasizing that this setting encompasses many scenarios discussed in existing works including $\mathcal{X} = \{0, 1\}$ and $\mathcal{X} = \mathbb{R}$ in [13] and $\mathcal{X}$ being a set of finite alphabets in [19]. At each time index t, the honest sensor k individually makes a local decision by mapping its own observations up to time t to an element in $\mathcal{X}$, and then chooses to report it or not according to the adopted reporting mechanism. Based on the received local reports from all sensors, the fusion center adopts a stopping rule to determine when to declare that the event has occurred. A change detection rule includes such a stopping rule and local rules at honest sensors. The M compromised sensors, on the other hand, try to disrupt/confuse the fusion center by sending attack signals in $\mathcal{X}$. We assume a very powerful attacker that knows the exact change-time ν and has access to the current and past observations of all nodes. The symbols sent by the compromised sensors at time t are then produced by g, a function (called an attack strategy) with inputs ν, $\mathbf{X}_1^t$, and the change detection rule. We denote by $\mathcal{G}$ the set of all attack strategies including all possible g with no more than M compromised sensors. Following [13], we analyze the performance of a rule by its worst-case expected detection delay and mean time to a false alarm in the sense of Lorden [4], under the worst attack strategy among $\mathcal{G}$. Specifically, let T be the stopping time of a rule, we define the performance metrics as follows.

- **Detection Delay:** The worst-case mean detection delay

$$\mathcal{D}[T] := \sup_{g \in \mathcal{G}, \nu} \operatorname{ess\,sup} \mathbb{E}_\nu^g[(T - \nu)^+ | \mathbf{X}_1^\nu], \tag{2}$$

  where $\mathbb{E}_\nu^g[.]$ means the expectation is taken w.r.t. P_0 when $t \leq \nu$ and w.r.t. P_1 when $t > \nu$ under the attack strategy $g \in \mathcal{G}$.

- **False Alarm:** Without any abnormal changes (i.e. $\nu = \infty$), the worst-case mean time to a false alarm is

$$\mathcal{A}[T] := \inf_{g \in \mathcal{G}} \mathbb{E}_{\infty}^{g}[T], \tag{3}$$

where $\mathbb{E}_{\infty}^{g}[.]$ means the expectation is w.r.t. P_0 for all t (i.e., $\nu = \infty$) under the attack strategy $g \in \mathcal{G}$.

The main theme of this paper is to investigate the optimal asymptotic behavior of how the expected detection delay scales with the mean time to a false alarm in the worst case. Specifically, for an optimal BDQCD rule with stopping time T satisfying $\mathcal{A}[T] \geq \gamma$, we want to characterize how $\mathcal{D}[T]$ grows with γ as $\gamma \to \infty$.

Remark II.1. *We would like to emphasize that our setting is slightly different from that in [13]. In [13], among honest sensors, some are affected and some are unaffected by the change. Similar scenarios with no Byzantine attack are also considered in the literature, with exactly one unknown sensor [20], [21] or a subset of all sensors [22], [23] being affected. For the unaffected, their observations are sampled i.i.d. from P_0 even after the change. In this paper, we do not consider unaffected sensors purely to avoid heavy notation. With slight modifications, our results can be easily extended to include unaffected sensors. This statement remains true even for the multi-hypothesis setting discussed later.*

B. Multi-Hypothesis BDQCD

For the multi-hypothesis version of BDQCD, we again consider a network with a fusion center and K distributed sensors. There is an *unknown* subset $\mathcal{N} \subset [K]$ of honest sensors, with the remaining $M := K - |\mathcal{N}|$ sensors being compromised. The fusion center tries to monitor an abrupt event and decide whether the event has occurred regardless of which type it is[2]. The observations of all K sensors are sequences of independent random variables with known distributions, subject to the same distribution change at an unknown but deterministic time ν. After this distribution change, there are Q different possible types. Specifically, let P_0 be the pre-change probability density function (PDF) and $P_1, \ldots, P_Q$ the post-change PDF corresponding to the states $1, \ldots, Q$, respectively. For each $k \in [K]$, we denote by X_t^k the observation made

[2]This is aligned with [6], [7] . In many applications, once a change has been detected, the operator can respond to it quickly and find out which type it is.

by sensor k at time t. We can now define $Q+1$ different hypotheses as follows. Under the hypothesis H_q, $q \in [Q]$, the random variables $X_1^k, X_2^k, \ldots, X_\nu^k$ are i.i.d. with the PDF P_0, while $X_{\nu+1}^k, X_{\nu+2}^k, \ldots$ are i.i.d. with the PDF P_q. Under the hypothesis H_0, X_t^k are i.i.d. with the PDF P_0 for all t. We write $\mathbf{X}_t = [X_t^1, \ldots, X_t^K]$ for each t and denote by $\mathbf{X}_{t_1}^{t_2}$ the collection of $\mathbf{X}_{t_1}, \mathbf{X}_{t_1+1}, \ldots, \mathbf{X}_{t_2}$ for each t_1, t_2 with $t_2 > t_1$.

As the binary case, there is a noiseless link of a finite or infinite number of bits associated with each sensor to the fusion center. At each time t, an honest sensor k makes a local decision individually by mapping its own observations up to time t to an element in $\mathcal{X}$ satisfying the bandwidth constraint. Depending on the adopted reporting mechanism and the bandwidth constraint, each sensor decides whether it should alarm the fusion center through the channel it is associated with. The M compromised sensors, on the other hand, try to disrupt/confuse the final decision of fusion center by sending attack signals which again belong to $\mathcal{X}$.

As [6], [7], let us define the sequence of alarm times $0 = T_0 < T_1 < T_2 < \ldots < T_\rho < \ldots$, where T_ρ is the alarm time using sensor observations after previous alarm time $T_{\rho-1}$, that is, $\mathbf{X}_{T_{\rho-1}+1}, \mathbf{X}_{T_{\rho-1}+2}, \ldots$; then the stopping time for type $q \in [Q]$ is defined as

$$T^q = \inf_{\rho \geq 1} \{T_\rho : \hat{q}_\rho = q\}, \tag{4}$$

where $\hat{q}_\rho$ is the decision at the fusion center declared at time T_ρ. Here, we use the convention $\inf\{\emptyset\} = \infty$ and it is possible that $T^q = \infty$, which corresponds to the case when the fusion center never declares change of type q. With a little abuse of notation, when only the first alarm time T_1 of a rule T matters, we sometimes simply write T_1 as T. Let g be an attack strategy of the M compromised sensors. We assume that the attacker knows ν, $\mathbf{X}_1^t$, and the global decision rule (including both the stopping rule at the fusion center and local rule at each sensor), and hence g is a function of these arguments. We also write $g = \emptyset$ when all the compromised sensors are absent. When a change under hypothesis H_q, $q \in [Q]$, happens at time ν and the strategy employed by the M compromised sensors is g, the underlying probability measure is denoted by $\mathsf{P}_\nu^{q,g}$. Moreover, when no change ever happens, i.e., $\nu = \infty$, we denote by $\mathsf{P}_\infty^{q=0,g}$ the underlying probability measure.

Following the single-sensor case [6], [7], we define the performance metrics as follows:

- **Detection Delay:** The worst-case mean detection delay is given by

$$\mathcal{D}[T] := \sup_{q \in [Q]} \sup_{g,\nu} \operatorname{ess\,sup} \mathbb{E}_\nu^{q,g}[(T-\nu)^+ | \mathbf{X}_1^\nu], \tag{5}$$

- **False Alarm or False Isolation:** The worst-case mean time to a false alarm or a false isolation is given by

$$\mathcal{A}[T] := \inf_{q\in[Q]^+} \inf_{g} \inf_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}_0^{q,g}[T^{\hat{q}}]. \tag{6}$$

Our objective is again to design fault-tolerant decision rules such that $\mathcal{D}[T]$ can be minimized, with large $\mathcal{A}[T] \geq \gamma$.

Remark II.2. *The detection delay defined in [6], [7] for the scenario with one single honest sensor is given by*

$$\sup_{q\in[Q]} \sup_{\nu} \operatorname{ess\,sup} \mathbb{E}_\nu^q[T-\nu | T>\nu, X_1^1, \ldots, X_\nu^1], \tag{7}$$

where $X_1^1, \ldots, X_\nu^1$ are the observations of the (single and honest) first sensor up to time ν. One may notice that (5) *is not quite of the same form as the above definition. In Lemma A.1 in Appendix A, we prove that even with multiple honest and compromised sensors, these two forms are equivalent and one is free to work with either of them.*

Before leaving this section, we quickly review the asymptotically optimal matrix CUSUM algorithm in [7] when there is just a single honest sensor $|\mathcal{N}| = 1, M = 0$, and $Q \geq 1$. For each hypothesis $q \in [Q]$, this honest sensor (with index $k = 1$) computes the CUSUM statistics $Y_t^k(q,j)$ for every $j \neq q \in [Q]^+$ at time t, recursively through $Y_0^k(q,j) = 0$ and

$$Y_t^k(q,j) = \left(Y_{t-1}^k(q,j) + \ell_t^k(q,j)\right)^+, \tag{8}$$

where $\ell_t^k(q,j) = \log \frac{P_q(X_t^k)}{P_j(X_t^k)}$ is the log-likelihood ratio (LLR) between P_q and P_j. The results are put into a $Q \times Q$ matrix $\mathbf{Y}_t$ with the qth row given by

$$\mathbf{Y}_t^k := [Y_t^k(q,0), \cdots, Y_t^k(q,j), \cdots, Y_t^k(q,Q)]. \tag{9}$$

Let $Y_{t,q}^k = \min_{j\in[Q]^+,\, j\neq q} Y_t^k(q,j)$ be the minimum of the qth row. The matrix CUSUM procedure in [7] locally determines that the event has occurred at the first time that any $Y_{t,q}^k, q \in [Q]$ exceeds a pre-defined threshold h. A hard decision is then alarmed, which means that the procedure terminates after this alarm and no other decisions will be further made.

III. Prior work

In this section, we review some prior results directly relevant to the present work. We again split our discussion into the binary and multi-hypothesis cases.

A. Binary BDQCD

For the considered binary BDQCD problem with $|\mathcal{N}| = 1$ and $M = 0$ (i.e., no compromised sensor and thereby no Byzantine attack), the problem reduces to the standard QCD problem for which it was shown in [4], [5] that Page's CUSUM procedure T_{single} [3] achieves the optimal scaling that for $\mathcal{A}[T_{\text{single}}] = \gamma$, the expected detection delay scales like $\mathcal{D}[T_{\text{single}}] \sim \log(\gamma)/I$ as $\gamma \to \infty$. For $M = 0$ and general $|\mathcal{N}|$, Mei in [19] developed a scheme $T_{\text{consensus}}$, called the consensus rule, where each sensor performs CUSUM according to its local observations and sends a binary report to the fusion center, which declares the occurrence of the event when all the $|\mathcal{N}|$ sensors simultaneously say so. It was then shown in [19] that this scheme is asymptotically optimal that under $\mathcal{A}[T_{\text{consensus}}] = \gamma$, the expected detection delay scales like

$$\mathcal{D}[T_{\text{consensus}}] \sim \frac{\log(\gamma)}{|\mathcal{N}|I}, \quad \text{as } \gamma \to \infty. \tag{10}$$

In [15], Banerjee and Fellouris proposed two families of stopping rules for the same $M = 0$ and general $|\mathcal{N}|$ case. In the first family of stopping rules, which we refer to as the one-shot d-th alarm, each sensor performs the CUSUM procedure locally and only reports an alarm once at the first time the local CUSUM statistic exceeds a predefined threshold; the fusion center then stops and declares the event as soon as receiving $d \leq |\mathcal{N}|$ reports. In the second family of stopping rules, which is referred to as the d-voting rule, each sensor again performs the CUSUM procedure locally but gets to report multiple times whenever the local CUSUM statistic exceeds the threshold. The fusion center then stops and declares the event as soon as receiving $d \leq |\mathcal{N}|$ reports *simultaneously*. The authors of [15] analyzed the second-order asymptotic performance and the results revealed that even though it was shown in [19] that the d-voting rule with $d = |\mathcal{N}|$ (i.e., the consensus rule) achieves the first-order asymptotic performance, it might be better in practice for it to wait for only the majority of sensors' reports, i.e., setting $d = \lceil (|\mathcal{N}| + 1)/2 \rceil$.

Very recently, in [13], binary BDQCD with general $|\mathcal{N}|$ and M was discussed and multiple schemes were analyzed. Among these schemes, the d-voting rule $\tau_{(d)}$, achieves the best scaling when $d = |\mathcal{N}|$ is chosen[3]. Specifically, it was shown in [13] that the following asymptotic performance can be achieved:

[3]This is also called the consensus rule in [13]. But it is noted that here, we only wait for $|\mathcal{N}|$, the number of honest sensors, responses rather than all K responses.

Theorem III.1 ([13, Theorem 26])**.** *Let d be an integer satisfying $M < d \leq |\mathcal{N}|$. For $\mathcal{A}[\tau_{(d)}] = \gamma$, the worst-case mean detection delay of the d-voting rule scales like*

$$\mathcal{D}[\tau_{(d)}] \sim \frac{\log \gamma}{(d-M)I}, \quad \text{as } \gamma \to \infty. \tag{11}$$

The best asymptotic performance reported in [13] is the above one with $d = |\mathcal{N}|$ (i.e., the consensus rule), which also coincides with another scheme in [13], low-sum-CUSUM that requires infinite bandwidth. This leads us to conjecture that (11), with $d = |\mathcal{N}|$, is the optimal first-order behavior. A tight converse is then necessary to verify this conjecture.

We would like to point out that a non-trivial converse can be obtained by revealing the identities of all $|\mathcal{N}|$ honest sensors and using the asymptotic optimality in [19], as detailed below.

Theorem III.2 (Simple converse)**.** *For any binary BCQCD rule T, with $\mathcal{A}[T] \geq \gamma$, the worst-case mean detection delay meets*

$$\mathcal{D}[T] \gtrsim \frac{\log \gamma}{|\mathcal{N}|I}, \quad \text{as } \gamma \to \infty. \tag{12}$$

Unfortunately, this converse is not tight compared to (11).

B. Multi-hypothesis BDQCD

To the best of our knowledge, the present work is the first to formulate and study the multi-hypothesis BDQCD. Prior to this work, the single honest sensor QCD problem with multiple hypothesis was first investigated in [6], in which Nikiforov extended Lorden's framework to include multiple post-change hypotheses. Nikiforov in [6] also proposed an algorithm based on the concept of generalized likelihood ratio and showed the asymptotic optimality of this algorithm. In [7], by cleverly switching the order of $\max$ and $\min$ in the algorithm in [6], Oskiper and Poor developed the matrix CUSUM algorithm that admits a recursive formula and hence can be efficiently implemented. Moreover, it was shown that, in addition to its low complexity, the matrix CUSUM procedure is also asymptotically optimal.

IV. Binary BDQCD

We consider the problem of binary DBQCD in this section. We first present the main result, that is, a tight converse to the first-order asymptotic performance of the worst-case detection delay, in Section IV-A. The proof of the main result is then given in Section IV-B.

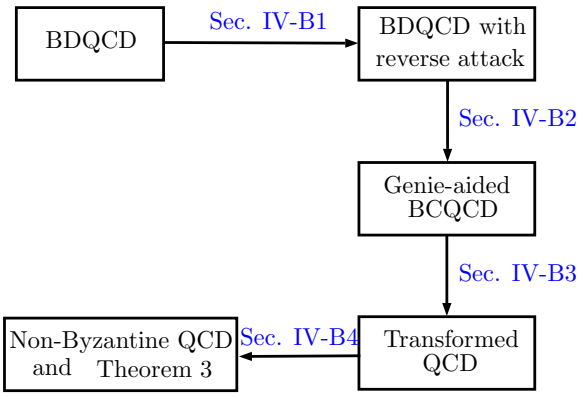


Fig. 1. Diagram of proof steps.

A. *Main results of this section*

Here, we present the main result of this section, which is a new converse of the first-order asymptotic performance of BDQCD.

