IEEE Control Systems Letters paper presented at
2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

A Framework for Bayesian Quickest Change Detection in General Dependent Stochastic Processes

Jasmin James[1], Jason J. Ford[2], and Timothy L. Molloy[3]

***Abstract*— In this paper we present a novel framework for quickly detecting a change in a general dependent stochastic process. We propose that any general dependent Bayesian quickest change detection (QCD) problem can be converted into a hidden Markov model (HMM) QCD problem, provided that a suitable state process can be constructed. The optimal rule for HMM QCD is then a simple threshold test on the posterior probability of a change. We investigate case studies that can be considered structured generalisations of Bayesian HMM QCD problems including: quickly detecting changes in statistically periodic processes and quickest detection of a moving target in a sensor network. Using our framework we pose and establish the optimal rules for these case studies. We also illustrate the performance of our optimal rule on real air traffic data to verify its simplicity and effectiveness in detecting changes.**

I. Introduction

Quickest change detection (QCD) problems are concerned with the quickest (on-line) detection of a persistent change in the mode of a process, most commonly between a "normal" and an "abnormal" mode. It is desirable to detect this change quickly (as soon as possible) subject to a constraint on the occurrence of false alarms. Information about the current mode that the process is in (normal or abnormal) is extracted from a series of quantitative observations (i.e., measurements corrupted by noise) [1]. When the observations suggest that the process is in the normal mode it is desirable to let the process continue. However, if the observations suggest that the process has changed then the aim is to detect this change quickly. As such, with each new observation, there is the decision of whether the process should continue or whether it should stop and a detection declared.

QCD problems arise in a wide variety of applications including quality control [2], and fault detection [2], [3]. There are various formulations of this problem that differ in assumptions on the point of change and optimality criteria. Early theoretical formulations for quickest change detection were developed by Shiryaev who assumed that the change point is a random variable with a known geometric distribution and the observations are independent and identically distributed (i.i.d.) [4]. Shiryaev established an optimal (stopping) rule that compares the posterior probability of a change with a threshold. Shiryaev's formulation has since been extended to encompass non-geometrically distributed change-times [5], [6] and dependent data (i.e., non-i.i.d. observations) [5], [7]–[9]. These generalisations (as well as Shiryaev's original formulation) are known as Bayesian formulations of QCD since the distribution of the change-time is used as a prior for detecting its occurrence.

Despite various (generalised) Bayesian QCD formulations appearing in the literature [6]–[8], establishing optimal detection rules for dependent data and arbitrary change-time distributions has remained a challenging problem and led to the development of weaker asymptotic optimality results that hold as the probability of false alarms vanishes. For the general non-i.i.d. case, [5] was able to show that Shiryaev's rule is asymptotically optimal. For hidden Markov models (HMMs), [9] was able to show that Shiryaev's rule is asymptotically optimal (under several regularity conditions). The asymptotic results of [5], however, are developed under conditions on the tail probabilities of the change-time distributions (and that the distributions of the change-times are independent of the state of the pre-change process) whilst [9] retains the strong geometric change-time conditions of Shirayav's original formulation. Recently [10] considered QCD for HMMs in a Bayesian setting and proved that Shiryaev's rule is an exact optimal solution.

In this paper we present a framework for quickly detecting a change in a general dependent stochastic process, proposing that any general dependent Bayesian QCD problem can be converted into an HMM QCD problem (which the optimal solution was established in [10]), when a suitable state process can be constructed. Specifically the key contributions of this paper are:

- A framework for posing Bayesian quickest change detection in general dependent stochastic processes as Bayesian quickest change detection in HMMs.
- An algorithmic process for formulating and solving Bayesian QCD problems as HMM QCD problems.
- Application of the framework and algorithmic process to multiple case studies and in a real-world example.

The rest of this paper is organised as follows: In Section II we formulate the Bayesian QCD problem, and in Section III we present our framework. In Section IV we present our case studies and real world examples. Finally, we provide concluding remarks in Section V.

[1]J. James is with the School of Mechanical and Mining Engineering, University of Queensland, St Lucia, QLD 4072, jasmin.martin@uq.edu.au
[2] J. J. Ford is with the School of Electrical Engineering and Robotics, Queensland University of Technology, 2 George St, Brisbane QLD, 4000 Australia. j2.ford@qut.edu.au. The work of Jason J. Ford was supported by the Queensland University of Technology's Centre for Robotics.
[3] T. Molloy is with the School of Engineering, Australian National University (ANU), Canberra, ACT 2601, Australia timothy.molloy@anu.edu.au.

