Quickest Detection in Multiple ON–OFF Processes

Qing Zhao, *Senior Member, IEEE*, and Jia Ye

***Abstract*—We consider the quickest detection of idle periods in multiple ON–OFF processes. At each time, only one process can be observed, and the observations are random realizations drawn from two different distributions depending on the current state (ON or OFF) of the chosen process. The objective is to catch an idle period in any of the ON–OFF processes as quickly as possible subject to a reliability constraint. We show that this problem presents a fresh twist to the classic problem of quickest change detection that considers only one stochastic process. A Bayesian formulation of the problem is developed for both infinite and finite number of processes based on the theory of partially observable Markov decision process (POMDP). While a general POMDP is PSPACE-hard, we show that the optimal decision rule has a simple threshold structure for the infinite case. For the finite case, basic properties of the optimal decision rule are established, and a low-complexity threshold policy is proposed which converges to the optimal decision rule for the infinite case as the number of processes increases. This problem finds applications in spectrum sensing in cognitive radio networks where a secondary user searches for idle channels in the spectrum.**

***Index Terms*—Bayesian formulation, cognitive radio, ON–OFF process, partially observable Markov decision process (POMDP), quickest change detection, spectrum sensing.**

I. INTRODUCTION

A. The Classic Quickest Change Detection

THE classic formulation of sequential quickest change detection dates back to 1931 [1] where the application of on-line quality control of a manufacturing process was considered. In the conventional setting, the problem is to detect an abrupt change in the distribution of a *single* stochastic process, which is available through a series of observations $\{X_t\}_{t\geq 1}$ drawn sequentially in a one-at-a-time manner. As illustrated in Fig. 1, before an unknown change point T_0, the observations $X_1, \ldots, X_{T_0-1}$ are i.i.d according to a distribution $f_0(x)$; after T_0, the observations $X_{T_0}, X_{T_0+1}, \ldots$, are i.i.d with a different distribution $f_1(x)$. The objective is to design a detection rule given by a stopping time T_d to detect as quickly as possible that the change has happened. The constraint is on the reliability of the detection that captures the frequency of false alarms. The essence of the problem is the tension between the objective and the constraint: the desired reliability can be achieved through the accumulation of measurements, which comes at the price of increasing the detection delay.

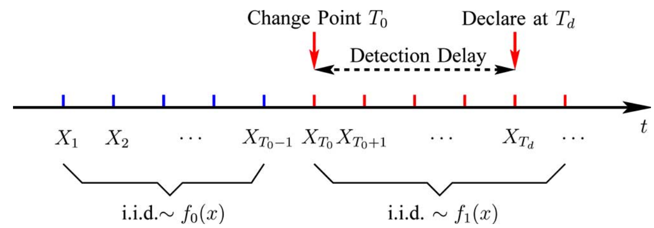

Fig. 1. Quickest change detection in a single stochastic process.

There are two standard formulations of the classic quickest change detection: Bayesian and minimax. The Bayesian formulation was developed by Shiryaev during the 1960s [2]–[4], where the change point T_0 is assumed to be random with a certain (known) prior distribution and the objective is to minimize the expected detection delay subject to an upper bound on the probability of raising a false alarm. Shiryaev considered the geometric/exponential prior distribution of the change point and established the optimal detection rule, which is often referred to as the Shiryaev–Roberts procedure [2]–[5]. Generalizations of the Shiryaev–Roberts procedure and its asymptotic optimality (as the probability of rasing a false alarm approaches 0) under arbitrary prior distributions and non-i.i.d observations have been obtained (see, for example, [6] and [7]).

The minimax formulation was proposed by Lorden in 1971 [8], in which the change point is assumed to be an unknown deterministic parameter and the objective is to minimize the worst-case conditional detection delay subject to a lower bound on the average run length to false alarm. It was shown in [8] that the well-known cumulative sum (CUSUM) algorithm proposed by Page in 1954 [9] is asymptotically (as the average run length to false alarm tends to infinity) minimax optimal with respect to Lorden's measure of the worst-case average detection delay. It was later shown by Moustakides that the minimax optimality of the CUSUM algorithm holds without the asymptotic condition [10]. In 1985, Pollak adopted a different measure of the worse-case average detection delay and showed that a randomized version of the Shiryaev–Roberts procedure is asymptotically optimal (within an additive $o(1)$ term) under this minimax formulation [11]. To date, it is not known what exactly minimizes Pollak's measure of the worst average detection delay in general. Recently, Polunchenko and Tartakovsky showed in [12] that Pollak's randomized version of the Shiryaev–Roberts procedure is not optimal by constructing a counterexample, for which the Shiryaev–Roberts procedure with a specific deterministic starting point is, in fact, optimal.

Manuscript received January 30, 2010; accepted August 21, 2010. Date of publication September 02, 2010; date of current version November 17, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dominic K. C. Ho. This work was supported by the Army Research Office under Grant W911NF-08-1-0467 and by the National Science Foundation under Grant CCF-0830685. Part of this work was presented at MILCOM, November 2008, at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2009, and at Allerton Conference, September 2009.

The authors are with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: qzhao@ucdavis.edu; jiaye@ucdavis.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2071869

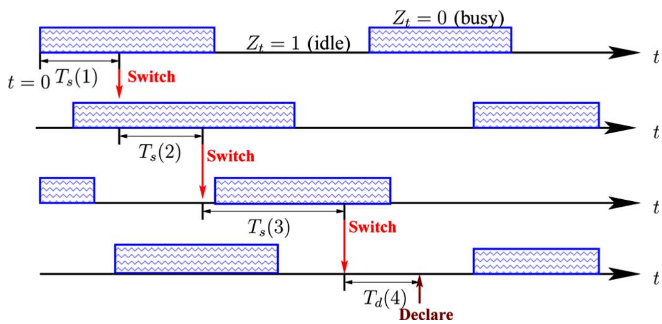

Fig. 2. Quickest detection in multiple ON–OFF processes ($T_s(l)$ denotes the time spent in the $l\mathrm{th}$ process before switching to the $(l+1)\mathrm{th}$ process; $T_d(L)$ denotes the time spent in the last process (the $L\mathrm{th}$ process) before declaring an idle period).

B. Quickest Detection in Multiple ON–OFF *Processes*

In this paper, we formulate a new form of quickest detection by considering multiple independent ON–OFF processes. The objective is to catch an idle/OFF period in any of the stochastic processes as quickly as possible subject to a reliability constraint. As illustrated in Fig. 2, assume that the user starts to monitor one of the processes at $t = 0$. At each time, the user needs to decide whether to continue in the current process, switch to a new process, or declare that the current process is idle which ends the decision horizon.