Theorem IV.1 (Tight converse)**.** *For any binary BDQCD rule T, with $\mathcal{A}[T] \geq \gamma$, the worst-case mean detection delay is lower bounded as*

$$\mathcal{D}[T] \gtrsim \frac{\log \gamma}{(|\mathcal{N}| - M)I}, \quad \text{as } \gamma \to \infty. \tag{13}$$

A sketch of the proof of the new converse is outlined in Fig. 1 and the details are given in the next subsection. To prove this theorem, we first note that if the optimal asymptotic scaling is lower bounded by $\eta(\gamma)$ under an attack strategy, then it is also lower bounded by $\eta(\gamma)$ under the *worst* attack. We thus proceed by constructing an attack strategy in Section IV-B1, called the *reverse attack*, which is later shown to be an asymptotically worst attack. We then, in Section IV-B2, construct a genie providing the identities of $|\mathcal{N}| - M$ out of $|\mathcal{N}|$ honest sensors and the local observations used for generating the local report at every sensor. After that, by absorbing the impact of the reverse attack into pre/post-change distributions, the problem is transformed into an equivalent centralized QCD problem in Section IV-B3 for which CUSUM is known to be optimal. Finally, in Section IV-B4, evaluating the CUSUM procedure for the transformed problem reveals the connection to another non-Byzantine QCD with only $|\mathcal{N}| - M$ honest sensors.

When comparing the main result of this section presented above and the achieviability result

Theorem III.1, one immediately characterizes the optimal first-order behavior of binary BDQCD as follows.

Corollary IV.1. *For an optimal binary BDQCD rule T^*, subject to $\mathcal{A}[T^*] \geq \gamma$, the first-order asymptotic worst-case mean detection delay is given by*

$$\mathcal{D}[T^*] \sim \frac{\log \gamma}{(|\mathcal{N}| - M)I}, \quad \textit{as } \gamma \to \infty. \tag{14}$$

Remark IV.1. *Supposed that, as in [13], there is a subset $\mathcal{B} \subseteq \mathcal{N}$ such that only sensors in $\mathcal{B}$ are affected by the change and those in $\mathcal{N} \setminus \mathcal{B}$ have observations drawn i.i.d. according to P_0 all the time. We can slightly alter our genie in our proof so that it reveals the identities of $|\mathcal{B}| - M$ affected sensors and the $|\mathcal{N} \setminus \mathcal{B}|$ unaffected sensors. One can then follow the same technique to prove the following converse,*

$$\mathcal{D}[T] \gtrsim \frac{\log \gamma}{(|\mathcal{B}| - M)I}, \quad \textit{as } \gamma \to \infty. \tag{15}$$

Moreover, setting $d = |\mathcal{B}|$ in the d-voting rule achieves the above first-order scaling; hence, the optimal first-order asymptotic performance of this setting is also characterized as

$$\mathcal{D}[T^*] \sim \frac{\log \gamma}{(|\mathcal{B}| - M)I}, \quad \textit{as } \gamma \to \infty. \tag{16}$$

B. Proof of the converse for binary BDQCD

The proof presented in this section follows closely the steps shown in Fig. 1.

1) The reverse attack: For the ease of presentation in this proof, we define $P_{0,1} = P_0$ and $P_{1,1} = P_1$. Recall that each honest sensor k's observation sequence X_t^k is drawn i.i.d. according to $P_{0,1}$ before the change time ν and i.i.d. according to $P_{1,1}$ after ν. We construct an attack strategy as follows. For each compromised sensors k', it generates a fake observation sequence $X_t^{k'}$, which is then input to the assigned local decision function for forming the fake report. The fake observation sequence is generated i.i.d. according to $P_{0,2}$ and $P_{1,2}$ before and after the change time ν, respectively. That is, the compromised sensors form fake reports according to observations based on wrong distributions. To establish the tight converse, we will set $P_{0,2} = P_{1,1} = P_1$ and $P_{1,2} = P_{0,1} = P_0$ in the very end of the proof; therefore, we call this attack strategy the "reverse attack". However, most of the steps in the proof stay valid for general densities $P_{0,2}$ and $P_{1,2}$. Next, we will show that under this reverse attack, for any detection rule with mean time to a false alarm no less than γ, the mean detection delay is lower-bounded by the RHS of (13). Note that by definition, the worst case delay in (2) will also be lower-bounded by (13) automatically.

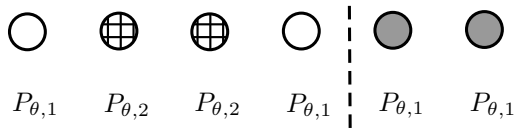
2) Genie-aided Byzantine centralized QCD: First note that the worst case happens when there are M compromised sensors. Also since the identities of the sensors are unknown, the fusion center cannot enhance the worst-case performance by selectively accepting reports. If the fusion center accepts reports from $K - K'$, $K' \leq |\mathcal{N}|$, sensors only, in the worst case, the problem reduces to the BDQCD with M compromised sensors and $|\mathcal{N}| - K'$ honest sensors, which results in a worse performance. Moreover, when $K' > |\mathcal{N}|$, we are left with only compromised sensors in the worst case, which is obviously worse than accepting all reports. We therefore only have to consider the fusion center taking reports from all K sensors for detection in what follows.

The K sensors are divided into three groups. Each of the first two groups consists of M sensors, while the last group contains $|\mathcal{N}| - M$ sensors. All sensors in the first and third groups are honest, while those in the second group are compromised. Assume that there is a genie giving away the identities of $|\mathcal{N}| - M$ honest sensors to the fusion center. For the rest M honest sensors and M compromised sensors, the identities are unknown to the fusion center. Without loss of generality, we assume that sensors in the first two groups have indices $[2M]$. We also give the observations used at each sensor (fake observations if the sensor is compromised) for generating its local report and the densities $P_{0,2}$ and $P_{1,2}$ to the fusion center. Let $s : [K] \to [2]$ be a function that assigns each sensor to index 1 or 2 (meaning "honest" or "compromised") in such a way that exactly M out of the first $2M$ sensors are assigned to index 2, and the last $|\mathcal{N}| - M$ sensors are all assigned to index 1. Let $\mathcal{S}$ be the collection of all possible assignments s. Clearly, there are total $|\mathcal{S}| = \binom{2M}{M}$ such assignments. For $\theta \in \{0, 1\}$, the product density under the compromised group assignment s is

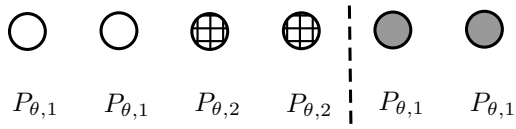
$$P_{\theta,s}(\mathbf{X}_t) = \prod_{k'=1}^{2M} P_{\theta,s(k')}(X_t^{k'}) \prod_{k=2M+1}^{K} P_{\theta,1}(X_t^k). \tag{17}$$

Now, we are facing a composite change detection problem, which we refer to as genie-aided Byzantine centralized QCD (BCQCD). Before the change time ν, the random vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_\nu$ are i.i.d. over time with density $P_{0,s}$ while $\mathbf{X}_{\nu+1}, \mathbf{X}_{\nu+1}, \ldots$ are generated with density $P_{1,s}$, for some $s \in \mathcal{S}$. In this genie-aided version, the fusion center knows everything about the compromised sensors except for their exact locations. With slight abuse of notations, as (2), the mean detection delay of this problem is given by

$$\mathcal{D}_{\text{genie}}[T] := \sup_{s\in\mathcal{S},\nu} \operatorname{ess\,sup} \mathbb{E}_\nu^s[(T-\nu)^+ | \mathbf{X}_1^\nu]; \tag{18}$$



a) First $2M$ sensors under assignment $\{1, 2, 2, 1\}$



b) First $2M$ sensors under assignment $\{1, 1, 2, 2\}$

Fig. 2. An example of genie-aided BCQCD with $|\mathcal{N}| = 4$ and $M = 2$ under two different assignments. Here, empty circles and crossed circles are honest and compromised sensors under the assignment s, respectively, and thus follow $P_{\theta,1}$ and $P_{\theta,2}$, respectively. Moreover, gray circles are those honest sensors whose identities are revealed to the fusion center by the genie; thereby, their observations follow $P_{\theta,1}$ always, regardless of assignment.

also as (3), the mean time to false alarm is

$$\mathcal{A}_{\text{genie}}[T] := \inf_{s\in\mathcal{S}} \mathbb{E}^s_\infty[T]. \tag{19}$$

Example IV.1. *An example of genie-aided BCQCD with $|\mathcal{N}| = 4$ honest sensors and $M = 2$ compromised sensors is provided in Fig. 2. In this figure, we use empty circles, crossed circles, and gray circles to represent honest sensors, compromised sensors, and honest sensors whose identities are revealed by the genie, respectively. The distribution that each sensor's observation follows under hypothesis $\theta \in \{0, 1\}$ is presented. In both Fig. 2-a) and Fig. 2-b), we note that the last $|\mathcal{N}| - M = 2$ sensors are always honest and their identities are revealed to the fusion center. Hence, their observations always follow P_θ independently. For the first $2M = 4$ sensors, the distributions under assignments $\{1, 2, 2, 1\}$ and $\{1, 1, 2, 2\}$ are shown in Fig. 2-a) and Fig. 2-b), respectively. We note that there are total $\binom{2M}{M} = 6$ different assignments and we only show 2 of them for demonstration.*

3) Transformed Centralized QCD: We transform the genie-aided BCQCD problem into an equivalent centralized QCD problem for which CUSUM is known to be optimal. Recall that unlike [14], now only $2M$ sensors' identities are unknown to the fusion center and we define "masked"-symmetric strategy as follows. Let $\tau_{2M}(\mathbf{X}_t)$ be the masked ordering map that puts the first $2M$ elements of its input $\mathbf{X}_t$ in descending order while keeps the other $|\mathcal{N}| - M$ positions unchanged. A decision rule $T(.)$ of the genie-aided BCQCD problem is said to be masked

symmetric if it can be represented as $T(\{\mathbf{X}_t\}_{t\geq 1}) = \tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})$ for some decision rule $\tilde{T}$.

In the transformed centralized QCD problem, the fusion center observes $\tilde{\mathbf{X}}_t = \tau_{2M}(\mathbf{X}_t)$ at time t. Let $\tilde{P}_\theta(\tilde{\mathbf{X}}_t)$ be the density of $\tau_{2M}(\mathbf{X}_t)$, where $\mathbf{X}_t$ is generated according to density $P_{\theta,s}, \theta \in \{0,1\}$. Before the change, the observations $\{\tilde{\mathbf{X}}_t\}$ follow $\tilde{P}_0$ while after the change, they follow $\tilde{P}_1$. Also,

$$\tilde{P}_\theta(\tilde{\mathbf{X}}_t) = \sum_{\mathbf{X}_t:\tau_{2M}(\mathbf{X}_t)=\tilde{\mathbf{X}}_t} P_{\theta,s}\left(\mathbf{X}_t\right), \tag{20}$$

where the equality follows from that the absolute value of the Jacobian of a permutation is always 1. Following the proof of part 1 of [14, Lemma 4.1], we can easily show that for all assignments $s \in \mathcal{S}$, the density $\tilde{P}_\theta(\tilde{\mathbf{X}}_t)$ does not depend on s. Suppose the change occurs at the time ν. Under hypothesis $\tilde{H}_1$, the random vectors $\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, \ldots, \tilde{\mathbf{X}}_\nu$ are drawn i.i.d. over time with density $\tilde{P}_0$ while $\tilde{\mathbf{X}}_{\nu+1}, \tilde{\mathbf{X}}_{\nu+2}, \ldots$ are generated i.i.d. with density $\tilde{P}_1$. Under hypothesis $\tilde{H}_0$, there is no change, i.e. $\nu = \infty$, and $\tilde{\mathbf{X}}_t$ are drawn i.i.d. with density $\tilde{P}_0$ for all t.

We first focus on a masked symmetric rule T, and show that the detection delay of the genie-aided BCQCD is identical to that of the transformed QCD problem, defined as $\mathcal{D}_{\text{trans}}[\tilde{T}] := \sup_\nu \operatorname{ess\,sup} \mathbb{E}_\nu[(\tilde{T}-\nu)^+|\tilde{\mathbf{X}}_1^\nu]$. Compared to the transformation in the one-shot hypothesis testing [14], our delay in (18) involves all pre-change observations by taking an essential supreme over the distributions of them. This difference complicates the transformation. Specifically, we will show

$$\begin{aligned}\mathcal{D}_{\text{genie}}[T] &= \sup_{s\in\mathcal{S},\nu} \operatorname{ess\,sup} \mathbb{E}_\nu^s[(\tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1}) - \nu)^+|\mathbf{X}_1^\nu] \\ &\overset{(a)}{=} \sup_{\nu} \operatorname{ess\,sup} \mathbb{E}_\nu[(\tilde{T}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1}) - \nu)^+|\tilde{\mathbf{X}}_1^\nu] = \mathcal{D}_{\text{trans}}[\tilde{T}].\end{aligned} \tag{21}$$

The first equality is from the definition in (18) and we will devote ourselves to proving equality

(21a). To this end, note that

$$\begin{aligned}
&\mathbb{E}^s_\nu[(\tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu]\\
=&\sum_{z=0}^{\infty} 1-\mathbb{P}\left(\left(\tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})-\nu\right)^+\leq z\big|\mathbf{X}_1^\nu\right)\\
=&\sum_{z=0}^{\infty} 1-\int 1_{\left\{\left(\tilde{T}(\{\tau_{2M}(\mathbf{x}_t)\}_{t\geq 1})-\nu\right)^+\leq z\right\}}\prod_{t=\nu+1}^{\nu+z} P_{1,s}(\mathbf{x}_t)d\mathbf{x}_{\nu+1}^{\nu+z}\\
\overset{(a)}{=}&\sum_{z=0}^{\infty} 1-\int 1_{\left\{\left(\tilde{T}(\{\tilde{\mathbf{x}}_t\}_{t\geq 1})-\nu\right)^+\leq z\right\}}\prod_{t=\nu+1}^{\nu+z} \tilde{P}_1(\tilde{\mathbf{x}}_t)d\tilde{\mathbf{x}}_{\nu+1}^{\nu+z}\\
=&\sum_{z=0}^{\infty} 1-\mathbb{P}\left(\left(\tilde{T}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1})-\nu\right)^+\leq z\big|\tilde{\mathbf{X}}_1^\nu\right)\\
=&\mathbb{E}_\nu[(\tilde{T}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1})-\nu)^+|\tilde{\mathbf{X}}_1^\nu],
\end{aligned}\tag{22}$$

where $\mathbb{P}$ is the associated probability measure and $1_{\{.\}}$ is the indicator function; and (a) follows from the change of variables in integration [24] and the fact that $\tilde{P}_1(.)$ does not depend on s.

Now, for a fixed ν and for each $s\in\mathcal{S}$, let $\mathbb{P}^s$ and $\tilde{\mathbb{P}}$ denote the probability measures on $\mathbb{R}^{K\times\nu}$ with densities specified by (17) and (20) with $\theta=0$, respectively. To establish (21a), observe that for any $x\in\mathbb{R}^{K\times\nu}$, from (22), we have

$$\begin{aligned}
&\mathbb{E}^s_\nu[(\tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu](x)=\\
&\mathbb{E}_\nu[(\tilde{T}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1})-\nu)^+|\tilde{\mathbf{X}}_1^\nu](\tau_{2M}(x)).
\end{aligned}\tag{23}$$

Let D_M denote $\operatorname{ess\,sup}\mathbb{E}^s_\nu[(\tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu]$ where the essential supremum is taken under $\mathbb{P}^s$. By definition, there exists $\Omega\subseteq\mathbb{R}^{K\times\nu}$ with $\mathbb{P}^s(\Omega)=1$ such that

$$D_M\geq\mathbb{E}^s_\nu[(\tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu](x)$$

for all $x\in\Omega$. By (23), we have

$$D_M\geq\mathbb{E}_\nu[(\tilde{T}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1})-\nu)^+|\tilde{\mathbf{X}}_1^\nu](\tau_{2M}(x)),\quad\forall x\in\Omega.$$

Note that

$$\tilde{\mathbb{P}}(\tau_{2M}(\Omega))=\int_{\tau_{2M}(\Omega)}\tilde{P}_0(y)dy=\int_\Omega P_{0,s}(x)dx=\mathbb{P}^s(\Omega)=1,$$

where the densities $\tilde{P}_0$ and $P_{0,s}$ are given by (20) and (17) respectively. We therefore conclude that $D_M\geq\operatorname{ess\,sup}\mathbb{E}_\nu[(\tilde{T}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1})-\nu)^+|\tilde{\mathbf{X}}_1^\nu]$ where the essential supremum here is taken under

$\tilde{\mathbb{P}}$. Noting that this is true for every $s \in \mathcal{S}$ gives the relation "$\geq$" in (21a). By using the same argument as above, but switching the roles of the left-hand side and right-hand side of (21a), we obtain the relation "$\leq$".