Copyright ©2024 IEEE

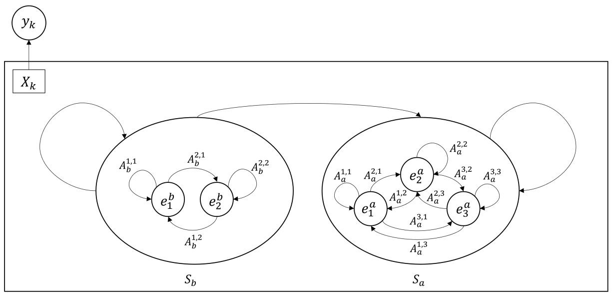

Fig. 1: An example of a HMM QCD problem where $S_b = \{e^b_1, e^b_2\}$ and $S_a = \{e^a_1, e^a_2, e^a_3\}$. The arrows represent the transitions of the HMMs.

II. Problem Formulation

For $k > 0$, let $y_k \in \mathcal{Y}$ be a sequence of (possibly dependent) random variables taking values in the set $\mathcal{Y} \subseteq \mathbb{R}^N$. We assume that the sequence y_k potentially contains an unknown (possibly random) change point $\nu \geq 1$ in the sense that before the change point the conditional density of y_k given $y_{[1,k-1]} \triangleq \{y_1, y_2, \ldots, y_{k-1}\}$ is $f^0_k(\cdot|y_{[1,k-1]})$ for $k < \nu$ and after the change point the conditional density is $f^1_k(\cdot|y_{[1,k-1]})$ for $k \geq \nu$. Under this change description, the joint probability density function of the general dependent process $y_{[1,k]}$ is given by

$$p_\nu(y_{[1,k]}) \triangleq \prod_{i=1}^{\nu-1} f^0_i(y_i|y_{[1,i-1]}) \prod_{j=\nu}^{k} f^1_j(y_j|y_{[1,j-1]}) \qquad (1)$$

where we define the initial densities $f^0_1(y_1|y_{[1,0]}) \triangleq f^0_1(y_1)$ and $f^1_1(y_1|y_{[1,0]}) \triangleq f^1_1(y_1)$, and if $k < \nu$ we define $\prod_{j=\nu}^{k} f^1_j(y_j|y_{[1,j-1]}) \triangleq 1$.

Significantly, the general dependent process model (1) can be specialised to describe abrupt changes in a range of important model classes in systems and control, including state-space models and Markov chains (see for example [5], [11]).

We observe y_k sequentially with the aim of detecting a change as soon as possible after the change time ν subject to a false alarm rate constraint.

A. Bayesian QCD Problem Formulation

Before we formally state our Bayesian HMM QCD problem, let us first introduce a probability measure space. Let $\mathcal{F}_k = \sigma(y_{[1,k]})$ denote the filtration generated by $y_{[1,k]}$. We will assume the existence of a probability space $(\Omega, \mathcal{F}, P_\nu)$ where we consider the set Ω consisting of all infinite sequences $\omega \triangleq (y_{[1,\infty]})$. Since Ω is separable and a complete metric space it can be endowed with a Borel σ-algebra $\mathcal{F} = \cup_{k=1}^{\infty} \mathcal{F}_k$ with the convention that $\mathcal{F}_0 = \{0, \Omega\}$, and P_ν is the probability measure constructed via applying Kolmogorov's extension theorem with the probability density function p_ν.

Under the Bayesian QCD formulation we consider the change time $\nu \geq 1$ to be an unknown random variable with prior distribution $\pi_k \triangleq P(\nu = k)$ for $k \geq 1$. This allows us to construct a new averaged measure $P_\pi(G) = \sum_{k=1}^{\infty} \pi_k P_k(G)$ for all $G \in \mathcal{F}$ and we let E_π denote the corresponding expectation operation. In this presentation, the geometric prior $\pi_k = (1-\rho)^{k-1}\rho$ with $\rho \in (0,1)$ as introduced by Shiryaev [4].

The classic formulation of Bayesian QCD seeks to find a stopping time $\tau \geq 1$ that solves the following constrained optimisation problem

$$\inf_{\tau \in T(\alpha)} E_\pi \left[(\tau - \nu)^+ \right] \qquad (2)$$

where $(\tau - \nu)^+ \triangleq \max(0, \tau - \nu)$ and $T(\alpha) \triangleq \{\tau : P_\pi(\tau < \nu) \leq \alpha\}$ denotes the set of stopping times satisfying a given probability of false alarm constraint $\alpha \in (0, 1-\rho)$.