One application of this problem is cognitive radio for opportunistic spectrum access, where a secondary user searches for channels temporarily unused by primary users in a spectrum consisting of multiple channels [13]. The objective is to detect, as quickly as possible, whether the sensed channel has become idle in order to maximize the transmission time before primary users reclaim the channel. The design constraint is on the interference to primary users, i.e., the probability of declaring a busy channel as idle.

Quickest detection in multiple ON–OFF processes has two fundamental differences from the classic quickest change detection problem. First, each ON–OFF process has *an infinite number* of change points. Second, the presence of multiple processes offers an additional dimension of freedom for the quickest detection: the user does not have to wait faithfully in a single channel for an idle period. As a consequence, in addition to a detection rule that balances the tradeoff between detection delay and detection reliability as in the classic quickest change detection, a switching rule is needed to determine when to abandon the current process and seek opportunities in a new process. The tradeoff here is between avoiding long realizations of busy periods and interrupting the accumulation of "evidence" (measurements) which is crucial in detecting changes in distribution. The switching rule needs to be designed jointly with the detection rule to achieve optimality.

C. Main Results

The contribution of this paper is twofold. First, we develop a Bayesian formulation of quickest detection in multiple ON–OFF processes, a problem that has not been formulated or studied in the literature. Second, we establish the basic structures of the optimal detection and switching rules in both the infinite and the finite regimes in terms of the number of ON–OFF processes.

In the infinite case, we consider a large number of homogeneous independent ON–OFF processes and the user always switches to a new process should it decide to abandon the current one. We formulate the problem as a partially observable Markov decision process (POMDP). While POMDPs are PSAPCE-hard in general [14], we show that for the problem at hand, the optimal decision rule has a simple threshold structure when the busy and idle times of the ON–OFF processes obey (potentially different) geometric/exponential distributions. The threshold structure is with respect to the *posterior* probability λ_t that the process currently being observed is idle at time t (given the entire observation history). Specifically, the user should switch to a new process when $\lambda_t \in [0, \eta_s)$, should continue observing the current process when $\lambda_t \in [\eta_s, \eta_d)$, and should declare that the current process is idle when $\lambda_t \in [\eta_d, 1]$, where η_s and η_d are, respectively, the switching and detection thresholds. Furthermore, we show that when process switching can be done instantaneously (i.e., negligible comparing with the time for taking a measurement), the optimal switching threshold η_s is the *prior* probability (before taking any measurements) that a process is idle, which is the fraction of idle time in the ON–OFF process. The detection threshold η_d is determined by the reliability constraint. When switching to a different process takes τ_s units of time, we show that the threshold structure of the optimal policy still holds. The only difference is that the optimal switching threshold η_s is smaller than the prior probability that a process is idle.

In the finite case, we address quickest detection with memory: switching back to a previously visited process is allowed, and measurements obtained during previous visits are taken into account in decision making. We show that this freedom of switching with memory significantly complicates the problem. The resulting POMDP changes from a one-dimensional problem to an N-dimensional problem, where N is the number of ON–OFF processes. Our objective is to establish the basic structure of the optimal decision rule and develop low-complexity policies with strong performance. In particular, we show that the optimal action of declaring always occurs in the process with the largest posterior probability of being in the idle state. The monotonicity of the detection threshold is also established. Based on the basic structure of the optimal policy, we propose a low-complexity threshold policy. Specifically, under the proposed policy, the user always observes the process with the largest posterior probability of being idle and declares when the largest posterior probability exceeds the detection threshold. The near optimal performance of this threshold policy is demonstrated by a comparison with a full-sensing scheme which defines an upper bound on the optimal performance. Furthermore, we show that this low-complexity policy converges to the optimal policy for the infinite case as the number N of processes increases.

Extensions to arbitrarily distributed busy and idle times, in particular, heavy-tail distributions are discussed. For heavy-tailed busy time, we show that the persistency property of heavy-tail distributions makes it particularly important to adopt a switching strategy to avoid realizations of exceptionally long busy periods. Extensions to heterogeneous ON–OFF processes and the impact of non-negligible switching time are also discussed.

D. Other Related Work

There is a large body of literature on various formulations of quickest change detection in a single stochastic process (see [15] for an overview of classic results). There are a number of recent results on variations of quickest change detection for the application of sensor networks and target tracking. For example, quickest change detection in distributed sensor systems was addressed in [16] and [17], where distributed sensors take i.i.d. measurements from a common stochastic process and transmit quantized data to a fusion center for final decision-making. Optimal and asymptotically optimal decentralized detection rules were obtained in [16] and [17] under different criteria. In [18], nonparametric change detection algorithms were developed for detecting a change in the spatial distribution of alarmed sensors in large-scale sensor networks. In [19], the CUSUM algorithm was adopted for detecting spawning targets and was used jointly with particle filters to handle radar signal processing, data association, and target tracking simultaneously. In the context of cognitive radio for opportunistic spectrum access, the CUSUM algorithm was applied in [20] and [21] for detecting the *return* of primary users (i.e., the starting point of a busy period) in a given *single* channel.

In [22] and [23], Shiryaev considered quickest detection of a target that may appear in one of N directions. Once appeared, the target does not leave or change direction. The resulting problem is thus to detect quickly and reliably a *single* random change point in *one* of N processes when all other $N-1$ processes contain no change. Shiryaev used the Wald's sequential analysis and the Neyman–Pearson method to solve the problem and characterized the average detection delay.

To our best knowledge, this work and the preceding conference versions are the first that address quickest detection in multiple ON–OFF processes.

II. Classic Quickest Change Detection Under Bayesian Formulation

In this section, we give a brief overview on the classic quickest change detection under Bayesian formulation and the Shiryaev–Roberts procedure [2]–[5], where the change point T_0 has a known geometric distribution parameterized by p. Specifically,

$$\Pr[T_0 = k] = p(1-p)^{k-1}(1-\lambda_0), \quad \forall k > 0$$

where $\lambda_0 \triangleq \Pr[T_0 = 0]$ is the probability that the change occurs before the observation starts. The problem of quickest change detection is to design a detection rule given by a stopping time T_d (see Fig. 1) under the following objective and constraint:

$$\min \mathbb{E}[(T_d - T_0)^+] \text{ subject to } \Pr[T_d < T_0] \leq \zeta \tag{1}$$

where $(T_d - T_0)^+ \triangleq \max\{0, T_d - T_0\}$ is the detection delay, $\Pr[T_d < T_0]$ the false alarm probability, and the expectation and the probability are averaged over the prior distribution of the change point.