We have shown that the detection delay of genie-aided BCQCD (18) is equal to that of transformed QCD under masked symmetric rules. One can similarly prove that the mean time to false alarm $\mathcal{A}_{\text{genie}}[T]$ in (19) is equal to that of the new problem $\mathcal{A}_{\text{trans}}[\tilde{T}] := \mathbb{E}_{\infty}[\tilde{T}]$. The rest is to show that for any fusion rule $T'(.)$, there is a masked symmetric rule $T(.)$ that is not worse than $T'(.)$. This is shown in Lemma A.2 in Appendix and then the transformation of QCD is established.

4) Establishing the converse: Let T^*, T^*_{genie}, and T^*_{trans} be optimal stopping rules for BDQCD, genie-aided BCQCD, and transformed QCD, respectively. So far, we have established the following relationships among the aforementioned problems

$$\mathcal{D}[T^*] \geq \mathcal{D}_{\text{genie}}[T^*_{\text{genie}}] = \mathcal{D}_{\text{trans}}[T^*_{\text{trans}}], \tag{24}$$

and

$$\mathcal{A}[T^*] \leq \mathcal{A}_{\text{genie}}[T^*_{\text{genie}}] = \mathcal{A}_{\text{trans}}[T^*_{\text{trans}}]. \tag{25}$$

We can therefore establish a converse bound by evaluating the performance of the transformed QCD, which is a standard QCD problem with observations following $\tilde{P}_0$ and $\tilde{P}_1$ before and after the change point ν, respectively. For such the problem, it is well known from [5], [19, Lemma 2] that an optimal strategy is Page's CUSUM procedure given by $\tilde{\sigma}(h) = \inf\{t \in \mathbb{N} : \tilde{Y}_t \geq h\}$, where $\tilde{Y}_t = (\tilde{Y}_{t-1} + \tilde{\ell}_t)^+$ with $\tilde{Y}_0 = 0$, and from (20)

$$\begin{aligned}
\tilde{\ell}_t &= \log \frac{\sum_{\mathbf{X}_t:\tau_{2M}(\mathbf{X}_t)=\tilde{\mathbf{X}}_t} P_{1,s}(\mathbf{X}_t)}{\sum_{\mathbf{X}_t:\tau_{2M}(\mathbf{X}_t)=\tilde{\mathbf{X}}_t} P_{0,s}(\mathbf{X}_t)} \\
&= \log \frac{\sum_{\pi\in\Pi_{2M}} P_{1,s}(\pi(\tilde{\mathbf{X}}_t))}{\sum_{\pi\in\Pi_{2M}} P_{0,s}(\pi(\tilde{\mathbf{X}}_t))} = \log \frac{\sum_{\pi\in\Pi_{2M}} P_{1,s\circ\pi^{-1}}(\tilde{\mathbf{X}}_t)}{\sum_{\pi\in\Pi_{2M}} P_{0,s\circ\pi^{-1}}(\tilde{\mathbf{X}}_t)} \\
&\overset{(a)}{=} \log \frac{\sum_{s'\in\mathcal{S}} P_{1,s'}(\tilde{\mathbf{X}}_t) \frac{2M!}{\binom{2M}{M}}}{\sum_{s'\in\mathcal{S}} P_{0,s'}(\tilde{\mathbf{X}}_t) \frac{2M!}{\binom{2M}{M}}}.
\end{aligned} \tag{26}$$

where $\pi : [K] \to [K]$ is a masked permutation function that permutes the first $2M$ entries while keeps the remaining $|\mathcal{N}| - M$ entries unchanged, Π_{2M} is the collection of all ($2M!$ in total) such π, and $\circ$ is the function composition operator; (a) follows from the fact that for a compromised group assignment s, summing over all the permuted versions $s \circ \pi^{-1}$ is equivalent to summing

over all the assignments s' with each s' being involved $2M!/\binom{2M}{M}$ times. In what follows, we set $P_{0,2} = P_{1,1}$ and $P_{1,2} = P_{0,1}$ according to the reverse attack described in Sec. IV-B1. We now rewrite the likelihood in (26) as

$$\tilde{\ell}_t = \log \frac{\sum_{s\in\mathcal{S}} P_{1,s}(\tilde{\mathbf{X}}_t)}{\sum_{s\in\mathcal{S}} P_{0,s}(\tilde{\mathbf{X}}_t)} \tag{27}$$

$$\overset{(a)}{=} \log \frac{\left(\sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{1,s(k')}(\tilde{X}_t^{k'})\right) \prod_{k=2M+1}^{K} P_{1,1}(\tilde{X}_t^k)}{\left(\sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{0,s(k')}(\tilde{X}_t^{k'})\right) \prod_{k=2M+1}^{K} P_{0,1}(\tilde{X}_t^k)}$$

$$\overset{(b)}{=} \log \frac{\prod_{k=2M+1}^{K} P_{1,1}(\tilde{X}_t^k)}{\prod_{k=2M+1}^{K} P_{0,1}(\tilde{X}_t^k)}, \tag{28}$$

where (a) follows from (17) and (b) is because of the fact that for every s, there exists a $\bar{s}$ such that $\bar{s}(k') = 2$ whenever $s(k') = 1$ and $\bar{s}(k') = 1$ whenever $s(k') = 2$; therefore,

$$\begin{aligned}\sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{1,s(k')}(\tilde{X}_t^{k'}) &= \sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{0,\bar{s}(k')}(\tilde{X}_t^{k'}) \\ &= \sum_{\bar{s}\in\mathcal{S}} \prod_{k'=1}^{2M} P_{0,\bar{s}(k')}(\tilde{X}_t^{k'}),\end{aligned} \tag{29}$$

where the first equality is from $P_{1,1} = P_{0,2}$ and $P_{1,2} = P_{0,1}$. Note that when $k = 2M+1 \ldots K$, $\tilde{X}_t^k$ is equal to the honest observation X_t^k in the BCQCD before transformation. Hence, the optimal test reduces to the standard centralized CUSUM procedure for the change detection with $|\mathcal{N}| - M$ honest sensors. Applying the results in [19] then shows (13).

5) Discussions: A byproduct obtained along the proof is that the "reverse attack" proposed in Section IV-B1 is an asymptotically worst attack for the original BDQCD problem. This observation is further exploited using the game theory, as described in Lemma VI.1 and Theorem VI.1 of Section VI.

Note that although the idea of the transformation in Sec. IV-B3 is inspired by [14], our proofs presented above are quite different. In [14], all sensors' identities are not revealed to the fusion center, while it is essential for us to construct a strong, but not so powerful, genie that reveals identities of some sensors for proving a tight converse, as in Sec. IV-B2. This new genie is the key to validate (28) in Sec. IV-B4, which shows that the optimal test under the proposed reverse attack constructed in Sec. IV-B1 only relates to the revealed $|\mathcal{N}|-M$ honest sensors. Furthermore, the transformation in Sec. IV-B3 is more involved than that in [14]. The difficulty comes from the fundamental difference between the one-shot hypothesis testing problem in [14] and our

sequential change detection that deals with observation sequences with an unknown change time. Finally, in the coming Section V-B, we extend our converse to the multiple-hypothesis case and face new challenges compared with the binary hypothesis problem in [14].

V. Multi-Hypothesis BDQCD

In this section, we consider the multi-hypothesis BDQCD. Again, we first present our main result of this section in Section V-A, which is the characterization of the asymptotic performance of the worst-case detection delay subject to a mean time to a false alarm or a false isolation. We then prove a converse for the considered problem in Section V-B, followed by the proposed stopping rules and their performance analysis in Section V-C. Throughout the section, we define for each pair of $q, j \in [Q]^+$, $q \neq j$, the KL divergence from P_j to P_q as

$$I(q,j) := \int \log\left(P_q(x)/P_j(x)\right) P_q(x) dx. \tag{30}$$

Let $\sigma^2(q,j)$ be the second moment of $I(q,j)$ defined as

$$\sigma^2(q,j) := \mathbb{E}_q\left[\left(\log\left(\frac{P_q(x)}{P_j(x)}\right) - I(q,j)\right)^2\right]. \tag{31}$$

We then make the following assumption:

Assumption V.1. *For any $q \in [Q]$,*

(i) $0 < I(q,j) < \infty$ *and* $\sigma^2(q,j) < \infty$, $\forall j \in [Q]^+, j \neq q$.

(ii) *Let $I^q := \min_{0\le j\le Q,\ j\neq q} I(q,j)$. Assume I^q admits a unique minimizer $j_q^* \in [Q]^+ \setminus \{q\}$.*

Now, consider $q = 0$.

(iii) *We define $I^0 := \min_{1\le j\le Q} I(j,0)$, and assume that I^0 admits a unique minimizer $j_0^* \in [Q]$.*

A. Main results of this section

Our first result for multi-hypothesis BDQCD with $Q+1$ hypotheses is the characterization of the converse as follows.

Theorem V.1. *Consider the multi-hypothesis BDQCD with $Q+1$ hypotheses. For any rule T with $\mathcal{A}[T] \ge \gamma$, the worst-case mean detection delay is lower bounded as*

$$\mathcal{D}[T] \gtrsim \frac{\log\gamma}{(|\mathcal{N}| - M)I^*}, \quad \textit{as } \gamma \to \infty. \tag{32}$$

where

$$I^* = \min_q I^q = \min_{q\in[Q]} \min_{j\in[Q]^+\setminus\{q\}} I(q,j) \tag{33}$$

This converse can be proved in a similar way to Theorem IV.1; hence, we only list the major differences between the two proofs in Section V-B.

For the achievability, we propose in Section V-C a family of stopping rules, called the simultaneous d-th alarm $\tau^s_{(d)}$, and show that this stopping rule achieves the first-order scaling of (32) when d is set to be $|\mathcal{N}|$. This simultaneous d-th alarm rule requires each sensor to send a Q-bit signal constantly through the noiseless link to the fusion center. To reduce the demanding bandwidth and energy requirements, another family of stopping rules, called multi-shot d-th alarm $\tau^m_{(d)}$, is proposed in Section V-C. This rule is more economic in that it only requires the sensor sending $\lceil \log_2 Q \rceil$-bit signal occasionally. In what follows, we present the asymptotic performance of the two proposed families of rules and refer the reader to Section V-C for their proofs.

Theorem V.2. *With the wort-case mean time to a false alarm or isolation no smaller than than γ, we have*
(a) among the proposed simultaneous d-th alarm rule $\tau^s_{(d)}$, the best first-order asymptotic worst-case mean detection delay is achieved when $d = |\mathcal{N}|$ and is given by

$$\mathcal{D}[\tau^s_{(|\mathcal{N}|)}] \lesssim \frac{\log\gamma}{(|\mathcal{N}|-M)I^*}, \quad \text{as } \gamma\to\infty; \tag{34}$$

(b) among the proposed multi-shot d-th alarm rule $\tau^m_{(d)}$, the first-order asymptotic worst-case mean detection delay when $d \geq M+1$ is given by

$$\mathcal{D}[\tau^m_{(d)}] \lesssim \frac{\log\gamma}{I^*}, \quad \text{as } \gamma\to\infty. \tag{35}$$

Combining the results in Theorems V.1 and V.2, we arrive at the following result of the optimal scaling for multi-hypothesis BDQCD with $Q+1$ hypotheses.

Corollary V.1. *For multi-hypothesis BDQCD with $Q+1$ hypotheses and the number of honest sensors $|\mathcal{N}| \geq M+1$, if the noiseless link of each sensor can support (at least) Q bits, the first-order asymptotic worst-case mean detection delay of an optimal stopping rule T^* subject to $\mathcal{A}[T^*] \geq \gamma$ is precisely*

$$\mathcal{D}[T^*] \sim \frac{\log\gamma}{(|\mathcal{N}|-M)I^*}, \quad \text{as } \gamma\to\infty. \tag{36}$$

Moreover, if $|\mathcal{N}| = M + 1$*, the optimal scaling*

$$\mathcal{D}[T^*] \sim \frac{\log \gamma}{I^*}, \quad \textit{as } \gamma \to \infty, \tag{37}$$

can be achieved by a stopping rule that requires only $\lceil \log_2 Q \rceil$*-bit links.*

B. Proof of Theorem V.1, the converse for multi-hypothesis BDQCD in Theorem V.2

For extending the converse from the binary case to the multi-hypothesis case with $Q + 1$ hypotheses, we encounter two main challenges. First, the asymptotically worst attack adopted in the binary case, namely the reverse attack, cannot be straightforwardly applied. As there are Q post-change distributions, we have to carefully choose one of them for swapping in order to make the attack strategy asymptotically worst. Second, after we manage to construct the attack strategy and complete the transformation, there are $Q+1$ hypotheses in the transformed QCD and hence Q^2 LLRs (see (8)) to be tracked in the asymptotically optimal matrix CUSUM procedure. It is difficult to make each LLR relate to only observations of (a subset of) honest sensors as we have done in (28) for the binary case.

In what follows, to solve the first issue, we modify the reverse attack in upcoming Section V-B1 by swapping the P_0 with the post-change distribution that is closest to P_0 in the sense of having the minimum KL divergence. To circumvent the second issue, we abandon the approach of evaluating the optimal detecting procedure in Section IV-B4 and directly perform the delay analysis based on [6, Theorem 2].

1) The reverse attack for $(Q+1)$*-hypotheses:* To prove this converse, under the true hypothesis $H_q, q \in [Q]^+$ (defined in Section II-B), we define $P_{q,1} = P_q$. The distribution $P_{q,2}$ for fake observation in the proposed attack strategy is constructed as follows. Let $q_m = \arg_{q\in[Q]^+} \min I_q$, with the corresponding $j^*_{q_m}$ defined in Assumption V.1 (ii) and (iii). We define

$$P_{q,2} = \begin{cases} P_{j^*_q}, & \text{if } q \neq j^*_{q_m}, \\ P_{q_m}, & \text{if } q = j^*_{q_m}. \end{cases} \tag{38}$$

The main intuition behind this choice is that we want the compromised sensors to follow the distribution that is closest to the true one under the KL divergence. This is done in the first case above. For the exception in the second case, i.e., $q = j^*_{q_m}$, it is for having the symmetry $P_{j^*_{q_m},2} = P_{q_m}$ and $P_{q_m,2} = P_{j^*_{q_m}}$, which will become handy later in the proof.

2) Genie-aided Byzantine centralized $(Q+1)$-hypotheses QCD: To establish the tight converse, we assume that a genie gives the fusion center the identities of $|\mathcal{N}|-M$ out of $|\mathcal{N}|$ honest sensors. Without loss of generality, we let these $|\mathcal{N}|-M$ sensors have the indices $2M+1,\ldots,K$, respectively. The genie also provides the fusion center with the observations used at each sensor for generating its local reports. Similarly to (17) in the binary case, with the help of this genie, the problem becomes the genie-aided BCQCD with distribution for each $q\in[Q]^+$ as

$$P_{q,s}(\mathbf{X}_t)=\prod_{k'=1}^{2M}P_{q,s(k')}(X_t^{k'})\prod_{2M+1}^{K}P_{q,1}(X_t^k). \tag{39}$$

For this genie-aided BCQCD, replacing $g\in\mathcal{G}$ with $s\in\mathcal{S}$ in (5) and (6), we obtain the detection delay and mean time to a false alarm or a false isolation given by

$$\mathcal{D}_{\text{genie}}[T]=\sup_{q\in[Q]}\sup_{s\in\mathcal{S},\nu}\operatorname{ess\,sup}\mathbb{E}_\nu^{q,s}[(T-\nu)^+|\mathbf{X}_1^\nu], \tag{40}$$

and

$$\mathcal{A}_{\text{genie}}[T]=\inf_{q\in[Q]^+}\inf_{s\in\mathcal{S}}\inf_{\hat{q}\in[Q]\setminus\{q\}}\mathbb{E}_0^{q,s}[T^{\hat{q}}], \tag{41}$$

respectively.