Alternatively, the Bayesian relaxation is to design a stopping time $\tau \geq 1$ with respect to the filtration generated by $y_{[1,k]}$ that minimises the following cost

$$J(\tau) \triangleq cE_\pi \left[(\tau - \nu)^+ \right] + P_\pi(\tau < \nu), \qquad (3)$$

where c is the penalty at each time step before declaring an alert at τ. If c can be found such that the solution to (3) achieves the probability of false alarm constraint constraints α with equality (that is, $P_\pi(\tau < \nu) = \alpha$) then a standard saddle-point argument [12, p. 220-1] establishes the duality gap is zero and the solution to (3) is also the solution to (2).

III. Proposed Framework

We now present our proposed framework for detecting a change in a general dependant process. Our framework is applicable to the broad class of general dependent stochastic processes with an underlying state process which can be modelled as a Markov chain (either exactly or estimated) and conditioned on the underlying states, the observations are conditionally independent.

A. State and Observation Process

To present our framework, consider two finite state spaces *before-change* $S_b \triangleq \{e^b_1, \ldots, e^b_{N_b}\}$ and *after-change* $S_a \triangleq \{e^a_1, \ldots, e^a_{N_a}\}$ where $e^b_i \in \mathbb{R}^{N_b}$ and $e^a_i \in \mathbb{R}^{N_a}$ are indicator vectors with 1 in the ith element and zeros elsewhere, and $N_b \geq 1$ and $N_a \geq 1$ are two integers. Consider also a process X_k for $k \geq 0$ that is able to randomly transition between states in the space of the current stage (within S_b or S_a) or able to transition to a state in the space of the next stage (e.g., from S_b to S_a) as seen in Figure 1. Here, X_k starts in the before stage; that is, $X_0 \in S_b$ with probability $p(X_0)$.

For $k < \nu$, $X_k \in S_b$ can be modelled a first-order time-homogeneous Markov chain described by the transition probabilities $A^{i,j}_b \triangleq P(X_{k+1} = e^b_i | X_{k+1} \in S_b, X_k = e^b_j)$ for $1 \leq i, j \leq N_b$. At some unknown time $k = \nu$ the process X_k transitions between stages in the sense $X_{\nu-1} \in S_b$ and $X_\nu \in S_a$ according to state change probabilities $A^{i,j}_\nu \triangleq P(X_{k+1} = e^a_i | X_{k+1} \in S_a, X_k = e^b_j)$ for $1 \leq i \leq N_a$ and $1 \leq j \leq N_b$. For $k > \nu$, $X_k \in S_a$ can be modelled as a first-order time-homogeneous Markov chain described by the transition probabilities $A_a^{i,j} \triangleq P(X_{k+1} = e_i^a | X_{k+1} \in S_a, X_k = e_j^a)$, for $1 \leq i,j \leq N_a$.

Working under the average measure, noting the change time $\nu = \inf\{k \geq 1 : X_k \in S_a\}$, we now consider the situation where process (1) can be modelled as

$$p_\pi(y_{[1,k]}) = \sum_{X_{[0,k]}} p_\pi(X_{[0,k]}) \prod_{i=1}^{\nu-1} b_b(y_i, \zeta(X_i)) \prod_{j=\nu}^{k} b_a(y_j, \zeta(X_j)). \quad (4)$$

where $\zeta(e_i) \triangleq i$ returns the index of the non-zero element of an indicator vector e_i^b or e_i^a and $b_b(y_k, i) \triangleq P(y_k | X_k = e_i^b)$ for $1 \leq i \leq N_b$ and $k < \nu$, and $b_a(y_k, i) \triangleq P(y_k | X_k = e_i^a)$ for $1 \leq i \leq N_a$ and $k \geq \nu$, with

$$p_\pi(X_{[0,k]}) = p(X_0) A_\nu^{\zeta(X_\nu),\zeta(X_{\nu-1})} \prod_{i=1}^{\nu-1} A_b^{\zeta(X_i),\zeta(X_{i-1})} \times \prod_{j=\nu+1}^{k} A_a^{\zeta(X_j),\zeta(X_{j-1})}$$

via the definition of X_k as a time-homogeneous Markov chain, with initial condition $p(X_0)$, using $X_{[0,k]} \triangleq \{X_0, \ldots, X_k\}$.

Then, with a slight adjustment of the construction of the sample space Ω to include infinite sequences $X_{[0,\infty]}$, we propose converting from (1) to the (4) process model i.e., converting the dynamics of a system to an HMM. This involves either uncovering exactly or via system identification approaches (such as in [13] but many exist) the underlying state and measurement processes. When this construction is possible, the posed QCD problem (3) can be embedded with an augmented HMM representation as shown in the following section.