Shiryaev showed that a sufficient statistic for quickest change detection is the posterior probability λ_t that the change has already occurred given the measurements obtained up to t:

$$\lambda_t \triangleq \Pr[T_0 \leq t | X_1, X_2, \ldots, X_t]. \tag{2}$$

Based on Bayes rule, the sufficient statistic λ_t can be computed recursively at each time t using the new observation $X_t = x$. The Shiryaev–Roberts procedure for quickest change detection is thus given by the following stopping rule on the posterior probability λ_t.

$$T_d = \inf\{t : \lambda_t \geq \eta_d\}, \tag{3}$$

where the detection threshold η_d is determined by the reliability constraint ζ given in (1). A closed-form expression of the detection threshold η_d is generally intractable. Setting $\eta_d = 1 - \zeta$ has been shown to be asymptotically optimal as the reliability constraint becomes more strict, i.e., as ζ approaches 0. Extensions of the Shiryaev–Roberts procedure to arbitrary prior distributions of the change point and its optimality under both Bayesian and minimax formulations and various asymptotic conditions have been studied extensively in the literature (see Section I-A).

III. Quickest Detection in Multiple Processes: The Infinite Regime

In this section, we consider the case where there are an infinite number of ON–OFF processes and the user always switches to a new process should it decide to abandon the current one. A Bayesian formulation of the problem is developed and the optimal decision rule is obtained.

A. Problem Statement

We consider an infinite number of independent homogeneous ON–OFF processes. Let $\{B_i\}_{i=-\infty}^{\infty}$ and $\{I_i\}_{i=-\infty}^{\infty}$ denote, respectively, the lengths of each busy and idle periods in a particular process. We assume that the busy periods $\{B_i\}_{i=-\infty}^{\infty}$ are i.i.d. with a geometric distribution parameterized by p_B, and the idle periods $\{I_i\}_{i=-\infty}^{\infty}$ are i.i.d. with a geometric distribution parameterized by p_I. The average busy and idle times are denoted by $m_B = 1/p_B$ and $m_I = 1/p_I$, respectively. Let λ_0 denote the fraction of idle time. It is given by

$$\lambda_0 = \frac{m_I}{m_B + m_I}. \tag{4}$$

This discrete-time model is equivalent to a continuous-time model where each ON–OFF process is a two-state continuous-time Markov chain (i.e., exponentially distributed busy and idle times) and the user samples the process at discrete times with a given sampling rate (see [24] on basics of Markov chains).

As illustrated in Fig. 2, let L be the total number of channels visited by the user until it declares, correctly or mistakenly, an idle period (for the example given in Fig. 2, $L = 4$). It is a random variable depending on the switching and detection rules and the random observations in each process. Let $T_s(l)\ (l = 1, \ldots, L-1)$ denote the time spent in the lth process before switching to the $(l+1)$th process. Let $T_d(L)$ denote the time spent in the last process (the Lth process) before declaring an idle period. The problem of quickest detection in multiple processes can be formulated as jointly choosing a sequence of switching rules $\{T_s(l)\}_{l=1}^{L-1}$ and a detection rule $T_d(L)$ under the following objective and constraint:

$$\begin{aligned} &\min \quad \mathbb{E}\left[\sum_{l=1}^{L-1} T_s(l) + T_d(L)\right] \\ &\text{subject to} \quad \Pr\left[Z_L\left(\sum_{l=1}^{L-1} T_s(l) + T_d(L)\right) = \text{busy}\right] \le \zeta \end{aligned} \tag{5}$$

where $\mathbb{E}[\sum_{l=1}^{L-1} T_s(l) + T_d(L)]$ reflects the expected waiting time before catching an idle period, and $Z_L(t)$ is the state of the Lth process at time t. We can see from (5) that quickest detection in multiple stochastic processes is fundamentally different from that in a single process, and is significantly more complex in that a sequence of stopping times $(T_s(1), T_s(2), \ldots, T_s(L-1), T_d(L))$ need to be designed.

B. A POMDP Formulation

In this section, we show that one approach to solving the problem given in (5) is to formulate it as a POMDP over a random horizon as detailed below.

1) State Space: The underlying system has three states: $Z_t \in \{0, 1, \Delta\}$, where 0 and 1 indicate, respectively, that the current process is busy and idle, Δ is an absorbing state, indicating the end of the decision horizon.

2) Action Space: There are three actions at each decision time: S (switch and take a measurement in a new process), C (continue taking measurements in the current process), and D (declare that the current process is idle).

3) State Transition: Due to the memoryless property of the geometric distributions of the busy and idle times, the underly dynamic system is Markovian. The transition probabilities under each action are given in Fig. 3. For example, when the current state is 0 and action C is taken, with probability $1 - p_B$ the system stays in 0, with probability p_B, the system transits to state 1. Whenever action D is taken, the system transits to the absorbing state Δ.

4) Observation Model: The observation at time t is X_t under actions S and C. The distribution of X_t is given by either $f_0(x)$ (if $Z_t = 0$) or $f_1(x)$ (if $Z_t = 1$) depending on the current state of the underlying system. Under action D, no observations are available.

5) Cost Structure: The actions of S and C have a unit cost that measures the delay in catching an idle period. Declaring a busy process as idle incurs a cost of γ that models the tradeoff between detection delay and detection reliability. It is set to satisfy the reliability constraint ζ given in (5). Note that it is not necessary to specify the value of γ based on ζ. As shown

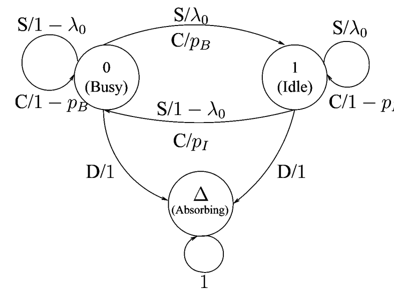

Fig. 3. The state transition diagram: the infinite regime.

in Section III-C, the value of γ is translated into a detection threshold chosen to satisfy the reliability constraint ζ in the optimal decision rule.

The objective is to choose actions sequentially in time to minimize the expected total cost over an infinite horizon, or equivalently, over a random horizon defined by the hitting time of the absorbing state Δ. It is clear from the cost structure that the expected total cost (excluding the potential cost of γ at the end of the decision horizon) is the expected delay in catching an idle period.

Since the underlying system state Z_t (the state of the current process at time t) is not directly observable from the measurements $\{X_k\}_{k=1}^t$, what we have here is a POMDP. Based on [25] and the i.i.d. nature of the observations, we know that a sufficient statistic for choosing the optimal action at each time is the information state or the belief value: the posterior probability λ_t that $Z_t = 1$ (the current process is idle) given the measurements obtained up to t:

$$\lambda_t \triangleq \Pr[Z_t = 1 | X_1, X_2, \ldots, X_t]. \tag{6}$$

Comparing (6) with (2), we can see that in the special case of a single stochastic process with a single change point T_0 as considered by Shiryaev, these two forms of sufficient statistic λ_t are equivalent.