3) Transformed Centralized $(Q+1)$-hypotheses QCD: We can now follow the proof of the binary case to transform the genie-aided BCQCD into a multiple-hypothesis QCD having distributions generalizing (20) as

$$\tilde{P}_q(\tilde{\mathbf{X}}_t)=\sum_{\mathbf{X}_t:\tau_{2M}(\mathbf{X}_t)=\tilde{\mathbf{X}}_t}P_{q,s}(\mathbf{X}_t), \tag{42}$$

where $q\in[Q]^+$ and τ_{2M} is again the masked ordering map. Let us define the KL divergence between transformed distributions $\tilde{P}_q$ and $\tilde{P}_j$ for $q\in[Q]$ and $j\in[Q]^+$ as

$$\tilde{I}(q,j)=\int\log\left(\frac{\tilde{P}_q(\tilde{\mathbf{x}})}{\tilde{P}_j(\tilde{\mathbf{x}})}\right)\tilde{P}_q(\tilde{\mathbf{x}})d\tilde{\mathbf{x}}. \tag{43}$$

Moreover, let $\tilde{I}_q^*=\min_{j\in[Q]^+\setminus\{q\}}\tilde{I}(q,j)$ and

$$\tilde{I}^*=\min_{q\in[Q]}\tilde{I}_q^*=\min_{q\in[Q]}\min_{j\in[Q]^+\setminus\{q\}}\tilde{I}(q,j). \tag{44}$$

Note that in the binary case (24), we have shown that the transformed QCD would have the same detection delay with the genie-aided BCQCD for any post-change distribution P_1.

Extending from the binary case (18) to the multi-hypothesis case (40), we just need to take an additional supremum over all post-change distributions P_q, $q \in [Q]$; therefore, the equivalence

$$\begin{aligned}\mathcal{D}_{\text{genie}}[T^*_{\text{genie}}] &= \mathcal{D}_{\text{trans}}[T^*_{\text{trans}}] \\ &:= \sup_{q\in[Q]} \sup_{\nu} \operatorname{ess\,sup} \mathbb{E}^q_\nu \left[\left(T^*_{\text{trans}}(\{\tilde{\mathbf{X}}_t\}_{t\geq 1}) - \nu \right)^+ \middle| \tilde{\mathbf{X}}^\nu_1 \right], \end{aligned} \tag{45}$$

still holds under multiple hypotheses for optimal rules T^*_{genie} and T^*_{trans} in genine-aided BCQCD and transformed QCD, respectively.

Regarding the mean time to a false alarm or false isolation, it is a bit more involved than the proof for detection delay since in addition to false alarm considered in Section IV-B3, we also need to deal with false isolation. We again first focus on a masked symmetric rule T satisfying $T(\{\mathbf{X}_t\}_{t\geq 1}) = \tilde{T}(\{\tau_{2M}(\mathbf{X}_t)\}_{t\geq 1})$ for some rule $\tilde{T}$. We aim to prove

$$\mathcal{A}_{\text{genie}}[T] = \mathcal{A}_{\text{trans}}[\tilde{T}] \tag{46}$$

Consider $\mathcal{A}_{\text{genie}}[T]$ in (41), for each $q \in [Q]^+$ and $s \in \mathcal{S}$, we have

$$\begin{aligned}\inf_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}^{q,s}_0[T^{\hat{q}}] &= \inf_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}^{q,s}_0 \left[\inf_{\rho\geq 1} \frac{T_\rho}{1_{\{\hat{q}_\rho = \hat{q}\}}} \right] \\ &= \inf_{\hat{q}\in[Q]\setminus\{q\}} \sum_{z=0}^{\infty} 1 - \mathbb{P}\left(\inf_{\rho\geq 1} \frac{T_\rho}{1_{\{\hat{q}_\rho=\hat{q}\}}} \leq z \right), \end{aligned} \tag{47}$$

where the first equality is from (4) and the convention $a/0 := \infty$ for positive $a \in \mathbb{R}$. Observe that

$$\begin{aligned}\mathbb{P}\left(\inf_{\rho\geq 1} \frac{T_\rho}{1_{\{\hat{q}_\rho=\hat{q}\}}} \leq z \right) &= \int 1_{\left\{ \inf_{\rho\geq 1} \frac{T_\rho(\{\mathbf{x}_t\}_{t\geq 1})}{1_{\{\hat{q}_\rho=\hat{q}\}}} \leq z \right\}} \prod_{t=1}^{z} P_{q,s}(\mathbf{x}_t) d\mathbf{x}_1^z \\ &= \int 1_{\left\{ \inf_{\rho\geq 1} \frac{\tilde{T}_\rho(\{\tau_{2M}(\mathbf{x}_t)\}_{t\geq 1})}{1_{\{\hat{q}_\rho=\hat{q}\}}} \leq z \right\}} \prod_{t=1}^{z} P_{q,s}(\mathbf{x}_t) d\mathbf{x}_1^z. \end{aligned} \tag{48}$$

Plugging (48) into (47), then (47) equals to

$$\begin{aligned}&\inf_{\hat{q}\in[Q]\setminus\{q\}} \sum_{z=0}^{\infty} 1 - \int 1_{\left\{ \inf_{\rho\geq 1} \frac{\tilde{T}_\rho(\{\tilde{\mathbf{x}}_t\}_{t\geq 1})}{1_{\{\hat{q}_\rho=\hat{q}\}}} \leq z \right\}} \prod_{t=1}^{z} \tilde{P}_{q,s}(\tilde{\mathbf{x}}_t) d\tilde{\mathbf{x}}_1^z \\ &= \inf_{\hat{q}\in[Q]\setminus\{q\}} \sum_{z=0}^{\infty} 1 - \mathbb{P}\left(\inf_{\rho\geq 1} \frac{\tilde{T}_\rho}{1_{\{\hat{q}_\rho=\hat{q}\}}} \leq z \right) \\ &= \inf_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}^{q,s}_0 \left[\inf_{\rho\geq 1} \frac{\tilde{T}_\rho}{1_{\{\hat{q}_\rho=\hat{q}\}}} \right] \\ &= \inf_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}^{q,s}_0[\tilde{T}^{\hat{q}}]. \end{aligned} \tag{49}$$

Taking infimum over $q \in [Q]^+$ and $s \in \mathcal{S}$ on the above shows (46) for a masked symmetric rule T. In Lemma A.3, we show that it suffices to consider masked symmetric rules as for any general rule T', there exists a symmetrized rule which is not worse than it in $\mathcal{A}_{\text{genie}}[T']$.

We have extended (25) and again shown that $\mathcal{A}_{\text{genie}}[T^*_{\text{genie}}] = \mathcal{A}_{\text{genie}}[T^*_{\text{trans}}]$ for optimal rules T^*_{genie} and T^*_{trans} in the multi-hypothesis cases and completed the transformation. To establish the converse, we now provide a converse to the asymptotic performance of the transformed QCD in Lemma A.4. That is, the first-order scaling of an optimal stopping rule T^*_{QCD} with $\mathcal{A}_{\text{trans}}[T^*_{\text{QCD}}] \geq \gamma$ is given by

$$\mathcal{D}_{\text{trans}}[T^*_{\text{QCD}}] \sim \frac{\log \gamma}{\tilde{I}^*}. \tag{50}$$

4) Establishing the converse by evaluating the transformed delay: To establish the converse, we now look into the structure of $\tilde{I}^*$ in (50), which is defined in (44). First, we note from (42)-(43) that for any pair of hypothesis indexes (q, j)

$$\begin{aligned}
\tilde{I}(q,j) &= \int \log\left(\frac{\sum_{\mathbf{x}:\tau_{2M}(\mathbf{x})=\tilde{\mathbf{x}}} P_{q,s}(\mathbf{x})}{\sum_{\mathbf{x}:\tau_{2M}(\mathbf{x})=\tilde{\mathbf{x}}} P_{j,s}(\mathbf{x})}\right) \tilde{P}_q(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \\
&\overset{(a)}{=} \int \log\left(\frac{\sum_{s\in\mathcal{S}} P_{q,s}(\tilde{\mathbf{x}}) \frac{2M!}{\binom{2M}{M}}}{\sum_{s\in\mathcal{S}} P_{j,s}(\tilde{\mathbf{x}}) \frac{2M!}{\binom{2M}{M}}}\right) \tilde{P}_q(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \\
&\overset{(b)}{=} \int \log\left(\frac{\sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{q,s(k')}(\tilde{x}^{k'}) \frac{2M!}{\binom{2M}{M}}}{\sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{j,s(k')}(\tilde{x}^{k'}) \frac{2M!}{\binom{2M}{M}}}\right) \tilde{P}_q(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \\
&\quad + \int \log\left(\frac{\prod_{2M+1}^{K} P_{q,1}(\tilde{x}^k)}{\prod_{2M+1}^{K} P_{j,1}(\tilde{x}^k)}\right) \tilde{P}_q(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \\
&\overset{(c)}{=} \int \log\left(\frac{\hat{P}_{q,s}(\tilde{x}^1, \ldots, \tilde{x}^{2M})}{\hat{P}_{j,s}(\tilde{x}^1, \ldots, \tilde{x}^{2M})}\right) \hat{P}_{q,s}(\tilde{x}^1, \ldots, \tilde{x}^{2M}) d\tilde{x}^1 \ldots d\tilde{x}^{2M} \\
&\quad + \int \log\left(\frac{\prod_{2M+1}^{K} P_{q,1}(\tilde{x}^k)}{\prod_{2M+1}^{K} P_{j,1}(\tilde{x}^k)}\right) \tilde{P}_q(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \\
&\overset{(d)}{\geq} \int \log\left(\frac{\prod_{2M+1}^{K} P_{q,1}(\tilde{x}^k)}{\prod_{2M+1}^{K} P_{j,1}(\tilde{x}^k)}\right) \prod_{2M+1}^{K} P_{q,1}(\tilde{x}^k) d\tilde{x}^{2M+1} \ldots d\tilde{x}^K \\
&= \sum_{k=2M+1}^{K} \int \log\left(\frac{P_{q,1}(x^k)}{P_{j,1}(x^k)}\right) P_{q,1}(x^k) dx^k \\
&\overset{(e)}{=} (|\mathcal{N}| - M) I(q,j),
\end{aligned} \tag{51}$$

where (a) follows from the same steps reaching (26), that is,

$$\sum_{\mathbf{x}:\tau_{2M}(\mathbf{x})=\tilde{\mathbf{x}}} P_{q,s}(\mathbf{x}) = \sum_{s\in\mathcal{S}} P_{q,s}(\tilde{\mathbf{x}}) \frac{2M!}{\binom{2M}{M}}; \tag{52}$$

(b) is because of the independence in (39); (c) holds by $\tilde{P}_q(\tilde{\mathbf{x}})$ equals to (52) and marginalizing $\tilde{x}^{2M+1}, \ldots, \tilde{x}^K$ out for the first integration, where we define a new PDF

$$\hat{P}_{q,s}(\tilde{x}^1, \ldots, \tilde{x}^{2M}) = \sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{q,s(k')}(\tilde{x}^{k'}) \frac{2M!}{\binom{2M}{M}}; \tag{53}$$

(d) follows by noting that the first integration in (c) is the KL divergence between $\hat{P}_{q,s}$ and $\hat{P}_{j,s}$, which is always non-negative [18], together with marginalizing $\tilde{x}^1, \ldots, \tilde{x}^{2M}$ in the second integration and the fact that the sensors in the third group are always honest in (42) (see (39)); and (e) follows from the selection of reverse attack in Section V-B1 and definition (30). Now, recall $q_m = \arg_q \min I_q$. For the pair $(q_m, j^*_{q_m})$ defined in Section V-B1, besides the inequality in (51), we can further show the following equality

$$\begin{aligned}
\tilde{I}(q_m, j^*_{q_m}) &= \int \log\left(\frac{\sum_{s\in\mathcal{S}} P_{q_m,s}(\tilde{\mathbf{x}})}{\sum_{s\in\mathcal{S}} P_{j^*_{q_m},s}(\tilde{\mathbf{x}})}\right) \tilde{P}_{q_m}(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \\
&\overset{(a)}{=} \int \log\left(\frac{\prod_{2M+1}^{K} P_{q_m,1}(\tilde{x}^k)}{\prod_{2M+1}^{K} P_{j^*_{q_m},1}(\tilde{x}^k)}\right) \prod_{2M+1}^{K} P_{q,1}(\tilde{x}^k) d\tilde{x}^{2M+1} \ldots d\tilde{x}^K \\
&= \sum_{k=2M+1}^{K} \int \log\left(\frac{P_{q_m,1}(x^k)}{P_{j^*_{q_m},1}(x^k)}\right) P_{q_m,1}(x^k) dx^k \\
&= (|\mathcal{N}| - M) I(q_m, j^*_{q_m}) = (|\mathcal{N}| - M) \min_{q\in[Q]} I_q,
\end{aligned} \tag{54}$$

where equality (a) follows from the symmetry enforced in the second case of the reverse attack in (38) and the steps for reaching (28). More specifically, in (51 b)

$$\sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{q_m,s(k')}(\tilde{x}^{k'}) = \sum_{s\in\mathcal{S}} \prod_{k'=1}^{2M} P_{j^*_{q_m},s(k')}(\tilde{x}^{k'})$$

from $P_{q_m,1} = P_{q_m} = P_{j^*_{q_m},2}$ and $P_{q_m,2} = P_{j^*_{q_m}} = P_{j^*_{q_m},1}$. Now plugging (51) and (54) into (44) shows that

$$\tilde{I}^* = (|\mathcal{N}| - M) I^*, \tag{55}$$

since $(|\mathcal{N}| - M) I^* \leq (|\mathcal{N}| - M) I(q, j) \leq \tilde{I}(q, j), \forall q, j$. Noting that $\mathcal{D}[T] \geq \mathcal{D}_{\text{trans}}[T^*_{\text{QCD}}] \sim \frac{\log\gamma}{(|\mathcal{N}|-M)I^*}$ under $\mathcal{A}[T] \geq r$, as $\gamma \to \infty$, it completes the proof of the converse part.

C. Proof of the acheivability for multi-hypothesis BDQCD in Theorem V.2

Now, we first describe the local decision rule at each honest sensor, and then propose two global fault-tolerant decision rules for the acheivability in part (a) and (b) of Theorem V.2 respectively.

1) Local decision rule: "Soft" Matrix CUSUM: Since the honest sensors are not allowed to cooperate with each other, it is natural to adopt the matrix CUSUM algorithm reviewed in Section II-B. Note that for the original matrix CUSUM in [7], since there is only one honest node, it makes perfect sense for the procedure to make a hard decision and terminate after the alarm; however, in our setting, the task is not done yet until the fusion center has determined the occurrence of the event. Therefore, we adapt the matrix CUSUM procedure to the "soft" version as follows. Whenever a $Y_{t,q}^k$ under (9) exceeds the threshold h at time index t, the hypothesis H_q is *softly* decided by informing the fusion center that this hypothesis is *acceptable* at the sensor k. Now each honest sensor may keep monitoring the event and report multiple hypotheses to the fusion center. Later in Sec. V-C3, this soft version will help us resolve the "undecidable event", which disables the fusion center to make a conclusive decision.