B. Augmented HMM

As some machinery for our framework, we define a new augmented state process $Z_k \in S$ where $S \triangleq \{e_1, \ldots, e_N\}$ where $e_i \in \mathbb{R}^N$ (are indicator vectors with 1 in the ith element and zero elsewhere) and $N = N_b + N_a$. This augmented state process combines the information of X_k and ν as follows. For $k < \nu$, $Z_k \in S$ is defined as

$$Z_k \triangleq \begin{bmatrix} X_k \\ \mathbf{0}_a \end{bmatrix},$$

and for $k \geq \nu$ as

$$Z_k \triangleq \begin{bmatrix} \mathbf{0}_b \\ X_k \end{bmatrix}.$$

where $\mathbf{0}_b$ and $\mathbf{0}_a$ are the zero vectors of size N_b and N_a, respectively. Following [10] Lemma 1 we note that the augmented process Z_k is a first-order time-homogeneous Markov chain with transition probabilities $A^{i,j} \triangleq P_\pi(Z_{k+1} = e_i | Z_k = e_j)$ that can be written as

$$A = \begin{bmatrix} (1-\rho)A_b & \mathbf{0}_{b\times a} \\ \rho A_\nu & A_a \end{bmatrix} \quad (5)$$

where $\mathbf{0}_{b\times a}$ is a $N_b \times N_a$ matrix of all zeros. Moreover with measurement matrix $B^{j,j}(y_k) \triangleq P_\pi(y_k | Z_k = e_j)$ of

$$B(y_k) = \text{diag}(b_b(y_k, 1), \ldots, b_b(y_k, N_b), b_a(y_k, 1), \ldots, b_a(y_k, N_a)) \quad (6)$$

then (Z_k, y_k) are the state and observation processes of a hidden Markov model with transition matrix A and measurement matrix B.

We let $\hat{Z}_k^i \triangleq P_\pi(Z_k = e_i | y_{[1,k]})$ denote the posterior probabilities of being in each of the states of Z_k with initial conditions $\hat{Z}_0$, where $\hat{Z}_0^i = P(Z_0 = e_i^b)$ for $i \in \{1, \ldots, N_b\}$ and $\hat{Z}_0^i = 0$ elsewhere. We can define the operation $M(Z) \triangleq \sum_{i=1}^{N_b} Z^i$ and no change posterior $\hat{M}_k^1 \triangleq M(\hat{Z}_k)$.

The optimal Bayesian QCD rule is then a simple threshold test

$$\tau^* = \inf\{k \geq 1 : \hat{M}_k^1 \leq h\} \quad (7)$$

presented in [10] in Theorem 1. We note that detection range and false alarm performance vary with the choice of the threshold parameter h. In practice, detection thresholds are often experimentally selected for a particular application or adaptively selected such as proposed in [14] on the basis of scene difficulty for vision based aircraft detection. The test statistic $\hat{M}_k^1$ can be calculated via the HMM filter recursion [15] of our augmented HMM as

$$\hat{Z}_k = N_k B(y_k) A \hat{Z}_{k-1} \quad (8)$$

with scalar normalisation $N_k \triangleq \langle \underline{1}, B(y_k) A \hat{Z}_{k-1} \rangle^{-1}$ where $\underline{1}$ is the $N \times 1$ vector of all ones. Algorithm 1 describes how to implement our optimal stopping rule (7) on a sequence of measurements $\{y_1, \ldots, y_k\}$.

Algorithm 1 Implementation of the Optimal Rule (7)

Require: $\{P(X_0 = e_1^i)\}$, $\{b(\cdot,\cdot)\}$, A_b, A_a, A_ν & ρ (found either exactly or through system identification techniques)
Require: $h \in (0,1)$ (selected for desired balance between detection delay and probability of false alarms)
  $\hat{X}_0^i \leftarrow P(X_0 = e_1^i)$ for $i \in \{1, \ldots, N(1)\}$
  $\hat{Z}_0 \leftarrow [\hat{X}_0', 0_{1\times N_b}]'$
  $A \leftarrow A_b, A_a, A_\nu$ & ρ using (5)
  repeat when y_k arrives
    $B(y_k) \leftarrow y_k$ using (6)
    $\hat{Z}_k \leftarrow B(y_k)$ & $\hat{Z}_{k-1}$ using (8)
    $\hat{M}_k^1 \leftarrow \sum_{i=1}^{N_b} \hat{Z}_k^i$
  until $\hat{M}_k^1 \leq h$ **return** $\tau = k$ using (7)

Simply, our framework involves the following:

- Seek to find or estimate an underlying state and measurement process that, conditioned on the underlying states, the observations are conditionally independent.
- Build up an augmented HMM QCD representation to embed the process.
- Apply existing HMM QCD optimal results from [10] which is a threshold test on the posterior.