It can be shown that λ_t has the following recursive update depending on the action a_{t-1} and the observation X_t:

$$\lambda_t = \begin{cases} \mathcal{T}(\lambda_0 | x), & a_{t-1} = \text{S},\ X_t = x \\ \mathcal{T}(\lambda_{t-1} | x), & a_{t-1} = \text{C},\ X_t = x \end{cases} \tag{7}$$

where $\mathcal{T}(\lambda|x)$ denotes the updated information state based on a new measurement x. Let $\bar{p} \triangleq 1 - p$ for $p \in [0, 1]$. We have

$$\begin{aligned} &\mathcal{T}(\lambda_{t-1}|x) \\ &\quad \triangleq \Pr[Z_t = 1 | \lambda_{t-1}, x] \\ &\quad = \frac{(\lambda_{t-1}\bar{p}_I + \bar{\lambda}_{t-1} p_B) f_1(x)}{(\lambda_{t-1}\bar{p}_I + \bar{\lambda}_{t-1} p_B) f_1(x) + (\lambda_{t-1} p_I + \bar{\lambda}_{t-1}\bar{p}_B) f_0(x)}. \end{aligned} \tag{8}$$

Note that when action S is taken, the new information state is updated from λ_0, the fraction of idle time which is the probability of hitting an idle period right away. We consider the nontrivial case where $\lambda_0 < 1 - \zeta$ so that the user cannot declare without taking any measurements. A policy π_∞ for detecting an idle period is a function that maps an information state $\lambda_t \in [0, 1]$ to

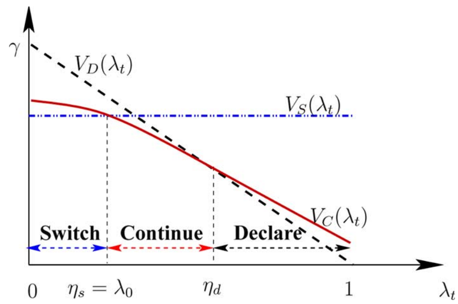

Fig. 4. The threshold structure of the optimal policy.

an action $a_t = \pi_\infty(\lambda_t)$. Quickest detection in the regime of infinite ON–OFF processes can thus be formulated as the following stochastic optimization problem:

$$\pi_\infty^* = \arg\min_{\pi_\infty} \mathbb{E}_{\pi_\infty}\left[\sum_{t=0}^{\infty} R_{\pi_\infty(\lambda_t)} | \lambda_0 = \frac{m_I}{m_B + m_I}\right] \quad (9)$$

where $R_{\pi_\infty(\lambda_t)}$ is the cost incurred under action $\pi_\infty(\lambda_t)$.

C. The Optimal Policy: A Threshold Policy

Referred to as the value function, $V(\lambda_t)$ denotes the minimum expected total remaining cost when the current information state is λ_t. It specifies the performance of the optimal policy π_∞^* starting from the information state λ_t. Let $V_S(\lambda_t)$ denote the expected total remaining cost when we take action S at the current time and then follow the optimal policy π_∞^*. Let $V_C(\lambda_t)$ and $V_D(\lambda_t)$ be similarly defined. We thus have

$$V(\lambda_t) = \min\{V_S(\lambda_t),\ V_C(\lambda_t),\ V_D(\lambda_t)\}. \quad (10)$$

From the cost structure, we obtain the following:

$$V_S(\lambda_t) = 1 + \int_x P(x;\lambda_0) V(\mathcal{T}(\lambda_0|x))\mathrm{d}x \quad (11)$$

$$V_C(\lambda_t) = 1 + \int_x P(x;\lambda_t) V(\mathcal{T}(\lambda_t|x))\mathrm{d}x \quad (12)$$

$$V_D(\lambda_t) = (1-\lambda_t)\gamma \quad (13)$$

where

$$P(x;\lambda) = (\lambda\bar{p}_I + \bar{\lambda}p_B)f_1(x) + (\lambda p_I + \bar{\lambda}\bar{p}_B)f_0(x) \quad (14)$$

is the probability density of observing x at time $t+1$ when the process has probability λ to be idle at time t. We then arrive at the following lemma which is the key to the threshold structure of the optimal policy.

Lemma 1: $V_S(\lambda_t) = V_C(\lambda_0)$ and is independent of λ_t. $V_D(\lambda_t)$ is linearly decreasing with λ_t. $V_C(\lambda_t)$ is concave and monotonically decreasing with λ_t when $p_B + p_I \leq 1$.

Proof: See Appendix A. ■

Lemma 1 establishes basic properties of the value functions under three possible current actions (see an illustration in Fig. 4). Based on Lemma 1, it is easy to see that $V_C(\lambda_t)$ and $V_S(\lambda_t)$ have a unique intersecting point $\eta_s = \lambda_0$. Noticing that $V_D(0) > V_C(0)$ and $V_D(1) < V_C(1)$, we can show that $V_C(\lambda_t)$ and $V_D(\lambda_t)$ have a unique intersecting point $\eta_d > \lambda_0$ (see Fig. 4). We thus have the following theorem.

Theorem 1: When $p_B + p_I \leq 1$, the quickest detection π_∞^* of an idle period in multiple ON–OFF processes is given by two thresholds η_s and $\eta_d \in (\eta_s, 1]$: switch to a new process if $\lambda_t \leq \eta_s$, continue in the current process if $\lambda_t \in (\eta_s, \eta_d)$, and declare if $\lambda_t \geq \eta_d$, i.e.,

$$\pi_\infty^*(\lambda_t) = \begin{cases} \text{S}, & \lambda_t \in [0, \eta_s) \\ \text{C}, & \lambda_t \in [\eta_s, \eta_d) \\ \text{D}, & \lambda_t \in [\eta_d, 1]. \end{cases} \quad (15)$$

Furthermore, the optimal switching threshold $\eta_s = \lambda_0$, the fraction of the idle time.

This simple threshold structure of the optimal policy π_∞^* agrees with our intuition: switch to a new process when the prospect of catching an idle period in a new process is better than staying in the current one (i.e., $\lambda_t \leq \lambda_0$). The condition of $p_B + p_I \leq 1$ generally holds. In particular, if the geometric distributions of the busy and idle times result from sampling exponential distributions (as it is often the case in practice where the processes are in continuous time), this condition is always satisfied. In the discrete case, as long as the average busy and idle times are longer than two samples, this condition would hold.[1]

The detection threshold η_d is chosen to satisfy the interference constraint ζ. Setting $\eta_d = 1 - \zeta$ always meets the constraint but potentially leads to suboptimal performance. Numerical methods developed for POMDP can be used for computing the optimal detection threshold (see, for example, [26]).

IV. Quickest Detection in Multiple Processes: The Finite Regime

In this section, we consider a finite number of ON–OFF processes and address switching with memory. The basic structure of the optimal decision rule are established, based on which a simple threshold policy with strong performance is proposed.