Formally, for the soft matrix CUSUM procedure, a hypothesis H_q is acceptable by the node k at time

$$\sigma_k^q(h) := \inf\left\{t \in \mathbb{N} : Y_{t,q}^k \geq h\right\}. \tag{56}$$

In contrast, for the original matrix CUSUM [7], a hypothesis H_q is hard decided at time $\sigma_k^q(h)$ if $\sigma_k^q(h)$ equals to

$$\sigma_k(h) := \min_{\hat{q}\in[Q]} \sigma_k^{\hat{q}}(h) \text{ and } q = \arg\max_{\hat{q}\in[Q]} \left(Y_{t,\hat{q}}^k|_{t=\sigma_k(h)}\right).$$

2) Global fault-tolerant decision rules: As a baseline, the one-shot rule which uses the original matrix CUSUM is first introduced as

One-shot d-th alarm: This family of rules is a direct extension of the one-shot rule for the binary case in [13], [15] to the multi-hypothesis setting. Each sensor adopts the original matrix CUSUM [7] as its local report mechanism and reports the first acceptable non-zero hypothesis as soon as the sensor finds it. The fusion center declares that an abrupt event has occurred at the first time that a hypothesis, say H_q, has received d local reports. It also declares that the hypothesis H_q is true.

Now, we propose two rules based on the "soft" matrix CUSUM.

i) Multi-shot d-th alarm $\tau^m_{(d)}(h)$: This family of rules requires each sensor to adopt the soft version of matrix CUSUM and to alarm whenever a hypothesis $H_{\hat{q}}$, $\hat{q} \in [Q]$, is acceptable. Formally, for each $k \in \mathcal{N}$, sensor k reports $H_{\hat{q}}$ at the time index $\sigma^{\hat{q}}_k(h)$, for every $\hat{q} \in [Q]$. In this reporting mechanism, we stipulate that for each sensor, every hypothesis can be reported at most once, and a reported hypothesis cannot be withdrawn. In other words, once reported by a sensor, a hypothesis will be promoted as a candidate by that sensor ever since. If a tie happens at an honest node k, then all the hypothesis indexes have the same $\sigma^{\hat{q}}_k(h)$ will be reported one after another, starting from the one with the largest $Y^k_{t,\hat{q}}$. For the case where two or more hypotheses have the same $Y^k_{t,\hat{q}}$, we break the tie randomly. Consecutive ties and/or multi-way ties can be easily resolved by equipping each node with a queue of size $Q-1$ and clearing the queue on the first-come first-serve basis. The fusion center declares that an abrupt event has occurred at the first time that a hypothesis, say H_q, has been deemed acceptable by d sensors. It also declares that the hypothesis H_q is true.

ii) Simultaneous d-th alarm $\tau^s_{(d)}(h)$: Each sensor constantly transmits Q bits local decision at time index t to indicate whether $H_{\hat{q}}$ is acceptable, $\forall \hat{q} \in [Q]$. The fusion center declares that an abrupt event of type q has occurred at the first time that a hypothesis, say H_q, has been simultaneously accepted by no less than d sensors.

We note that the three families of rules have different bandwidth and/or energy requirements. The one-shot scheme is the most bandwidth- and energy-efficient one as it requires each link to support $\lceil \log_2 Q \rceil$ bits and this link is used only once. The multi-shot scheme also requires links to support $\lceil \log_2 Q \rceil$ bits, but each link may be used up to Q times. As for the simultaneous rule, it requires each link to support Q bits and each link is constantly used. Also, it is worth noting that the soft matrix CUSUM reduces to the original CUSUM adopted in [13] when $Q = 1$, i.e. binary hypothesis. Thus, the proposed multi-shot and simultaneous d-th alarm include the one-shot and voting rules in [13], [15] as special cases, respectively. Moreover, depending on the application at hand, the fusion center can opt to stop only once or multiple times. When the fusion center chooses to stop only once as [13], $T_\rho = \infty$ in (4) for $\rho > 1$. For the scenario where the fusion center makes multiple alarms, it restarts with the same global decision rule after each alarm at T_ρ. Our upcoming Proposition V.1 applies to both scenarios.

3) Performance analysis: We now carry out the worst-case analysis on the performance of the multi-shot d-th alarm and simultaneous d-th alarm rules. For the one-shot d-th alarm, we point out a notable difference from the binary counterpart [13] which significantly degrade the performance from the converse in Theorem V.1. The *undecidable event* may happen: it is possible that there is no non-zero hypothesis index with enough local alarms for making a decision, even though all honest sensors have raised alarms; thereby, the detection delay is infinity. Unfortunately, even more advanced multi-shot d-th alarm can only achieve the converse when $|\mathcal{N}| = M + 1$ from the upcoming analysis. When $|\mathcal{N}| = M+1$, if the one-shot d-th alarm is used, the compromised sensors can easily trigger the undecidable event.

To further characterize (5) and (6) for the two proposed rules, we will prove asymptotic dominance results in upcoming Lemma V.1, which greatly simplifies the delay analysis. Intuitively, although there are multiple sensors and $Q + 1$ hypotheses, for each honest sensor being considered, we only have to examine the statistics between the q-th hypothesis and the one that is "*closest*" to q, for every $q \in [Q]$. Note that this intuition also comply with our asymptotic converse. Before introducing Lemma V.1 and the complete analysis, some definitions and a proposition regarding the compromised sensors will be provided first. Similarly to [13], from (56), we define the ordered time indexes $\sigma^q_{(1)}(h) \leq \ldots \leq \sigma^q_{(|\mathcal{N}|)}(h)$, for all $q \in [Q]$, over $|\mathcal{N}|$ honest sensors as if there is no compromised sensor, for softly deciding hypothesis H_q (cf. the qth row of CUSUM matrix (9)) We also let $S^q_\ell(h)$ be the first time that the hypothesis H_q is simultaneously softly-decided by ℓ honest sensors, defined as

$$\inf\left\{t \in \mathbb{N} : Y^k_{t,q} \geq h \ \forall k \in \mathcal{L}, \text{for some } \mathcal{L} \subset [\mathcal{N}], |\mathcal{L}| = \ell\right\}. \tag{57}$$

Finally, we will use $\mathbb{E}^q_\nu$ to represent the expectation when the change with hypothesis index q happens at time ν and the compromised sensors are absent.

To continue the worst-case analysis in (5) and (6), recall that all the compromised sensors know the actual ν and the actual hypothesis q. They can then collaboratively attack/confuse the fusion center. Thus, it is obvious that choosing any $d \leq M$ is bad for false alarm or false isolation in (6), while any $d > |\mathcal{N}|$ is bad for detection delay in (5). We therefore confine the choice of d to some reasonable region and obtain the following result.

Proposition V.1. *Fix $h > 0$. For any positive integer Q, and $d \in \{M+1, ..., |\mathcal{N}|\}$, for multi-shot*

d-th alarm, we have

$$\mathcal{A}[\tau^m_{(d)}(h)] \geq \min_{q\in[Q]^+} \min_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}_0^{q,\emptyset}[\sigma^{\hat{q}}_{(d-M)}(h)], \tag{58}$$

$$\mathcal{D}[\tau^m_{(d)}(h)] \leq \max_q \mathbb{E}_0^{q,\emptyset}[\sigma^q_{(d)}(h)] + Q - 1; \tag{59}$$

while for simultaneous d-th alarm

$$\mathcal{A}[\tau^s_{(d)}(h)] \geq \min_{q\in[Q]^+} \min_{\hat{q}\in[Q]\setminus\{q\}} \mathbb{E}_0^{q,\emptyset}[S^{\hat{q}}_{d-M}(h)]; \tag{60}$$

$$\mathcal{D}[\tau^s_{(d)}(h)] \leq \max_q \mathbb{E}_0^{q,\emptyset}[S^q_d(h)]. \tag{61}$$

Proof: See Appendix B. ■

Based on Proposition V.1, in what follows, we provide explicit upper bounds on the detection delay. Lemma V.1 is shown first, which states that for each hypothesis $q \in [Q]$, although there are total $Q+1$ hypotheses, one only has to worry about the one that is "*closest*" to q in terms of the KL divergence. By writing $\mathbb{P}_q$ for $\mathsf{P}_0^{q,g=\emptyset}$ and recall the definition of j^*_q in Assumption V.1, we have:

Lemma V.1. *Suppose h is large enough and Assumption V.1 holds. For any $q \in [Q]$, it holds $\mathbb{P}_q$-a.s. that*

(i) The first time H_q is softly decided at the honest sensor k, $\sigma^q_k(h)$ in (56)*, equals to*

$$\sigma^{q,j^*_q}_k(h) := \inf\{t \in \mathbb{N} : Y^k_t(q, j^*_q) \geq h\}. \tag{62}$$

(ii) For any $|\mathcal{N}| \geq d \geq 1$, the first time H_q is simultaneously softly-decided by d honest sensors, $S^q_d(h)$, equals to

$$S^{q,j^*_q}_d(h) := \inf\left\{t \in \mathbb{N} : Y^{(|\mathcal{N}|-d+1)}_t(q, j^*_q) \geq h\right\}, \tag{63}$$

where $Y^{(1)}_t(q, j^*_q) \leq \ldots \leq Y^{(|\mathcal{N}|)}_t(q, j^*_q)$ are the ordered CUSUM statistics of $Y^k_t(q, j^*_q)$, for hypotheses q and j^*_q at time t.

Proof: See Appendix C. ■

We are now ready to present the results on the asymptotic delay performance. Let $Z_{(1)}$, $Z_{(2)}$, …, $Z_{(|\mathcal{N}|)}$ be the order statistics of independent standard normal random variables. For each $d \in \{1, 2, \ldots, |\mathcal{N}|\}$, we denote by ξ_d the expected value of $Z_{(d)}$. Moreover, for each $q \in [Q]$, we set $D^q_{d:|\mathcal{N}|} := \xi_d \sqrt{\frac{\sigma^2(q,j^*_q)}{I^q}}$.

Theorem V.3. *Suppose Assumption V.1 holds. As $h \to \infty$, for any $q \in [Q]$ and $1 \leq d \leq |\mathcal{N}|$, we have*

$$\mathbb{E}_0^{q,\emptyset}[\sigma_{(d)}^q(h)] = \frac{h}{I^q} + D_{d:|\mathcal{N}|}^q \sqrt{h}(1+o(1)), \tag{64}$$

and the detection delay of the multi-shot d-th alarm in (59) *is upper-bounded as*

$$\mathcal{D}[\tau_{(d)}^m(h)] \leq \max_q \left(\frac{h}{I^q} + D_{d:|\mathcal{N}|}^q \sqrt{h}(1+o(1)) \right). \tag{65}$$

Proof: See Appendix D. ■

Theorem V.4. *Suppose Assumption V.1 holds. As $h \to \infty$, for any $q \in [Q]$ and $1 \leq d \leq |\mathcal{N}|$, we have*

$$\mathbb{E}_0^{q,\emptyset}\left[S_d^q(h)\right] \leq \frac{h}{I^q} + D_{d:|\mathcal{N}|}^q \sqrt{h}(1+o(1)), \tag{66}$$

and the detection delay of the simultaneous d-th alarm in (61) *is upper-bounded as*

$$\mathcal{D}[\tau_{(d)}^s(h)] \leq \max_q \left(\frac{h}{I^q} + D_{d:|\mathcal{N}|}^q \sqrt{h}(1+o(1)) \right). \tag{67}$$

Proof: See Appendix D. ■

The asymptotic performance of the mean time to false alarm/isolation of the proposed families of rules are given in the following.

Theorem V.5. *Fix $h > 0$. If $M < d \leq |\mathcal{N}|$, the mean time to a false alarm or a false isolation $\mathcal{A}[\tau_{(d)}^m(h)]$ in* (58) *for the multi-shot d-th alarm is lower-bounded by*

$$\frac{d-M}{(d-M+1)} \begin{pmatrix} |\mathcal{N}| \\ d-M \end{pmatrix}^{\frac{-1}{d-M}} \exp(h). \tag{68}$$

Proof: See Appendix E. ■

Theorem V.6. *Fix $h > 0$. If $M < d \leq |\mathcal{N}|$, the mean time to a false alarm or a false isolation $\mathcal{A}[\tau_{(d)}^s(h)]$ in* (60) *is lower-bounded by*

$$\frac{1}{2} \begin{pmatrix} |\mathcal{N}| \\ d-M \end{pmatrix}^{-1} \exp\left((d-M)h\right). \tag{69}$$

Proof: See Appendix E. ■

Although the delay upper bounds in Theorems V.3 and V.4 are identical, we will show the superiority of the simultaneous rule when detection delay and mean time to false alarm/isolation

are jointly considered, which also validates Theorem V.2. For simultaneous d-th alarm $\tau^s_{(d)}(h)$, one can ensure $\mathcal{A}[\tau^s_{(d)}](h) \geq \gamma$ from (69) by selecting local threshold

$$\frac{1}{d-M}\left(\log\gamma + \log\left(2\left(\begin{array}{c}|\mathcal{N}|\\ d-M\end{array}\right)\right)\right). \tag{70}$$

Also by plugging h in (70) into (67), with $\gamma \to \infty$,

$$\mathcal{D}[\tau^s_{(d)}(h)] \lesssim \max_q \left(\frac{\log\gamma}{(d-M)I^q}\right). \tag{71}$$

Then part (a) of Theorem V.2 is valid since $d = |\mathcal{N}|$ minimizes the right hand side above. Moreover, for the multi-shot d-th alarm $\tau^m_{(d)}(h)$, one can ensure $\mathcal{A}[\tau^m_{(d)}(h)] \geq \gamma$ from (68) by selecting local threshold

$$h = \log\gamma + \frac{1}{d-M}\log\left(\begin{array}{c}|\mathcal{N}|\\ d-M\end{array}\right) + \log\left(\frac{d-M+1}{d-M}\right). \tag{72}$$

Plugging h in (72) into (65) and let $\gamma \to \infty$, we have

$$\mathcal{D}[\tau^m_{(d)}(h)] \lesssim \max_q \left(\frac{\log\gamma}{I^q}\right). \tag{73}$$

This validates part (b) of Theorem V.2.

Remark V.1. *Apart from being the building block of our proof, Lemma V.1 also reveals another practical benefit. It basically confirms that for large h, for each row of the matrix CUSUM, i.e., each abnormal hypothesis H_q, an honest sensor only has to compute and update the CUSUM statistics $Y^k_t(q, j^*_q)$. This is particularly useful in applications where sensors are subject to stringent energy constraints. Moreover, although Lemma V.1 only shows asymptotic optimality of this approach, we provide in Fig. 3 an example showing that the intuition of updating only the closest hypothesis also applies to small h. In this example, we consider $Q = 2$ with P_0, P_1, P_2 the PDFs of Gaussian random variables with means $0, 1, -1$ and same variances σ^2, respectively. Instead of the original CUSUM matrix, we use the following reduced one*

$$\begin{bmatrix} Y^k_t(1,0) & \infty \\ \infty & Y^k_t(2,0). \end{bmatrix} \tag{74}$$

As shown in Fig. 3, the detection performances of full CUSUM matrix and the reduced one are almost identical, as expected.

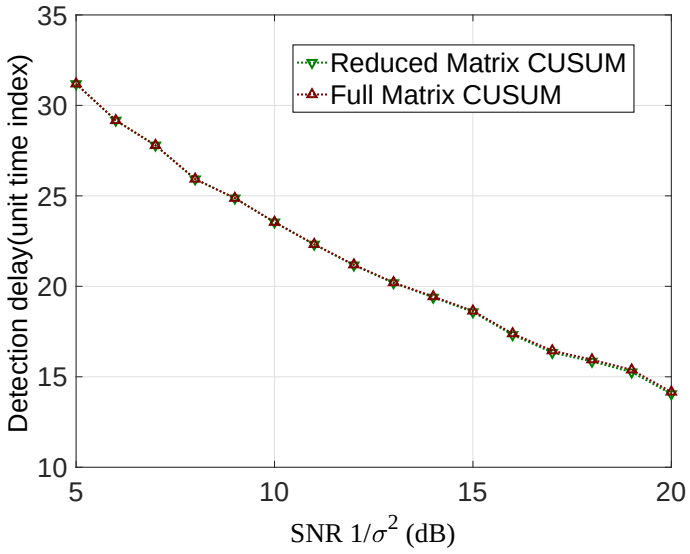


Fig. 3. Detection delay of full Matrix CUSUM and that of the reduced one in (74) for an honest sensor. The local threshold is set such that $\exp(h) = 10^4$ ($h \approx 9.21$) and each curve is calculated based on 20000 realizations.