We assume that the pre-change and post-change distributions are known exactly. However, our framework is still of

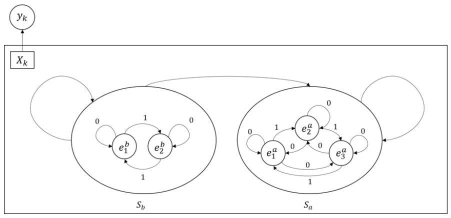

Fig. 2: An example of QCD in periodic processes case as a HMM detection problems, where $T_1 = 2$ and $T_2 = 3$. The probability of transitioning between the two HMMs is ϱ and the probabilities of transitioning between the states is represented in A_ν.

value when this assumption does not hold since both the HMM filter and Bayesian detector are known to degrade gracefully (in terms of performance) in the presence of model uncertainty and mismatch, as shown in [16] and [17]. For example, it opens the possibility of employing asymptotically minimax robust Bayesian QCD solutions when there is uncertainty about the parameter of the post-change conditional densities (cf. [17]).

IV. Case Studies

In this section we apply our proposed framework and elegantly establish optimal rules for important quickest detection problems by showing that they are structured versions of the general HMM QCD problem solved in [10]. We also illustrate the performance of our optimal rule on real air traffic data to verify its simplicity and effectiveness in detecting changes.

A. Case Study 1: QCD in Statistically Periodic Processes.

Periodic behaviour is present in a wide variety of applications including: power grid monitoring where the usage of power varies with low usage during the nighttime hours and high usage during the daytime hours [18], traffic monitoring where the intensity of traffic has a periodic behaviour over days and weeks [19], neural spike patterns in brain-computer interface studies, social networks, and many more [19]. This problem of QCD in a statistically periodic process is considered in [19] where they established an algorithm that can asymptotically minimise the average detection delay subject to a constraint on the probability of false alarm. More recently this problem is considered in [20] where they were able to show that a stopping rule based on a periodic sequence of thresholds is exactly optimal for their associated cost under some strict constraints including the pre- and post-change cycles are the same length and synchronised. A contribution of this paper is establishing that an optimal rule for QCD of statistically periodic processes is a simple threshold test (improving the results of [19] which only determined asymptotic optimality, and without the cycle constraints of [20]).

1) Optimal Result: Here we follow the notation of [19] and set up our model of an independent and periodically identically distributed (i.p.i.d) stochastic process. Consider a random variable y_k that is independent and has density f_k for $k \geq 1$, with period T_1 such that $f_{k+T_1} = f_k \; \forall k \geq 1$. We aim to detect a deviation from this periodic behaviour.

Let us consider another set of densities g_k for $k \geq 1$, with period T_2 such that $g_{k+T_2} = g_k \; \forall k \geq 1$. We assume that at an unknown changes time ν, the process y_k changes from being governed by periodic statistical properties $(f_1, \dots, f_{T_1})$ to being governed by the new set of densities $(g_1, \dots g_{T_2})$, where the probability that the modified behaviour begins at a specific location i in the cycle of the densities $(g_1, \dots g_{T_2})$ is given by the probability mass function $p_g(i)$ for $i \in \{1, \dots, T_2\}$. It is assumed the prior distribution of the change event at time k is described by a geometric prior $(1-\rho)^{k-1}\rho$, for some $\rho \in (0, 1)$, as introduced by Shiryaev in the context of the QCD problem [4].

Note that the densities $(g_1, \dots g_{T_2})$ need not be all different from the set of pre-change densities $(f_1, \dots, f_{T_1})$, but we assume that there exists at least one such that $g_i \neq f_i$ for some i. Our goal is to quickly detect when $y_k \sim g_k$ by seeking to design a stopping time that satisfies our desire to detect this change as early as possible subject to a constraint on false alarms.

We now show that the problem of quickly detecting a change in a periodic process can be considered a structured version of the HMM QCD problem and is solved by the optimal rule established in Theorem 1 [10].