A. An N-Dimensional POMDP

Consider N homogeneous and independent ON–OFF processes. At each time instant, based on all the observations obtained so far, the user decides whether to declare or to continue taking measurements, and in which process such an action should be taken. We show below that this freedom of switching with memory significantly complicates the problem. The resulting POMDP changes from a one-dimensional problem to an N-dimensional problem.

1) State Space: The system state at time t is given by $\{Z_1(t), \ldots, Z_N(t)\}$, where $Z_i(t) \in \{0, 1\}$ denotes the state of the ith process at time t. We then augment the state space by an absorbing state Δ which indicates the end of the decision horizon.

2) Action Space: The action space is $\{C_i, D_i\}_{i=1}^N$ where C_i and D_i denote, respectively, the action of continuing taking measurements and the action of declaring in the ith process.

[1] In the special case of $p_B + p_I = 1$, the underlying Markov system reduces to an i.i.d. system. Past observations in a process do not bear any information for inferring the current busy/idle state of this process. One should simply decide whether to declare based on the current observation, and staying in the same process or is equivalent to switching to a new process.

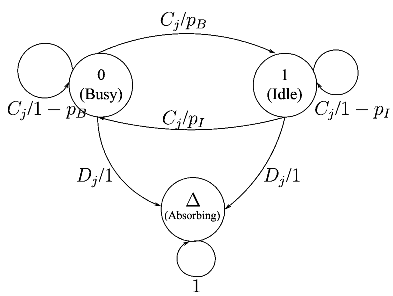

Fig. 5. The state transition diagram: the finite regime.

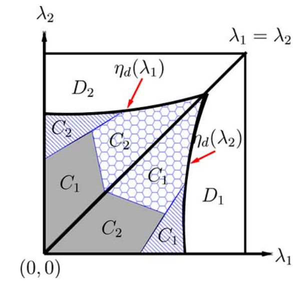

Fig. 6. The basic structure of the optimal policy: the finite regime.

3) State Transition: The state transition of the ith process under all possible actions $\{C_j, D_j\}_{j=1}^N$ are given in Fig. 5. Note that the busy/idle state of the ith process evolves independently of the action as long as the user decides to continue sensing. Whenever the action of declaring is taken, the system state transits to Δ.

4) Observation Model: The observation at time t is X_t under action C_i. The distribution of X_t is given by either $f_0(x)$ or $f_1(x)$ depending on the current state $Z_i(t)$. Under action D_i, no observation is available.

5) Cost: Action C_i has a unit cost. Declaring a busy process as idle incurs a cost of γ.

6) A Sufficient Statistic: The information state or the belief vector [25] as a sufficient statistic is now an N-dimensional vector: $\mathbf{\Lambda}(t) \triangleq [\lambda_1(t), \ldots, \lambda_N(t)]$, where $\lambda_i(t)$ is the posterior probability that the ith process is idle given all the past measurements.

Given the action a_t and the observation X_t at time t (if observation is available), the belief value of the ith process can be updated as follows:

$$\lambda_i(t) = \begin{cases} \mathcal{T}(\lambda_i(t-1)|x) & a_{t-1} = C_i,\ X_t = x \\ \mathcal{T}(\lambda_i(t-1)) & a_{t-1} = C_j,\ j \neq i \end{cases} \tag{16}$$

where $\mathcal{T}(\lambda|x)$ denotes the updated belief based on a new measurement x as given in (8), and $\mathcal{T}(\lambda)$ denotes the updated belief based purely on the underlying Markov chain defined by the geometric distributions of the busy and idle times, i.e.,

$$\mathcal{T}(\lambda) = \lambda\bar{p}_I + \bar{\lambda}p_B. \tag{17}$$

A policy π_N for detecting an idle period in N ON–OFF processes is a function that maps a belief vector $\mathbf{\Lambda}(t)$ to an action $a_t = \pi_N(\mathbf{\Lambda}(t))$. We arrive at the following stochastic control problem:

$$\pi_N^* = \arg\min_{\pi_N} \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} R_{\pi_N(\mathbf{\Lambda}(t))} | \mathbf{\Lambda}(0)\right] \tag{18}$$

where $R_{\pi_N(\mathbf{\Lambda(t)})}$ is the cost incurred under action $\pi_N(\mathbf{\Lambda(t)})$ and $\mathbf{\Lambda}(0) = \{\lambda_0, \ldots, \lambda_0\}$.

B. Basic Structures of the Optimal Policy

In this section, we establish the basic structure of the optimal policy. For simplicity of the presentation, we first consider the case of $N = 2$. Extensions to $N > 2$ is straightforward. The time index in $\lambda_i(t)$ is omitted. Let $V(\lambda_1, \lambda_2)$ denote the value function. We have

$$\begin{aligned} V(\lambda_1, \lambda_2) = \min\{&V_{C_1}(\lambda_1, \lambda_2),\ V_{C_2}(\lambda_1, \lambda_2), \\ &V_{D_1}(\lambda_1, \lambda_2),\ V_{D_2}(\lambda_1, \lambda_2)\} \end{aligned} \tag{19}$$

where $V_{C_i}(\lambda_1, \lambda_2)$ and $V_{D_i}(\lambda_1, \lambda_2)$ are the value functions for a given current action. From the cost structure, we obtain the following:

$$\begin{aligned} V_{C_1}(\lambda_1, \lambda_2) &= 1 + \int_x P(x; \lambda_1) V(\mathcal{T}(\lambda_1|x), \mathcal{T}(\lambda_2)) \mathrm{d}x \\ V_{C_2}(\lambda_1, \lambda_2) &= 1 + \int_x P(x; \lambda_2) V(\mathcal{T}(\lambda_1), \mathcal{T}(\lambda_2|x)) \mathrm{d}x \\ V_{D_i}(\lambda_1, \lambda_2) &= (1 - \lambda_i)\gamma \end{aligned} \tag{20}$$

where $P(x; \lambda)$ is given in (14). We then have the following lemma that characterizes the value functions.

Lemma 2:

L2.1: When $p_B + p_I \leq 1$, $V_{C_i}(\lambda_1, \lambda_2)$ are concave and monotonically decreasing with λ_i for all i.

L2.2: $V_{C_i}(\lambda_1, \lambda_2)$ are symmetric with respect to the plane $\lambda_1 = \lambda_2$, i.e., $V_{C_1}(\lambda_1, \lambda_2) = V_{C_2}(\lambda_2, \lambda_1)$.

L2.3: $V_{D_i}(\lambda_1, \lambda_2)$ is linearly decreasing with λ_i for all i.

Proof: See Appendix B. ■

Based on Lemma 2, we obtain the following basic structure of the optimal policy.