VI. Game-theoretic formulation

In this section, we formulate a leader-follower Stackelberg game [16, Section 3.6] for the considered BDQCD, where the fusion center and honest sensors act as the leader and the compromised sensors act as the follower. It turns out that the characterization of the first-order optimality in the previous sections will help us characterize the Stacklberg equilibrium.

In the game, the information available to the two players are as follows:

- The follower knows the leader's strategy g_1, all the current and past local observations $\mathbf{X}_1^t$, the change time ν, and the actual hypothesis.
- The leader is oblivious of the exact indexes of compromised sensors, but knows the maximum number of compromised sensors M.

Now we define the strategy spaces of the two players. With $Q+1$ hypotheses, we assume the noiseless link of each sensor can support (at least) Q bits and use the $K \times 1$ vector $\hat{\boldsymbol{\lambda}}_t \in \mathbb{Z}_{2^Q}^K$ to denote the Q-bit local decisions at time t, where $\mathbb{Z}_{2^Q}$ is the integer ring modulo 2^Q. The compromised sensors cooperatively form the attack vector $\mathbf{e}_t \in \mathbb{Z}_{2^Q}^K$ which has at most M non-zero components, reflecting that there are at most M compromised sensors. At time t, the fusion

center receives

$$\hat{\boldsymbol{\lambda}}_t + \mathbf{e}_t, \tag{75}$$

where the addition is over $\mathbb{Z}_{2^Q}$. A strategy g_1 of the leader at time t includes a local decision rule that maps $X_1^k, \ldots, X_t^k$ into the kth entry of $\hat{\boldsymbol{\lambda}}_t$ at each sensor $k \in [K]$ (the leader treats all K sensors as honest), and a stopping rule at the fusion center which maps $K \times t$ matrix $[\hat{\boldsymbol{\lambda}}_1 + \mathbf{e}_1, \ldots, \hat{\boldsymbol{\lambda}}_t + \mathbf{e}_t]$ to a decision $\hat{q}_t \in [Q]^+$. An alarm is fired if $\hat{q}_t \neq 0$. The stopping time for type $q \in [Q]$ is given in (4) and again let T be the first alarm time. A strategy g_2 of the follower at time t is the vector $\mathbf{e}_t$ in (75), where all K elements are from $\mathbb{Z}_{2^Q}$ but elements in a subset of indices with size $|\mathcal{N}|$ (corresponding to indexes of honest sensors) is deterministically 0. We use $\mathcal{G}_1$ and $\mathcal{G}_2$ to denote the pure-strategy spaces of aforementioned g_1 and g_2, respectively.

We first focus on the binary case as in Section IV, i.e., $Q = 1$, and refer to the game as the *binary BDQCD Stackelberg game*. From the detection delay in (2) and false alarm in (3), we define the corresponding performance metrics under strategies (g_1, g_2) as

$$\begin{aligned} \mathcal{D}(g_1, g_2) &:= \sup_{\nu} \operatorname{ess\,sup} \mathbb{E}_{\nu}^{g_2}[(T-\nu)^+ | \mathbf{X}_1^{\nu}], \quad \text{and} \\ \mathcal{A}(g_1, g_2) &:= \mathbb{E}_{\infty}^{g_2}[T], \end{aligned} \tag{76}$$

respectively, where we note that the stopping time T is a function of (g_1, g_2). Since we wish the mean time to false alarm to be larger than a given γ, the cost for the leader is then defined as

$$J^1(g_1, g_2) \triangleq \lim_{\gamma \to \infty} \left(\frac{\mathcal{D}(g_1, g_2)}{\log \gamma} + I_- \left(\gamma - \mathcal{A}(g_1, g_2) \right) \right), \tag{77}$$

where function $I_-(u)$, as defined in [25, Section 11.2], is ∞ when $u > 0$ and zero otherwise. Note that the change point ν is known at the follower, and thus g_2 can be different before and after the change time ν. The cost for the follower is the mean time to false alarm $J^2(g_1, g_2) \triangleq \mathcal{A}(g_1, g_2)$ as the follower wants to sabotage detection by making false alarm more frequent.

For our game, the Stackelberg equilibrium strategy for the leader and cost are defined as follows. As [16, Definition 4.1], we define

Definition VI.1. *Fix $\varepsilon > 0$. For any $g_1 \in \mathcal{G}_1$, the set $R_\varepsilon^2(g_1) \subseteq \mathcal{G}_2$ defined by*

$$R_\varepsilon^2(g_1) = \left\{ g_2 \in \mathcal{G}_2 : J^2(g_1, g_2) \leq \inf_{\xi \in \mathcal{G}_2} J^2(g_1, \xi) + \varepsilon \right\} \tag{78}$$

is the ε-optimal response set of the follower to the strategy g_1 of the leader.

Definition VI.2. *The Stackelberg cost of the leader is defined as*

$$J^{1*} = \inf_{g_1 \in \mathcal{G}_1} \inf_{\varepsilon > 0} \sup_{g_2 \in R^2_\varepsilon(g_1)} J^1(g_1, g_2). \tag{79}$$

For any $\varepsilon > 0$, a strategy $g_\varepsilon^{1*} \in \mathcal{G}_1$ is an ε-Stackelberg equilibrium strategy for the leader if

$$\inf_{\bar{\varepsilon} > 0} \sup_{g_2 \in R^2_{\bar{\varepsilon}}(g_\varepsilon^{1*})} J^1(g_\varepsilon^{1*}, g_2) \leq J^{1*} + \varepsilon$$

Based on the above definitions, we first prove the following lemma which results in the game solution later in Theorem VI.1.

Lemma VI.1. *For the binary BDQCD Stackelberg game, if there exists a pure strategy $\acute{g}_2 \in \mathcal{G}_2$ that results in a lower bound $J^1(g_1, \acute{g}_2) \geq \eta$ for any $g_1 \in \mathcal{G}_1$, then*

$$\sup_{g_2 \in R^2_\varepsilon(g_1)} J^1(g_1, g_2) \geq J^1(g_1, \acute{g}_2) \geq \eta \tag{80}$$

for any $\varepsilon > 0$.

Proof. We prove that $\sup_{g_2 \in R^2_\varepsilon(g_1)} J^1(g_1, g_2) \geq J^1(g_1, \acute{g}_2)$. If $\acute{g}_2 \in R^2_\varepsilon(g_1)$, then the inequality is trivial. If not, for any $g_2 \in R^2_\varepsilon(g_1)$ we have $\mathcal{A}(g_1, g_2) \leq \mathcal{A}(g_1, \acute{g}_2)$. If $\mathcal{A}(g_1, \acute{g}_2) < \gamma$, then $J^1(g_1, g_2) = J^1(g_1, \acute{g}_2) = \infty$ from (77). Now consider $\mathcal{A}(g_1, \acute{g}_2) \geq \gamma$. If $\mathcal{A}(g_1, g_2) < \gamma \leq \mathcal{A}(g_1, \acute{g}_2)$, from (77),

$$J^1(g_1, \acute{g}_2) \leq J^1(g_1, g_2) = \infty.$$

Otherwise, if $\gamma \leq \mathcal{A}(g_1, g_2) \leq \mathcal{A}(g_1, \acute{g}_2)$, then

$$I_-\left(\gamma - \mathcal{A}(g_1, \acute{g}_2)\right) = I_-\left(\gamma - \mathcal{A}(g_1, g_2)\right) = 0,$$

we construct an attack $\hat{g}_2$ acting as g_2 when $\nu = \infty$ and as $\acute{g}_2$ otherwise. This $\hat{g}_2$ will lie in $R^2_\varepsilon(g_1)$ and result in $J^1(g_1, \hat{g}_2) = J^1(g_1, \acute{g}_2)$, which results in $J^1(g_1, \acute{g}_2) = J^1(g_1, \hat{g}_2) \leq \sup_{g_2 \in R^2_\varepsilon(g_1)} J^1(g_1, g_2)$. □

Theorem VI.1. *For the binary BDQCD Stackelberg game, the Stackelberg cost J^{1*} is $\frac{1}{(|\mathcal{N}|-M)I}$ when $|\mathcal{N}| > M$ and zero elsewhere; and for any $\varepsilon > 0$, the $|\mathcal{N}|$-voting rule is the ε-Stackelberg equilibrium strategy for the leader.*

Proof. From Lemma VI.1, for the converse $J^{1*} \geq \frac{1}{(|\mathcal{N}|-M)I}$ it suffices to construct an attack $\acute{g}_2$ such that for any $g_1 \in \mathcal{G}_1$, we have

$$J^1(g_1, \acute{g}_2) \geq \frac{1}{(|\mathcal{N}| - M)I}, \tag{81}$$

when $|\mathcal{N}| > M$. This is valid from the proof of Theorem IV.1 by choosing $\acute{g}_2$ as the proposed reverse attack. On the other hand, the achievability comes from (11). That is, the $|\mathcal{N}|$-voting rule, which uses only 1 bit from each sensor, achieves

$$\max_{g_2 \in R_\varepsilon^2\left(\tau^s_{(|\mathcal{N}|)}(h)\right)} J^1\left(\tau^s_{(|\mathcal{N}|)}(h), g_2\right) = \frac{1}{(|\mathcal{N}| - M)I}, \quad \forall \varepsilon > 0, \tag{82}$$

by selecting local threshold h for the worst case attack (where all compromised sensors send "1" always) such that

$$\mathcal{A}(\tau^s_{(|\mathcal{N}|)}(h), g_2) \geq \gamma, \ \ \forall g_2 \in R_\varepsilon^2\left(\tau^s_{(|\mathcal{N}|)}(h)\right). \tag{83}$$

Then $J^{1*} \leq \frac{1}{(|\mathcal{N}|-M)I}$ and it concludes the proof. □

We now consider $Q > 1$ and define the multi-hypothesis BDQCD Stackelberg game. From the detection delay defined in (5) and false alarm/isolation defined in (6), we define the corresponding performance metrics under strategy (g_1, g_2) as

$$\mathcal{D}(g_1, g_2) := \sup_{\nu} \sup_{q \in [Q]} \operatorname{ess\,sup} \mathbb{E}_\nu^{q, g_2}[(T - \nu)^+ | \mathbf{X}_1^\nu]. \tag{84}$$

$$\mathcal{A}(g_1, g_2) := \inf_{q \in [Q]^+} \inf_{\hat{q} \in [Q] \setminus \{q\}} \mathbb{E}_0^{q, g_2}[T^{\hat{q}}]. \tag{85}$$

Unfortunately in this case, we are unable to prove results similar to Lemma VI.1 with the follower's cost $J^2(g_1, g_2) = \mathcal{A}(g_1, g_2)$. The main difficulty is that when $\nu = 0$, one needs to consider jointly the mean time to a false alarm and that to a false isolation, when it comes to constructing an attack $\acute{g}_2$. Therefore, we instead define $J^2(g_1, g_2) = -J^1(g_1, g_2)$ and show the following result.

Theorem VI.2. *For multi-hypothesis BDQCD Stackelberg game, the Stackelberg cost is* $J^{1*} = \frac{1}{(|\mathcal{N}|-M)I^*}$ *when* $|\mathcal{N}| > M$ *and zero elsewhere; and for any* $\varepsilon > 0$*, the simultaneous* $|\mathcal{N}|$*-th alarm is the* ε*-Stackelberg equilibrium strategy for the leader.*

Proof. In this case, J^{1*} in (79) becomes

$$\inf_{g_1 \in \mathcal{G}_1} \sup_{g_2 \in \mathcal{G}_2} J^1(g_1, g_2) \tag{86}$$

If we can construct an attack $\acute{g}_2$ such that for any $g_1 \in \mathcal{G}_1$, $J^1(g_1, \acute{g}_2) \geq \frac{1}{(|\mathcal{N}|-M)I}$, and then $\sup_{g_2 \in \mathcal{G}_2} J^1(g_1, g_2) \geq J^1(g_1, \acute{g}_2) \geq \frac{1}{(|\mathcal{N}|-M)I}$ by definition. The game solution simply follows from the proof of Theorem V.1 and V.2 (a) by choosing $\acute{g}_2$ as our reverse attack for the multi-hypothesis case in Section V-B1. □

VII. Conclusions

In this paper, the problem of BDQCD has been studied, where a fusion center sequentially monitors an abrupt event via distributed sensors which might be compromised. Both the binary hypothesis and multi-hypothesis cases have been considered. For the binary case, a novel converse bound for the first-order asymptotic detection delay performance in the large mean time to a false alarm regime has been proved. By comparing the converse bound and the first-order scaling achieved by the existing consensus rule, we have characterized the fundamental limit of binary BDQCD in the large mean time to a false alarm regime (or the small false alarm rate regime in a sense). For the multi-hypothesis BDQCD, the novel converse has been generalized from the binary case and the optimal first-order asymptotic performance has again been characterized. Along with establishing this fundamental result, two novel families of stopping rules have been proposed, namely the multi-shot d-th alarm and the simultaneous d-th alarm. The former is much more energy-efficient and bandwidth efficient while the latter can achieve asymptotically optimal performance under sufficient link bandwidth whenever there are more honest sensors than compromised ones. Finally, a leader-follower Stackelberg game has been formulated based on the BDQCD problem discussed. The asymptotically optimal stopping rule and the asymptotically worst attack proposed for BDQCD have led us to the game solution, in which the leader adopts the proposed asymptotically optimal stopping rule (i.e., the simultaneous rule) and the follower employs the corresponding asymptotically worst attack.

Appendix A
Lemmas

In this appendix, some useful lemmas are presented and their proofs are given.

Lemma A.1. *For any decision rule T, $\nu \geq 0$, $q \in [Q]$, and attack strategy g,*

$$\operatorname{ess\,sup} \mathbb{E}_{\nu}^{q,g}[(T-\nu)^{+} \mid \mathbf{X}_1^{\nu}] = \operatorname{ess\,sup} \mathbb{E}_{\nu}^{q,g}[T-\nu \mid T>\nu,\ \mathbf{X}_1^{\nu}]. \tag{87}$$

Proof: For any decision rule T, $\nu \geq 0$, $q \in [Q]$, and attack strategy g, consider the subset S of $\mathbb{R}^{K\times\nu}$ on which $T > \nu$. By the definition of S, observe that

$$\mathsf{P}_{\nu}^{q,g}[T>\nu \mid \mathbf{X}_1^{\nu}] = \begin{cases} 1, & \text{on } S, \\ 0, & \text{on } S^c. \end{cases} \tag{88}$$

Also, note that $\mathbb{E}_\nu^{q,g}[(T-\nu)^+ \mid \mathbf{X}_1^\nu]$ is well-defined on $\mathbb{R}^{K\times\nu}$, and by definition constantly zero on S^c. On the other hand,

$$\mathbb{E}_\nu^{q,g}[T-\nu \mid T>\nu,\ \mathbf{X}_1^\nu] \tag{89}$$

is well-defined only on S; specifically, since $\mathsf{P}_\nu^{q,g}[T>\nu \mid \mathbf{X}_1^\nu]=0$ on S^c, the conditional probability needed to evaluate (89) is not well-defined on S^c.