Theorem 1. For QCD in statistically periodic processes described in Section IV-A under the cost criterion (3), there is an optimal rule with stopping time τ^*, and threshold point $h \geq 0$ given by

$$\tau^* = \inf\{k \geq 1 : \hat{M}_k^1 \leq h\}. \quad (9)$$

Proof. Let us define our set of periodic spaces $S_b = \{e_1^b, \dots, e_{N_b}^b\}$ and $S_a = \{e_1^a, \dots, e_{N_a}^a\}$ where $N_b = T_1$ and $N_a = T_2$. For $k > 0$, we consider a random process X_k whose statistical properties change at some (unknown) time $\nu > 1$, in the sense that for $0 < k < \nu$, $X_k \in S_b$, whilst for $k \geq \nu$, $X_k \in S_a$. As seen in Figure 2, we model transitions of S_b as a first-order time-homogeneous Markov chain described by the $A_b^{i,j}$. Note that due to the constrained nature of the periodic problem $A_b^{i+1,i} = 1$ for $1 < i < N_b$ and $A_b^{1,N_b} = 1$, and all other elements 0. For $X_k \in S_a$ transitions between elements of S_a can also be modelled as a first-order time-homogeneous Markov chain described by the transition probability $A_a^{i,j}$, for $1 \leq i, j \leq N_a$. Once again, due to the constrained nature of the periodic problem $A_a^{i+1,i} = 1$ for $1 < i < N_a$ and $A_a^{1,N_a} = 1$.

At time $k = \nu$, with probability ρ, the process transitions between sets according to state change probabilities $A_\nu^{i,j} = p_g(i)$ for $1 \leq i \leq N_a$, and $1 \leq j \leq N_b$.

Finally, for each $k > 0$, X_k is indirectly observed via measurements y_k generated by conditional observation densities $b_b(y_k, i) = f_i$ for $1 \leq i \leq T_1$ and $k < \nu$ and $b_a(y_k, i) = g_i$ for $1 \leq i \leq T_2$ and $k \geq \nu$. This periodic process QCD problem can be represented as an augmented HMM as presented in Section III-B with constrained transitions, and has change event having a state independent prior

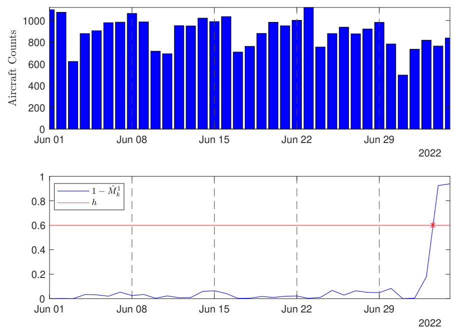

Fig. 3:
(Top) the air traffic counts for days of the week starting the 1st of June 2022. The black dashed lines signify a new week. (Bottom) the test statistic of our optimal rule (7) (inverted for presentation) which detects that a change has occurred on the 3rd of July 2022 with an experimentally selected threshold of $h = 0.6$.

$\pi_k = (1-\rho)^{k-1}\rho$ for some $\rho \in (0,1)$. Applying Theorem 1 [10] thus gives the optimal stopping rule as (9). □

This exactly optimal result improves on the asymptotic results of [19] without the strict cycle constraints of [20].

2) Example on Real Data: We now apply our optimal rule to real data to demonstrate its effectiveness in detecting changes. We applied our data to flight counts of daily air traffic in Brisbane, Queensland, Australia (freely accessible via https://opensky-network.org/). We estimated the pre-change distributions $N_b = 7$ (from e_1^b = Monday to e_7^b = Sunday) as Gaussian using aircraft counts from training data collected between 6th July - 31st August 2022. For the post-change distributions we set $N_a = 2$ possible post-changes states that were 50 counts above and below the mean pre-change aircraft count. We set the state transitions as

$$A_b = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, A_a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and}$$

$$A_\nu = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}.$$

We calculated the test statistic via the HMM filter (8).

Figure 3 (top) shows the air traffic counts for days of the week starting the 1st of June 2022 to the 5th of July 2022. The counts exhibit periodic behaviour over several weeks (weeks are illustrated by the black dotted lines). There is a change in the periodic behaviour from the 3rd of July 2022, where the aircraft count appears to trend lower. We believe this corresponds to the July floods [21] which would have impacted flights and runways. Figure 3 (bottom) illustrates our optimal rule (7) in blue (note we have inverted the test statistic for presentation), and a threshold of $h = 0.6$ (experimentally selected). Our optimal rule effectively detects this change in aircraft counts on the 3rd of July 2022 (counts appeared to revert to normal after the 6th of July, hence their use as pre-change training data).