Theorem 2: When $p_B + p_I \leq 1$, the optimal action $a^*(\lambda_1, \lambda_2)$ under the belief vector (λ_1, λ_2) is in the following form:

$$a^*(\lambda_1, \lambda_2) = \begin{cases} D_1, & \text{if } \lambda_1 \geq \lambda_2,\ \lambda_1 > \eta_d(\lambda_2) \\ D_2, & \text{if } \lambda_1 < \lambda_2,\ \lambda_2 > \eta_d(\lambda_1) \\ C_1 \text{ or } C_2, & \text{Otherwise} \end{cases} \tag{21}$$

where $\eta_d(\lambda) : [0, 1] \to [0, 1]$ is the detection threshold which is monotonically increasing in λ. Furthermore, the action of continuing has a symmetric structure, i.e.,

$$a^*(\lambda_1, \lambda_2) = C_1 \Leftrightarrow a^*(\lambda_2, \lambda_1) = C_2. \tag{22}$$

Proof: See Appendix C. ■

Fig. 6 illustrates the basic structure of the optimal policy given in Theorem 2. The threshold interface $\eta_d(\lambda)$ between action C and D and its monotonicity are illustrated. The partition of the region for action C into C_1 and C_2 is unknown except its symmetric structure.

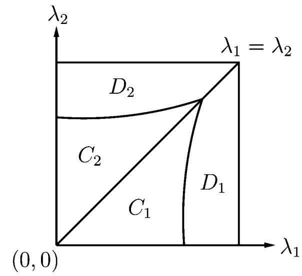

Fig. 7. A low-complexity threshold policy $\hat{\pi}_N$ for the finite regime.

The result holds for the general case of $N > 2$. Specifically, under the optimal policy, the user always declares on the process with the largest belief value, and the declaring threshold is monotonically increasing with the belief values of the rest $N-1$ processes. The proof follows the same line as in the case of $N = 2$.

C. A Low-Complexity Threshold Policy

In this section, we propose a low-complexity threshold policy based on the basic structure of the optimal policy established in the previous subsection.

As illustrated in Fig. 7 where we consider $N = 2$, under the proposed threshold policy $\hat{\pi}_N$, when both λ_1 and λ_2 are below their corresponding detection thresholds $\eta_d(\lambda_2)$ and $\eta_d(\lambda_1)$, the user continues taking observations on the process with a larger belief value. Otherwise, the user declares that an idle state has been reached on the process with a larger belief value. In the case of general N, let $\mathbf{\Lambda}^{-i} = \{\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_N\}$. We have

$$\hat{\pi}_N(\mathbf{\Lambda}) = \begin{cases} C_i, & \text{if } i = \arg\max_{1\le j\le N}\{\lambda_j\} \text{ and } \lambda_i < \eta_d(\mathbf{\Lambda}^{-i}) \\ D_i, & \text{if } i = \arg\max_{1\le j\le N}\{\lambda_j\} \text{ and } \lambda_i \ge \eta_d(\mathbf{\Lambda}^{-i}). \end{cases} \tag{23}$$

We show in the theorem below that this low-complexity threshold policy converges to the optimal policy π^*_∞ for the case of infinite processes.

Theorem 3: As the number N of processes increases, $\hat{\pi}_N$ converges to the optimal policy π^*_∞ for detecting an idle period in an infinite number of ON–OFF processes.

Proof: See Appendix D. ∎

V. Extensions and Discussions

A. Extension to Arbitrary Distributions

We now discuss extensions to ON–OFF processes with arbitrarily distributed busy and idle times. We focus on the infinite regime and show how the threshold policy π^*_∞ can be implemented for quickest detection in arbitrarily distributed ON–OFF processes. When the busy and idle times are not geometrically distributed or the observations are not i.i.d., the posterior probability λ_t is no longer a sufficient statistic in general. While the optimality of π^*_∞ may be lost in this case, it can still be implemented, and simulation results demonstrate its strong performance as shown in Section VI.

Assume that the busy and idle times have distribution $\{g_0(k)\}_{k=1}^\infty$ and $\{g_1(k)\}_{k=1}^\infty$, respectively. We are particularly interested in scenarios where the busy time distribution g_0 is a heavy-tail distribution. For example, in the context of cognitive radio for opportunistic spectrum access, the connection time of the primary users may have a heavy-tail distribution (see [27] and [28] on the self-similar nature of the communication traffic and its relation to heavy-tail distributed file size).

A commonly used heavy-tail distribution is the Pareto distribution:

$$g_0(k) = \begin{cases} 0, & k < a \\ a^\alpha(k^{-\alpha} - (k+1)^{-\alpha}), & k \ge a \end{cases}$$

where $a > 0$ is the minimum busy time and α is the tail index. We consider $\alpha > 1$ so that the busy time has a finite mean but potentially infinite variance. It is easy to show that for a Pareto distributed busy time B, $\forall\, s > 0$,

$$\Pr[B > \tau + s \mid B > \tau] \nearrow 1 \quad \text{as} \quad \tau \to \infty.$$

This is the persistency property of heavy-tail distributions. In other words, a busy time that has lasted longer than a certain threshold is more likely to persist into the future, and such exceptionally large realizations, albeit rare, dominate the average behavior. This persistency property of heavy-tail distribution makes it crucial to design an optimal switching rule to avoid exceptionally long busy periods.

To implement π^*_∞, we only need to compute the belief value λ_t defined in (6). For arbitrarily distributed busy and idle times, however, the belief can no longer be computed recursively as given in (7). At each t, all the observations obtained in the current process are needed for computing λ_t. For notation simplicity, assume that the time index is reset to $t = 1$ when the user switches to a new process, i.e., $X_1, \ldots, X_t$ are obtained from the same process. We thus have (24), shown at the bottom of the page, where $\Pr[Z_1 = z_1, \ldots, Z_t = z_t]$ can be computed based

$$\lambda_t = \frac{\sum_{(z_1,\ldots,z_{t-1})\in\{0,1\}^{t-1}} \prod_{j=1}^{t-1} f_{z_j}(X_j) f_1(X_t) \Pr[Z_1 = z_1, \ldots, Z_{t-1} = z_{t-1}, Z_t = 1]}{\sum_{(z_1,\ldots,z_t)\in\{0,1\}^t} \prod_{j=1}^{t} f_{z_j}(X_j) \Pr[Z_1 = z_1, \ldots, Z_t = z_t]} \tag{24}$$

on the distributions of the busy and idle times. Specifically, let $r_{0,i}$ denote the length of the $i\mathrm{th}$ segment of 0's in $(z_1, \ldots, z_t)$, and $r_{1,i}$ the length of the $i\mathrm{th}$ segment of 1's in $(z_1, \ldots, z_t)$. We have

$$\Pr[Z_1 = z_1, \ldots, Z_t = z_t] = \begin{cases} (1-\lambda_0)\underline{g_0}(r_{0,1}) \prod_{i>1} g_0(r_{0,i}) \prod_{i\geq 1} g_1(r_{1,i}), & \text{if } z_1 = 0 \\ \lambda_0 \underline{g_1}(r_{1,1}) \prod_{i\geq 1} g_0(r_{0,i}) \prod_{i>1} g_1(r_{1,i}), & \text{if } z_1 = 1 \end{cases} \tag{25}$$

where $\underline{g_0}$ and $\underline{g_1}$ are, respectively, the distributions of the residual busy and idle times when the user starts observing a process (after switching). Note that $\underline{g_0}$ and $\underline{g_1}$ are different from g_0 and g_1 (except when g_0 and g_1 are geometric) since the time instant at which the user starts observing a process is not synchronized with the starting point of a busy or idle period. Based on the distribution of the so-called forward renewal time or the residual life of a renewal interval [29], we have

$$\underline{g_0}(k) = \frac{1}{m_B} \sum_{l=k}^{\infty} g_0(l), \quad k > 0. \tag{26}$$