Now, by the law of total expectation,

$$\begin{aligned}
&\mathbb{E}_\nu^{q,g}[(T-\nu)^+ \mid \mathbf{X}_1^\nu] \\
&\quad= \mathbb{E}_\nu^{q,g}[(T-\nu)^+ \mid T>\nu,\ \mathbf{X}_1^\nu]\ \mathsf{P}_\nu^{q,g}[T>\nu \mid \mathbf{X}_1^\nu] \\
&\qquad+ \mathbb{E}_\nu^{q,g}[(T-\nu)^+ \mid T\le\nu,\ \mathbf{X}_1^\nu]\ \mathsf{P}_\nu^{q,g}[T\le\nu \mid \mathbf{X}_1^\nu] \\
&\quad= \mathbb{E}_\nu^{q,g}[T-\nu \mid T>\nu,\ \mathbf{X}_1^\nu] \quad \text{on } S,
\end{aligned} \tag{90}$$

where the last equality follows from (88). Since $\mathbb{E}_\nu^{q,g}[(T-\nu)^+ \mid \mathbf{X}_1^\nu]$ is constantly zero on S^c, the desired result is a direct consequence of (90). ■

Lemma A.2. *In binary genie-aided BCQCD, for any general (not necessarily masked symmetric) fusion rule $T'(\{\mathbf{X}_t\}_{t\ge1})$, there is a masked symmetric rule $T(\{\mathbf{X}_t\}_{t\ge1})$ that is not worse than $T'(\{\mathbf{X}_t\}_{t\ge1})$.*

Proof: The proof is a constructive one. Recall we have defined $\pi:[K]\to[K]$ a masked permutation function that permutes the first $2M$ entries while keeps the remaining $|\mathcal{N}|-M$ entries unchanged. Let Π_{2M} be the collection of all ($2M!$ in total) such π. For $\mathbf{X}_t=[X_t^1,\dots,X_t^K]$, we slightly abuse the notation to write $\pi(\mathbf{X}_t)=[X_t^{\pi(1)},\dots,X_t^{\pi(K)}]$. Let

$$T(\{\mathbf{X}_t\}_{t\ge1}) = \frac{1}{2M!}\sum_{\pi\in\Pi_{2M}} T'(\{\pi(\mathbf{X}_t)\}_{t\ge1}). \tag{91}$$

Following the proof of part 1 of [14, Lemma 4.2], one can show that $T(\{\mathbf{X}_t\}_{t\ge1})$ is indeed a masked symmetric strategy. Now the detection delay $\mathcal{D}_{\text{genie}}[T(\{\mathbf{X}_t\}_{t\ge1})]$ is

$$\sup_{s\in\mathcal{S},\nu} \operatorname{ess\,sup} \mathbb{E}_\nu^s\left[\left(\frac{1}{2M!}\sum_{\pi\in\Pi_{2M}} T'(\{\pi(\mathbf{X}_t)\}_{t\ge1})-\nu\right)^+ \middle| \mathbf{X}_1^\nu\right].$$

Thus $\mathcal{D}_{\text{genie}}[T(\{\mathbf{X}_t\}_{t\geq 1})]$ is no longer than

$$\begin{aligned}
&\frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\sup_{s\in\mathcal{S},\nu}\operatorname{ess\,sup}\mathbb{E}_\nu^s\left[(T'(\{\pi(\mathbf{X}_t)\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu\right]\\
&\overset{(a)}{=}\frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\sup_{s\in\mathcal{S},\nu}\operatorname{ess\,sup}\mathbb{E}_\nu^{s\circ\pi^{-1}}\left[(T'(\{\mathbf{X}_t\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu\right]\\
&\overset{(b)}{=}\frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\sup_{s'\in\mathcal{S},\nu}\operatorname{ess\,sup}\mathbb{E}_\nu^{s'}\left[(T'(\{\mathbf{X}_t\}_{t\geq 1})-\nu)^+|\mathbf{X}_1^\nu\right]\\
&=\frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\mathcal{D}_{\text{genie}}[T'(\{\mathbf{X}_t\}_{t\geq 1})]=\mathcal{D}_{\text{genie}}[T'(\{\mathbf{X}_t\}_{t\geq 1})].
\end{aligned}\tag{92}$$

Note that essential supremum of the right-hand side of (a) is taken under the probability measure whose density is specified by (17) under $\theta=0$ and the compromised group assignment $s\circ\pi^{-1}$. Then (a) can be proved similar to (21a) by the fact

$$\begin{aligned}
&\mathbb{E}_\nu^s\left[(T'(\{\pi(\mathbf{X}_t)\}_{t\geq 1})-\nu)^+\;\middle|\;\mathbf{X}_1^\nu\right](x)=\\
&\mathbb{E}_\nu^{s\circ\pi^{-1}}\left[(T'(\{\mathbf{X}_t\}_{t\geq 1})-\nu)^+\;\middle|\;\mathbf{X}_1^\nu\right](\pi(x)),\forall x\in\mathbb{R}^{K\times\nu}
\end{aligned}$$

since the permutation $\pi(.)$ is one-to-one; and (b) is due to the fact that $\{s\circ\pi^{-1}|s\in\mathcal{S}\}=\mathcal{S}$. We can similarly show that for the mean time to false alarm of the new rule, $\mathcal{A}_{\text{genie}}[T(\{\mathbf{X}_t\}_{t\geq 1})]\geq\mathcal{A}_{\text{genie}}[T'(\{\mathbf{X}_t\}_{t\geq 1})]$. Then we conclude that the masked symmetric strategy $T(\{\mathbf{X}_t\}_{t\geq 1})$ is at least as good as $T'(\{\mathbf{X}_t\}_{t\geq 1})$. ■

Lemma A.3. *In multiple-hypothesis genie-aided BCQCD, for any general (not necessarily masked symmetric) fusion rule $T'(\{\mathbf{X}_t\}_{t\geq 1})$, there is a masked symmetric rule $T(\{\mathbf{X}_t\}_{t\geq 1})$ that has longer mean time to a false alarm or a false isolation than $T'(\{\mathbf{X}_t\}_{t\geq 1})$.*

Proof: Again, the symmetrized rule T is formed as (91) in the binary BCQCD. It is obvious

that

$$\begin{aligned}
\mathcal{A}_{\text{genie}}[T] &= \inf_{q\in[Q]^+}\inf_{s\in\mathcal{S}}\inf_{\hat{q}\in[Q]\setminus\{q\}}\mathbb{E}_0^{q,s}[T^{\hat{q}}(\{\mathbf{X}_t\}_{t\geq 1})] \\
&= \inf_{q\in[Q]^+}\inf_{s\in\mathcal{S}}\inf_{\hat{q}\in[Q]\setminus\{q\}}\mathbb{E}_0^{q,s}\left[\frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}T'^{\hat{q}}(\{\pi(\mathbf{X}_t)\}_{t\geq 1})\right] \\
&\geq \frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\inf_{q\in[Q]^+}\inf_{s\in\mathcal{S}}\inf_{\hat{q}\in[Q]\setminus\{q\}}\mathbb{E}_0^{q,s}\left[T'^{\hat{q}}(\{\pi(\mathbf{X}_t)\}_{t\geq 1})\right] \\
&= \frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\inf_{q\in[Q]^+}\inf_{s\in\mathcal{S}}\inf_{\hat{q}\in[Q]\setminus\{q\}}\mathbb{E}_0^{q,s\circ\pi^{-1}}\left[T'^{\hat{q}}(\{\mathbf{X}_t\}_{t\geq 1})\right] \\
&= \frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\inf_{q\in[Q]^+}\inf_{s'\in\mathcal{S}}\inf_{\hat{q}\in[Q]\setminus\{q\}}\mathbb{E}_0^{q,s'}\left[T'^{\hat{q}}(\{\mathbf{X}_t\}_{t\geq 1})\right] \\
&= \frac{1}{2M!}\sum_{\pi\in\Pi_{2M}}\mathcal{A}_{\text{genie}}[T'] = \mathcal{A}_{\text{genie}}[T'].
\end{aligned} \tag{93}$$

As a remark, the above proof generalizes that for the mean time to false alarm part for the binary BCQCD which is omitted in Lemma A.2. ■

Lemma A.4. *For any centralized multi-sensor and $(Q+1)$-hypothesis QCD rule $\tilde{T}$, subject to $\mathcal{A}[\tilde{T}]\geq\gamma$, the detection delay is lower-bounded by*

$$\mathcal{D}[\tilde{T}]\gtrsim\frac{\log\gamma}{\tilde{I}^*},\ \textit{as}\ \gamma\to\infty,$$

where $\tilde{I}^*$ is defined in (44)*.*

Proof: One can follow the same arguments in the proof of [6, Theorem 2] to show this Lemma. Specifically, for any $q\in[Q]$, take an arbitrary $\varepsilon_q\in(0,1)$. We extend the sequence of additional stopping variables $\tilde{T}_{a,0}:=0<\tilde{T}_{a,1}<\tilde{T}_{a,2}<...$ introduced in the beginning of the proof of [6, Theorem 2] into multi-sensor version as

$$\begin{aligned}
\tilde{T}_{a,i+1} &= \max_{q\in[Q]}\tilde{T}^q_{a,i+1}, \\
\tilde{T}^q_{a,i+1} &= \inf\left\{n\geq\tilde{T}_i+1:\frac{\tilde{P}_q(\tilde{\mathbf{X}}_{\tilde{T}_i+1})...\tilde{P}_q(\tilde{\mathbf{X}}_n)}{\tilde{P}_0(\tilde{\mathbf{X}}_{\tilde{T}_i+1})...\tilde{P}_0(\tilde{\mathbf{X}}_n)}\leq\varepsilon_q\right\}.
\end{aligned}$$

Then the rest of the proof simply follows [6]. ■

APPENDIX B
PROOF OF PROPOSITION V.1

To prove (58) and (60), we note that in the worst case, all the compromised sensors can raise alarms about the same hypothesis continuously. This implies that as soon as $d - M$, $d \in \{M+1, ..., |\mathcal{N}|\}$, honest sensors raise alarms of the same hypothesis, the compromised sensors can cooperatively enforce a false alarm event. If the fusion center stops only once, for the false isolation, it may declare the correct decision before the first $\sigma^{\hat{q}}_{(d-M)}(h)$ for $\tau^m_{(d)}(h)$ (or $S^{\hat{q}}_{d-M}(h)$ for $\tau^s_{(d)}(h)$) and then $T^{\hat{q}} = \infty$ in (6), which results in a lower bound instead of equality in (58) (and (60)). When the fusion center stops multiple times, recall that in both $\tau_d(h)$ and $\tau^s_{(d)}(h)$, the fusion center resets local CUSUM matrices of all sensors to the all zero matrix after each stop time. The mean of false alarm or isolation time $T^{\hat{q}}$ of $\tau_d(h)$ (respectively $\tau^s_{(d)}(h)$) is clearly lower bounded by that obtained by applying $\tau_d(h)$ (respectively $\tau^s_{(d)}(h)$) but without reset, which corresponds to the right hand side of (58) (respectively (60)).

For the detection delay (5), the worst-case attack $\mathcal{G}$ for both the proposed stopping algorithms happens when all compromised sensors always output local decisions corresponding to H_0. Moreover, we note that both the global decision rules mentioned above are non-decreasing functions in each entry of local CUSUM matirx $Y^k_t(q,j)$ in (8) and the worst CUSUM statistic that the pre-change observations can impose is $Y^k_t(q,j) = 0, t \leq \nu$. Hence, for the proposed algorithms, in the worst case, T is not a function of previous observations and Lemma 3 in [13] can be applied to show the equivalence between (5) and

$$\mathcal{D}[T] = \sup_{q\in[Q]} \sup_{g} \mathbb{E}^{q,g}_0[T], \tag{94}$$

which corresponds to the scenario where the change occurs at $t = 0$. We therefore only have to consider as the worst-case expected detection delay in the sequel. For mulit-shot $\mathcal{D}[\tau^m_{(d)}(h)]$ in (59), the fusion center has to wait for d honest sensors accepting the true H_q. However, false isolation $q_f \neq q$ may still happens if $\sigma^{q'}_{(d)}(h) < \sigma^{q}_{(d)}(h)$, $q' \in [Q] \setminus \{q\}$. Moreover, the longest extra delay caused by ties is $Q - 1$. Thus we have the upper-bound in (59). The upper-bound for simultaneous $\mathcal{D}[\tau^s_{(d)}(h)]$ in (61) can be obtained similarly.

APPENDIX C
PROOF OF LEMMA V.1

Fix $q \in [Q]$ and $k \in \mathcal{N}$. For any $j \in [Q]^+$ with $j \neq q$, the CUSUM statistics $Y_t^k(q,j)$ at sensor k can be decomposed as $Y_t^k(q,j) = Z_t^k(q,j) + \xi_t^k(q,j)$, where

$$Z_t^k(q,j) := \sum_{s=1}^{t} \log\left(\frac{P_q(X_s^k)}{P_j(X_s^k)}\right), \xi_t^k(q,j) := -\min_{0\le s<t} Z_s^k(q,j).$$

Under $\mathbb{P}_q$, $Z^k(q,j)$ is a random walk with drift $I(q,j) > 0$ and variance $\sigma^2(q,j) < \infty$. It follows that $Z_t^k(q)$, defined as $(Z_t^k(q,1), \ldots, Z_t^k(q,q-1), Z_t^k(q,0), Z_t^k(q,q+1), \ldots, Z_t^k(q,Q))$, is a Q-dimensional random walk. Also $Y_t^k(q)$, which is similarly defined as $Z_t^k(q)$ by replacing $Z_t^k(q,j)$ with $Y_t^k(q,j)$, is a Q-dimensional perturbed random walk, as discussed in [26, Section 6.10][4].

Now for any $j \in [Q]^+$ with $j \notin \{q, j_q^*\}$, at time index $\sigma_k^{q,j_q^*}(h)$ in (62), the CUSUM statistics for hypotheses (q,j)

$$\frac{Y^k_{\sigma_k^{q,j_q^*}(h)}(q,j)}{h} \to \frac{I(q,j)}{I(q,j_q^*)} = \frac{I(q,j)}{I^q} \quad \text{as } h \to \infty, \qquad \mathbb{P}_q\text{-a.s.},$$

by [26, Theorem 10.1, p.206]. This, together with Assumption V.1, implies that it holds $\mathbb{P}_q$-a.s. that $\forall j \in [Q]^+ \setminus \{q, j_q^*\}$

$$\frac{Y^k_{\sigma_k^{q,j_q^*}(h)}(q,j)}{h} > 1, \tag{95}$$

as h is large enough. Now, observe that from (56), $\sigma_k^q(h) = \inf\left\{t \in \mathbb{N} : \min_{0\le j\le Q,\ j\neq q} Y_t^k(q,j) \ge h\right\}$, and the RHS equals to

$$\begin{aligned}&\inf\left\{t \in \mathbb{N}\colon \min_{0\le j\le Q,\ j\notin\{q,j_q^*\}} Y_t^k(q,j) \ge h \text{ and } Y_t^k(q,j_q^*) \ge h\right\}\\ &= \sigma_k^{q,j_q^*}(h), \quad \text{as } h \text{ is large enough}, \qquad \mathbb{P}_q\text{-a.s.},\end{aligned}$$

where the last line follows from (95). Since this relation is true for all $k \in \mathcal{N}$ and $\mathcal{N}$ is a finite set, we conclude that $\sigma_k^q(h) = \sigma_k^{q,j_q^*}(h)$ for all $k \in \mathcal{N}$ as h is large enough, $\mathbb{P}_q$-a.s. This concludes the proof for part (i).

[4]Note that while the exposition in [26, Section 6.10] focuses on two-dimensional perturbed random walks, the same results there can be generalized to multi-dimensional cases as stated in [26, Remark 10.1, p. 208].

For part (ii), from (57), $S_d^q(h)$ is equal to

$$\inf\big\{t\in\mathbb{N}:\min_{0\le j\le Q,\ j\neq q} Y_t^k(q,j)\ge h\ \ \forall k\in\mathcal{L},\\ \text{for some } \mathcal{L}\subset[\mathcal{N}],\ |\mathcal{L}|=d\big\}.$$

Then from (95), $S_d^q(h)$ becomes

$$\begin{aligned}&\inf\big\{t\in\mathbb{N}: Y_t^k(q,j_q^*)\ge h\ \forall k\in\mathcal{L}, \text{for some}\,\mathcal{L}\subset[\mathcal{N}],|\mathcal{L}|=\ell\big\}\\ &=\inf\left\{t\in\mathbb{N}: Y_t^{(K-d+1)}(q,j_q^*)\ge h\right\}.\end{aligned}$$

Then as $h\to\infty$, $\mathbb{P}_q$-a.s. we have $S_d^q(h)=S_d^{q,j_q^*}(h)$.

APPENDIX D
PROOFS OF THEOREMS V.3 AND V.4

We first prove Theorem V.3.