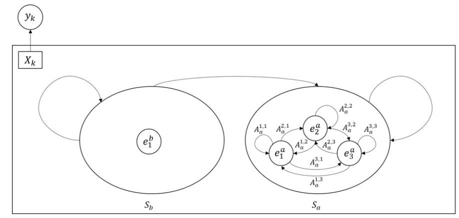

Fig. 4: An example of quickest detection of a moving target in a sensor network, where $L = 3$.

B. Case Study 2: Quickest detection of a Moving Target in a Sensor Network

The problem of quickly detecting a moving target in a sensor network is significant in various important applications including intrusion detection in computer networks and security systems [22]. In [22] it is established that a windowed test based on a generalised likelihood ratio approach is asymptotically optimal for quickest detection of a moving target in a sensor network. A contribution of this paper is establishing that the optimal Bayesian rule for QCD of a moving target in a sensor network is a simple threshold test (a non-Bayesian version of this problem is considered in [22], a Bayesian approach may be advantageous in applications with prior information or mean time to failure constraints).

1) Optimal Result: Following [22], we formulate the moving target QCD problem. Consider a sensor network of L nodes denoted by $\mathcal{L} = \{1, \ldots, L\}$. Let $y_k = [y_k^1, \ldots, y_k^L]'$ denote the observation vector at time k where y_k^ℓ denotes the measurement obtained by sensor $\ell \in \mathcal{L}$ at time k. For simplicity of presentation, we consider y_k^ℓ as scalar but these results can be generalised for more complex measurements. The samples obtained by sensor ℓ are i.i.d. according to a $f_b^\ell(\cdot)$ for all $\ell \in \mathcal{L}$ and are independent across different sensors. For $k < \nu$, the joint density of $y_{[1,k]}$ is thus

$$f_b(y_1, \ldots, y_k) = \prod_{j=1}^{k} \prod_{i=1}^{L} f_b^i(y_j^i) \quad (10)$$

At time ν a target appears and, at each time instant $k \geq \nu$, one of the sensors is affected by the target which we denote by $S_k \in \mathcal{L}$. It is assumed the prior distribution of the target's appearance is geometric. At the change time, ν, the probability that the target appears at a specific node i is given by the probability mass function $p_{\mathcal{L}}(i)$. Conditioning on S_k the joint distribution of y_k is given by

$$f^\ell(y_k) \triangleq \left(\prod_{i \neq \ell} f_b^i(y_k^i) \right) f_a^\ell(y_k^\ell) \quad (11)$$

where $f_a^\ell(\cdot)$ denotes the density of the affected sensor ℓ. We highlight that at each time instant there is only one target affected and the affected sensor changes with time as the target moves around the sensor network. Such that, if an affected sensor ℓ becomes unaffected, then its distribution goes back to its pre-change model. We make the assumption that as the target moves around in the network the affected sensor evolves as a Markov chain. Our goal is to quickly detect when a sensor in the network is affected by the target. To do this we seek to design a stopping time that satisfies our desire to detect this sensor as early as possible subject to a constraint on false alarms.

We now show that the problem of quickly detecting a moving target in a sensor network can be considered a structured version of the HMM QCD problem and is solved by the optimal rule from Theorem 1 [10].

Theorem 2. For quickest detection of a moving target in a sensor network with pre- and post-change densities described by (10) and (11) respectively. Then, for the cost criterion (3), there is an optimal rule with stopping time τ^*, and threshold point $h \geq 0$ given by (9).

Proof. We first define our spaces. As seen in Figure 4, let us model $S_b \triangleq \{e_1^b\}$ where none of the sensors are affected as a single state HMM and $S_a \triangleq \{e_1^a, \ldots, e_{N_a}^a\}$ where $N_a = L$ to represent the L possible states that the target could be. For example let e_1^a denote when the target is affecting sensor $\ell = 1$. For $k > 0$, we consider a random process X_k whose statistical properties change at some (unknown) time $\nu > 1$, in the sense that for $0 < k < \nu$, $X_k \in S_b$, whilst for $k \geq \nu$, $X_k \in S_a$. We model transitions of S_a as a first-order time-homogeneous Markov chain described by the $A_a^{i,j}$, for $1 \leq i, j \leq N_a$.