$\underline{g_1}(k)$ can be similarly obtained from g_1. For Pareto distributed busy time, we have

$$\underline{g_0}(k) = \begin{cases} \frac{\alpha-1}{a\alpha}, & 0 < k \leq a \\ \frac{a^{\alpha-1}}{\alpha}((k-1)^{-(\alpha-1)} - k^{-(\alpha-1)}), & k > a \end{cases}$$

which remains to be a heavy-tail distribution with a tail index of $\alpha - 1$ (a heavier tail).

A recursive implementation of π^*_∞ can be obtained under the assumption that the user will declare an idle period before the process changes back to a busy period (i.e., the user experiences at most a single change from busy to idle in each observed process). This assumption holds approximately when channel idle periods are longer than the detection time with high probability, which is the scenario when opportunistic spectrum access is feasible. Under this assumption, we can obtain a recursive implementation by considering the likelihood ratio $\hat{\lambda}_t$ defined similarly to the case of quickest detection in a single process [6]. Specifically, the likelihood ratio $\hat{\lambda}_t$ can be recursively updated as follows:

$$\hat{\lambda}_t = \begin{cases} \frac{\lambda_0}{1-\lambda_0}, & t = 0 \\ \left(\frac{\underline{G^c_0}(t-1)}{\underline{G^c_0}(t)} \hat{\lambda}_{t-1} + \frac{g_0(t)}{\underline{G^c_0}(t)} \right) \frac{f_1(X_t)}{f_0(X_t)}, & t > 0 \end{cases} \tag{27}$$

where $\underline{G^c_0}(t) = \sum_{k=t+1}^{\infty} \underline{g_0}(k)$ is the complement cumulative distribution of the residual busy time.

The threshold policy π^*_∞ on λ_t is equivalent to a threshold policy on $\hat{\lambda}_t$: switch to a new process when $\hat{\lambda}_t \leq \hat{\eta}_s$, continue in the current process when $\hat{\lambda}_t \in (\hat{\eta}_s, \hat{\eta}_d)$, and declare when $\hat{\lambda}_t \geq \hat{\eta}_d$, where the switching threshold $\hat{\eta}_s = \lambda_0/(1-\lambda_0)$, and the detection threshold can be set to $\hat{\eta}_d = (1-\zeta)/\zeta$.

Extensions to non-i.i.d. observations can be similarly obtained when the joint distributions of the observations are used in (24) to calculate λ_t.

B. Impact of the Switching Time

We have assumed that process switching time is negligible. We show here that the threshold structure of the optimal policy π^*_∞ still holds for the general case with an arbitrary switching time τ_s. The only difference is that the optimal switching threshold η_s is smaller than λ_0 when $\tau_s > 0$. This can be shown by noticing that $V_S(\lambda_t) = \tau_s + V_C(\lambda_0)$, i.e., the horizontal line in Fig. 4 is raised up by τ_s and intersects with $V_C(\lambda_t)$ at a point smaller than λ_0. It is possible that when the switching time τ_s is sufficiently large, the optimal policy is to never switch.

C. Extension to Heterogeneous ON–OFF Processes

The N-dimensional POMDP formulation of quickest detection in N ON–OFF processes can be readily extended to handle heterogeneous ON–OFF processes. In this case, the belief value $\lambda_i(t)$ of the $i\mathrm{th}$ process is updated based on (16) by using the parameters $p_B^{(i)}$, $p_I^{(i)}$, $f_0^{(i)}$, and $f_1^{(i)}$ of this particular process. While the basic threshold structure of the optimal policy given in Theorem 2 may be lost, the proposed low-complexity policy $\hat{\pi}_N$ can still be applied and has been observed to achieve near optimal performance in simulation examples.

VI. Simulation Examples

In this section, we study the performance of π^*_∞ and $\hat{\pi}_N$ through simulation examples. We consider the application of cognitive radio for opportunistic spectrum access. The primary signals are modeled as Gaussian signals in Gaussian noise, i.e., $f_0(x)$ and $f_1(x)$ are Gaussian distributions with zero mean and variances σ_0^2 and σ_1^2, respectively. The signal-to-noise ratio (SNR) is given by $10\log(\sigma_0^2 - \sigma_1^2)/\sigma_1^2$. In all the examples, the detection threshold η_d is set to $1-\zeta$.

A. Infinite Regime

We first consider geometrically distributed busy and idle times. Shown in Fig. 8 (left) is the expected time to catch an idle period as a function of λ_0, the fraction of idle time in each channel. Compared to the strategy that stays in a single channel, π^*_∞ offers significant improvement over a large range of λ_0. When λ_0 is close to 1, these two strategies have comparable performance. This is because when λ_0 is close to 1, with high probability the user will hit an idle period immediately in the first channel. This is confirmed in Fig. 8 (right), which plots the expected number of channels visited under π^*_∞ as a function of λ_0. This figure also shows that even when λ_0 is small, the user only needs to visit a small number of channels before catching an idle one. Yet a significant reduction in detection time is achieved over the single-channel strategy. This observation also suggests that we do not require a large number of channels to see the optimality of π^*_∞.