Proof: Fix a $1\le d\le|\mathcal{N}|$. From [15, Theorem 3.1], we know that as $h\to\infty$,

$$\mathbb{E}_0^{q,\emptyset}[\sigma_{(d)}^{q,j_q^*}(h)]=\frac{h}{I^q}+D^q_{d:|\mathcal{N}|}\sqrt{h}(1+o(1)). \tag{96}$$

Since $\sigma_{(d)}^q(h)$ and $\sigma_{(d)}^{q,j_q^*}(h)$ are both nonnegative and non-decreasing in h, the monotone convergence theorem yields

$$\lim_{h\to\infty}\mathbb{E}_0^{q,\emptyset}[\sigma_{(d)}^q(h)]=\mathbb{E}_0^{q,\emptyset}\left[\lim_{h\to\infty}\sigma_{(d)}^q(h)\right]=\mathbb{E}_0^{q,\emptyset}\left[\lim_{h\to\infty}\sigma_{(d)}^{q,j_q^*}(h)\right]=\lim_{h\to\infty}\mathbb{E}_0^{q,\emptyset}\left[\sigma_{(d)}^{q,j_q^*}(h)\right], \tag{97}$$

where the second equality follows from Part (i) of Lemma V.1. The above two equations together show (64). Finally, plugging (64) into (59) and observing that $Q-1$ vanishes as $h\to\infty$ results in (65). ■

We then provide a proof to Theorem V.4.

Proof: Fix a $1\le d\le|\mathcal{N}|$. From [15, Theorem 3.2], it follows that as $h\to\infty$,

$$\mathbb{E}_0^{q,\emptyset}[S_d^{q,j_q^*}(h)]\le\frac{h}{I^q}+D^q_{d:|\mathcal{N}|}\sqrt{h}(1+o(1)). \tag{98}$$

Since $S_d^q(h)$ and $S_d^{q,j_q^*}(h)$ are both nonnegative and non-decreasing in h, the monotone convergence theorem yields

$$\lim_{h\to\infty}\mathbb{E}_0^{q,\emptyset}[S_d^q(h)]=\mathbb{E}_0^{q,\emptyset}\left[\lim_{h\to\infty}S_d^q(h)\right]=\mathbb{E}_0^{q,\emptyset}\left[\lim_{h\to\infty}S_d^{q,j_q^*}(h)\right]=\lim_{h\to\infty}\mathbb{E}_0^{q,\emptyset}\left[S_d^{q,j_q^*}(h)\right], \tag{99}$$

where the second equality follows from Part (ii) of Lemma V.1. The previous two equations together then give (66). Finally, plugging (66) into (61) results in (67). ■

APPENDIX E
PROOFS OF THEOREMS V.5 AND V.6

We first provide a proof to Theorem V.5.

Proof: We assume that whenever a tie happens, every competing hypothesis becomes acceptable simultaneously at the fusion center. This would only make the mean time to a false alarm or a false isolation smaller; hence, is valid for deriving lower bounds.

Let q_a be the actual hypothesis index and

$$q^* = \arg \min_{\hat{q} \in [Q] \setminus \{q_a\}} \mathbb{E}_0^{q_a,\emptyset}[\sigma^{\hat{q}}_{(d-M)}(h)]. \tag{100}$$

Recall that $\mathsf{P}_0^{q_a,\emptyset}$ is the probability measure when the change of type H_{q_a} happens at $\nu = 0$ and the compromised sensors are absent. With a fixed d, we have

$$\mathbb{E}_0^{q_a,\emptyset}[\tilde{\sigma}^{q^*}_{(d-M)}(h)] = \sum_{t=0}^{\infty} \mathsf{P}_0^{q_a,\emptyset}\left(\tilde{\sigma}^{q^*}_{(d-M)}(h) > t\right). \tag{101}$$

Now, let $\mathcal{N}_q(s) \triangleq \{k \in \mathcal{N} : \sigma_k^q(h) \le s\}$ be the set of honest sensor indices with $\sigma_k^q(h) \le s$. For every $t \in \mathbb{N}$, the event $\sigma^{q^*}_{(d-M)}(h) \le t$ happens if and only if the following is true,

$$\bigcup_{s=1}^{t} \left\{ \left(\bigcap_{q \neq q^*}^{Q} |\mathcal{N}_q(s)| < d - M \right) \bigcap \left\{ |\mathcal{N}_{q^*}(s)| \ge d - M \right\} \right\}. \tag{102}$$

Then, we have

$$\begin{aligned} \mathsf{P}_0^{q_a,\emptyset}\left(\sigma^{q^*}_{(d-M)}(h) \le t\right) &\le \mathsf{P}_0^{q_a,\emptyset}\left(\bigcup_{s=1}^{t} |\mathcal{N}_{q^*}(s)| \ge d - M \right) \\ &= \mathsf{P}_0^{q_a,\emptyset}\left(|\mathcal{N}_{q^*}(t)| \ge d - M \right) \end{aligned} \tag{103}$$

Also, we know that $|\mathcal{N}_{q^*}(t)| \ge d-M$ happens if and only if there are sensor indices $k_1, \ldots, k_{d-M} \in \mathcal{N}$ with $\sigma_{k_j}^{q^*}(h) \le t$ for $j \in [d-M]$. We further bound (103) by union bound as follows,

$$\begin{aligned} \mathsf{P}_0^{q_a,\emptyset}\left(|\mathcal{N}_{q^*}(t)| \ge d - M \right) &\le \sum_{k_1,\ldots,k_{d-M} \in \mathcal{N}} \mathsf{P}_0^{q_a,\emptyset}\left(\bigcap_{j=1}^{d-M} \sigma_{k_j}^{q^*}(h) \le t \right) \\ &= \sum_{k_1,\ldots,k_{d-M} \in \mathcal{N}} \prod_{j=1}^{d-M} \mathsf{P}_0^{q_a,\emptyset}\left(\sigma_{k_j}^{q^*}(h) \le t \right) \\ &= \begin{pmatrix} |\mathcal{N}| \\ d - M \end{pmatrix} \left(\mathsf{P}_0^{q_a,\emptyset}\left(\sigma_1^{q^*}(h) \le t \right) \right)^{d-M}. \end{aligned} \tag{104}$$

where the first and second equalities are from the independent and identical distributions of different sensor observations, respectively.

Note that from the definition of matrix CUSUM in (56), it follows that

$$\begin{aligned}
\mathsf{P}_0^{q_a,\emptyset}\left(\sigma_1^{q^*}(h)\le t\right) &= \mathsf{P}_0^{q_a,\emptyset}\left(\bigcup_{s=1}^{t}\left\{\left\{Y^1_{s,q^*}\ge h, q^*=\arg\max_{q'\in[Q]} Y^1_{t,q'}\right\}\bigcap_{s'=1}^{s-1}\max_{q'\in[Q]} Y^1_{s',q'}<h\right\}\right)\\
&\le \sum_{s=1}^{t}\mathsf{P}_0^{q_a,\emptyset}\left(\left\{Y^1_{s,q^*}\ge h, q^*=\arg\max_{q'\in[Q]} Y^1_{t,q'}\right\}\bigcap_{s'=1}^{s-1}\max_{q'\in[Q]} Y^1_{s',q'}<h\right)\\
&\le \sum_{s=1}^{t}\mathsf{P}_0^{q_a,\emptyset}\left(Y^1_{s,q^*}\ge h\right)=\sum_{s=1}^{t}\mathsf{P}_0^{q_a,\emptyset}\left(\bigcap_{0\le j\le Q, j\neq q^*} Y^1_s(q^*,j)\ge h\right)\\
&\le \sum_{s=1}^{t}\mathsf{P}_0^{q_a,\emptyset}\left(Y^1_s(q^*,0)\ge h\right). \qquad (105)
\end{aligned}$$

Now, we know from [19] that $\mathsf{P}_0^{q_a,\emptyset}\left(Y^1_s(q^*,0)\ge h\right)\le e^{-h}$. Thus, from (104) and (105), we have

$$\mathsf{P}_0^{q_a,\emptyset}\left(\sigma^{q^*}_{(d-M)}(h)\le t\right)\le \binom{|\mathcal{N}|}{d-M} t^{d-M}e^{-(d-M)h}. \qquad (106)$$

Plugging (106) into (101) results in

$$\begin{aligned}
\mathbb{E}_0^{q_a,\emptyset}[\sigma^{q^*}_{(d-M)}(h)] &> \sum_{t=0}^{\infty}\left(1-\binom{|\mathcal{N}|}{d-M} t^{d-M}e^{-(d-M)h}\right)^+\\
&\ge \int_0^{\infty}\left(1-\binom{|\mathcal{N}|}{d-M} t^{d-M}e^{-(d-M)h}\right)^+ dt, \qquad (107)
\end{aligned}$$

where the second inequality comes from the non-increasing property in t of

$$1-\binom{|\mathcal{N}|}{d-M} t^{d-M}e^{-(d-M)h}. \qquad (108)$$

Finally, noticing that the lower bound in (107) is not a function of the actual hypothesis q_a concludes the proof of $\mathcal{A}[\tau^m_{(d)}(h)]$. ■

In what follows, we present a proof to Theorem V.6.

Proof: Again, let q_a be the actual hypothesis and let

$$q^*=\arg\min_{\hat{q}\in[Q]\setminus\{q_a\}}\mathbb{E}_0^{q_a,\emptyset}[S^{\hat{q}}_{d-M}(h)]. \qquad (109)$$

With a fixed d, we have

$$\mathbb{E}_0^{q_a,\emptyset}[S^{\,q^*}_{(d-M)}(h)]=\sum_{t=0}^{\infty}\mathsf{P}_0^{q_a,\emptyset}\left(S^{\,q^*}_{(d-M)}(h)>t\right). \qquad (110)$$

Note that for every $t \in \mathbb{N}$, the event $S^{q^*}_{(d-M)}(h) \leq t$ happens if and only if the following event is true,

$$\bigcup_{s=1}^{t} \left\{ \left(\bigcap_{q \neq q^*}^{Q} Y_{s,q}^{(K-(d-M)+1)} < h \right) \bigcap Y_{s,q^*}^{(K-(d-M)+1)} \geq h \right\}. \tag{111}$$

Then, we have

$$\begin{aligned} \mathsf{P}_0^{q_a,\emptyset} \left(S^{q^*}_{(d-M)}(h) \leq t \right) &\leq \mathsf{P}_0^{q_a,\emptyset} \left(\bigcup_{s=1}^{t} Y_{s,q^*}^{(K-(d-M)+1)} \geq h \right) \\ &\leq \sum_{s=1}^{t} \mathsf{P}_0^{q_a,\emptyset} \left(Y_{s,q^*}^{(K-(d-M)+1)} \geq h \right) \end{aligned} \tag{112}$$

Also, we know that event $Y_{s,q^*}^{(K-(d-M)+1)} \geq h$ happens if and only if there are $d-M$ sensors with indexes $k_1, \ldots, k_{d-M} \in \mathcal{N}$ which have local decisions q^* at time index s. Therefore,

$$\begin{aligned} \mathsf{P}_0^{q_a,\emptyset} \left(Y_{s,q^*}^{(K-(d-M)+1)} \geq h \right) &= \sum_{k_1,\ldots,k_{d-M} \in \mathcal{N}} \prod_{j=1}^{d-M} \mathsf{P}_{q_a}^{\emptyset} \left(Y_{s,q^*}^{k_j} \geq h \right) \\ &= \begin{pmatrix} |\mathcal{N}| \\ d-M \end{pmatrix} \left(\mathsf{P}_0^{q_a,\emptyset} \left(Y_{s,q^*}^{1} \geq h \right) \right)^{d-M} \end{aligned} \tag{113}$$

where the first and second equalities are from the independent and identical distributions of different sensor observations, respectively. Now as in (105), it follows that

$$\mathsf{P}_0^{q_a,\emptyset} \left(Y_{s,q^*}^{1} \geq h \right) \leq \mathsf{P}_0^{q_a,\emptyset} \left(Y_{s}^{1}(q^*, 0) \geq h \right) \leq e^{-h} \tag{114}$$

Thus, from (112)-(114),

$$\mathsf{P}_0^{q_a,\emptyset} \left(S^{q^*}_{(d-M)}(h) \leq t \right) \leq \begin{pmatrix} |\mathcal{N}| \\ d-M \end{pmatrix} t e^{-(d-M)h}. \tag{115}$$

Plugging (115) into (110) and noticing that the bound in (115) is independent of q_a completes the proof for the lower bound on $\mathcal{A}(\tau^s_{(d)}(h))$. ■

REFERENCES

[1] Y.-C. Huang, S.-C. Lin, and Y.-J. Huang, "A tight converse to the asymptotic performance of Byzantine distributed sequential change detection," in *Proc. IEEE ISIT*, 2019.

[2] Y.-J. Huang, S.-C. Lin, and Y.-C. Huang, "On Byzantine distributed sequential change detection with multiple hypotheses," in *Proc. IEEE ISIT*, 2019.

[3] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, Jun. 1954.

[4] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, no. 6, p. 18971908, Dec. 1971.

[5] G. V. Moustakides, "Optimal stopping times for detecting changes in distribution," *Ann. Statist.*, vol. 14, no. 4, p. 13791387, 1986.

[6] I. V. Nikiforov, "A generalized change detection problem," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 171–187, Jan. 1995.

[7] T. Oskiper and H. V. Poor, "Online activity detection in a multiuser environment using the matrix CUSUM algorithm," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 477–493, 2002.

[8] V. V. Veeravalli and T. Banerjee, "Quickest change detection," in *Academic Press Library in Signal Processing*, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, vol. 6, ch. 6, pp. 209–255.

[9] J. Wurm, Y. Jin, Y. Liu, S. Hu, K. Heffner, F. Rahman, and M. Tehranipoor, "Introduction to cyber-physical system security: A cross-layer perspective," *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 3, no. 3, pp. 215–227, July-Sept. 2017.

[10] Y.-P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A primer on 3GPP narrowband internet of things," *IEEE Comm. Mag.*, vol. 55, no. 3, pp. 117–123, March 2017.

[11] "Operations and maintenance saving from advanced metering infrastructure - initial results," Technical Report, U.S. Dept. Energy, Office Elect. Del. Energy Rel., Dec. 2012. [Online]. Available: http://energy.gov/sites/prod/files/AMI%5FSavings%5FDec2012Final.pdf

[12] E. Bayraktar and L. Lai, "Byzantine fault tolerant distributed quickest change detection," *SIAM J. Control Optim.*, vol. 53, no. 2, pp. 575–591, 2015.

[13] G. Fellouris, E. Bayraktar, and L. Lai, "Efficient Byzantine sequential change detection," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3346–3360, May 2018.

[14] W.-N. Chen and I.-H. Wang, "Anonymous heterogeneous distributed detection: Optimal decision rules, error exponents, and the price of anonymity," *IEEE Trans. Inf. Theory*, 2019, to appear.

[15] S. Banerjee and G. Fellouris, "Decentralized sequential change detection with ordered CUSUMs," in *Proc. IEEE ISIT*, 2016, pp. 36–40.

[16] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1999, vol. 23.

[17] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2009.

[18] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[19] Y. Mei, "Information bounds and quickest change detection in decentralized decision systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2669–2681, Jul. 2005.

[20] A. G. Tartakovsky, "Asymptotic performance of a multichart cusum test under false alarm probability constraint," in *Proc. 44th IEEE Conf. Decision Control*, 2005, pp. 320–325.

[21] A. G. Tartakovsky and V. V. Veeravalli, "Change-point detection in multichannel and distributed systems," in *Applied Sequential Methodologies: Real-World Examples with Data Analysis*, N. Mukhopadhyay, S. Datta, and S. Chattopadhyay, Eds. Marcel Dekker, 2004, vol. 173, ch. 17, pp. 339–370.

[22] Y. Mei, "Efficient scalable schemes for monitoring a large number of data streams," *Biometrika*, vol. 97, pp. 419–433, Jun. 2010.

[23] Y. Xie and D. Siegmund, "Sequential multi-sensor change-point detection," *Ann. Statist.*, vol. 41, no. 2, p. 670692, Dec. 2013.

[24] P. Billingsley, *Probability and measure*, 3rd ed. Wiley-Interscience., 2008.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[26] A. Gut, *Stopped random walks*, 2nd ed., ser. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2009, limit theorems and applications.