At time $k = \nu$, with probability ρ, the process transitions between sets according to state change probabilities $A_\nu^{i,j} = p_{\mathcal{L}}(\mathcal{B}(i))$ for $1 \leq i \leq N_a$, and $1 \leq j \leq N_b$. Finally, for each $k > 0$, X_k is indirectly observed by measurements y_k generated by conditional observation densities $b_b(y_k, 1) = f_b(y_k)$ when $k < \nu$ and $b_a(y_k, i) = f^\ell(y_k)$ for $1 \leq i \leq N_a$ and $k \geq \nu$. Given that this process can be represented as an augmented HMM as presented in Section III-B with constrained transitions, and has change event having a state independent prior $\pi_k = (1-\rho)^{k-1}\rho$ for some $\rho \in (0,1)$, we can apply Theorem 1 [10] which gives that our optimal stopping rule is (9). □

This exactly optimal result contrasts with the non-Bayesian asymptotic result of [22]. This optimal rule additionally has an efficient recursive form compared to the non-recursive GLR rule of [22].

V. Conclusion

In this paper we presented a framework for quickly detecting a change in a general dependent stochastic process and a process for converting a general dependent QCD problem into an HMM QCD problem. We investigated case studies and established that these change detection problems are structured generalisations of the HMM QCD problem and can be exactly solved by the optimal HMM QCD rule which is a threshold test on the posterior efficiently calculated via a HMM filter recursion.

VI. Acknowledgements

The authors would like to thank Aaron McFadyen for help with air traffic data preparation for Section IV.A.

References

[1] M. Basseville, "Detecting changes in signals and systems - A survey," *Automatica*, vol. 24, no. 3, pp. 309–326, May 1988.

[2] M. Basseville and I. Nikiforov, *Detection of Abrupt Change Theory and Application*, 04 1993, vol. 15.

[3] N. Vaswani, "Additive change detection in nonlinear systems with unknown change parameters," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 859–872, 2007.

[4] A. N. Shiryaev, *Optimal Stopping Rules*. Springer-Verlag Berlin Heidelberg, 2008, vol. 8.

[5] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory of Probability & Its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

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

[7] S. Dayanik and C. Goulding, "Sequential detection and identification of a change in the distribution of a Markov-modulated random sequence," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3323–3345, July 2009.

[8] B. Yakir, *Optimal detection of a change in distribution when the observations form a Markov chain with a finite state space*, ser. Lecture Notes–Monograph Series. Hayward, CA: Institute of Mathematical Statistics, 1994, vol. Volume 23, pp. 346–358.

[9] C. Fuh and A. G. Tartakovsky, "Asymptotic Bayesian theory of quickest change detection for hidden Markov models," *IEEE Transactions on Information Theory*, vol. 65, no. 1, pp. 511–529, Jan 2019.

[10] J. J. Ford, J. James, and T. L. Molloy, "Exactly optimal Bayesian quickest change detection for hidden Markov models," *Automatica*, vol. 157, p. 111232, 2023.

[11] T. L. Molloy and J. J. Ford, "Asymptotic minimax robust quickest change detection for dependent stochastic processes with parametric uncertainty," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6594–6608, 2016.

[12] D. G. Luenberger, *Optimization by Vector Space Methods*, 1st ed. USA: John Wiley & Sons, Inc., 1997.

[13] J. Lai and J. J. Ford, "Relative entropy rate based multiple hidden Markov model approximation," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 165–174, 2010.

[14] T. L. Molloy, J. J. Ford, and L. Mejias, "Adaptive Detection Threshold Selection for Vision-based Sense and Avoid," in *International Conference on Unmanned Aircraft Systems (ICUAS'17)*, Jun 2017 (In press).

[15] R. Elliott, L. Aggoun, and J. Moore, *Hidden Markov Models: Estimation and Control*. Springer-Verlag, 1995.

[16] O. Techakesari, J. J. Ford, and D. Nešić, "Practical stability of approximating discrete-time filters with respect to model mismatch," *Automatica*, vol. 48, no. 11, pp. 2965–2970, 2012.

[17] T. L. Molloy and J. J. Ford, "Asymptotic minimax robust quickest change detection for dependent stochastic processes with parametric uncertainty," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6594–6608, 2016.

[18] Y. C. Chen, T. Banerjee, A. D. Domínguez-García, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, 2016.

[19] T. Banerjee, P. Gurram, and G. Whipps, "Bayesian quickest detection of changes in statistically periodic processes," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2204–2208.

[20] T. Banerjee, P. Gurram, and G. T. Whipps, "A Bayesian theory of change detection in statistically periodic random processes," *IEEE Transactions on Information Theory*, vol. 67, no. 4, pp. 2562–2580, 2021.

[21] BBC. (2022) Sydney floods: Tens of thousands told to evacuate. [Online]. Available: https://www.bbc.com/news/world-australia-62027248

[22] G. Rovatsos, S. Zou, and V. V. Veeravalli, "Quickest detection of a moving target in a sensor network," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2399–2403.