We now consider the case where the idle time has a geometric distribution and the busy time has a truncated Pareto distribution with a truncation window length W. The distribution of the residual busy time is given by

$$\underline{g_0}(k) = \begin{cases} \frac{(\alpha-1)(1-(\frac{a}{W})^\alpha)}{\alpha a^\alpha (a^{-(\alpha-1)} - W^{-(\alpha-1)})} & 1 \leq k \leq a \\ \frac{(k-1)^{-(\alpha-1)} - k^{-(\alpha-1)} - W^{-\alpha}(\alpha-1)}{\alpha(a^{-(\alpha-1)} - W^{-(\alpha-1)})} & a+1 \leq k \leq W \\ 0 & k > W. \end{cases}$$

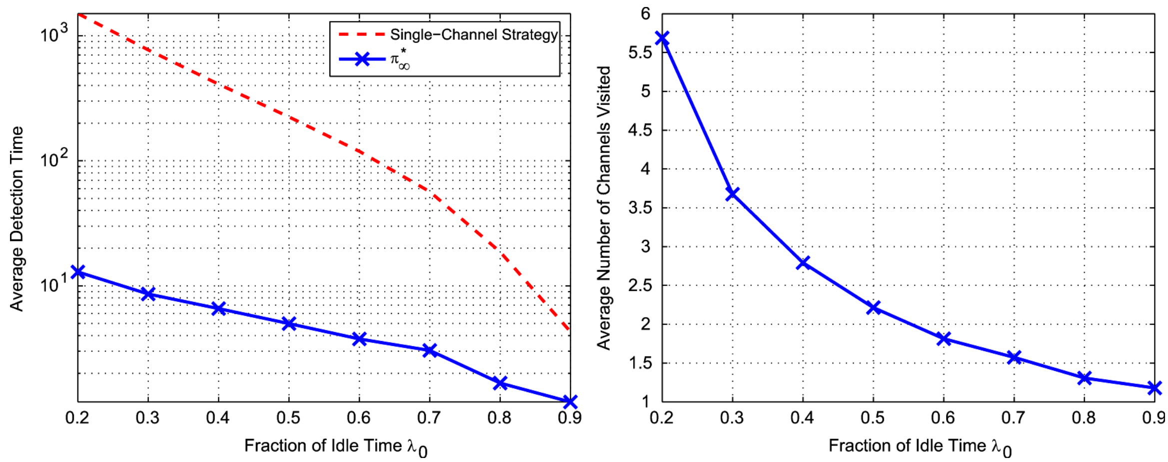

Fig. 8. Average detection time and average number of channels visited ($m_I = 500, m_B = (1-\lambda_0)/\lambda_0 m_I, \zeta = 0.1, \mathrm{SNR} = 10$).

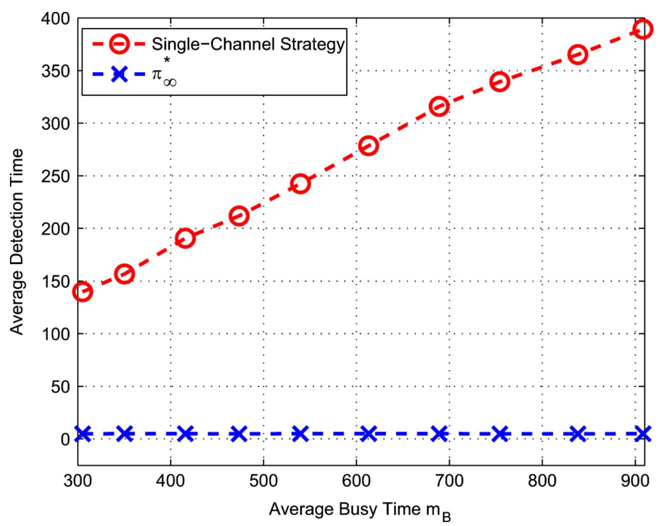

Fig. 9. Average detection time versus m_B for Pareto distributed busy time ($\lambda_0 = 0.5, a = 200, \alpha = 1.01, m_I = \lambda_0/(1-\lambda_0)m_B, \zeta = 0.1, \mathrm{SNR} = 10$).

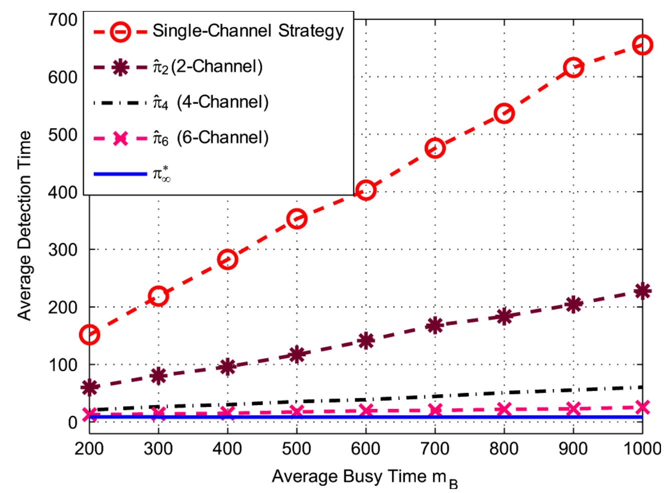

Fig. 10. Average detection time versus m_B. ($\lambda_0 = 0.3, p_I = p_B(1-\lambda_0)/\lambda_0, \mathrm{SNR} = 10, \zeta = 0.1$).

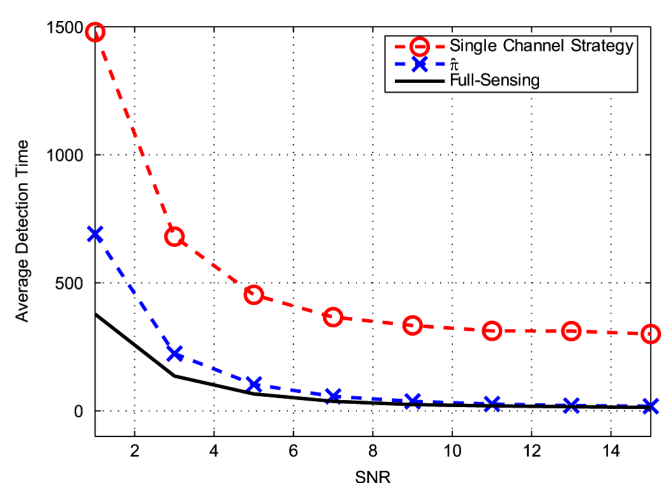

Fig. 11. Average detection time versus SNR ($m_I = 500, m_B = 600, \zeta = 0.1, N = 4$).

Fig. 9 shows the expected detection time as a function of the average busy time m_B. Specifically, we increase both the average busy time m_B (by increasing the truncation window W) and the average idle time m_I while keeping the fraction λ_0 of idle time fixed. We observe that the average detection time of the single-channel strategy increases quickly with m_B, as expected. In sharp contrast, π^*_∞ maintains the same average detection time, independent of the average busy time in every channel. This is due to the channel switching mechanism in π^*_∞ that avoids large realizations of busy period.

B. Finite Regime

In Fig. 10, we show the performance of $\hat{\pi}_N$ for different N. We increase both m_B and m_I while keeping λ_0 fixed. We observe that the average detection time decreases quickly as N increases, and the performance of $\hat{\pi}_N$ quickly converges to that of the optimal policy π^*_∞ for the infinite-channel case.

Shown in Fig. 11 is the performance of $\hat{\pi}_N$ at different SNR. We use the full-sensing scheme (i.e., the user can sense all channels simultaneously) as a benchmark, which provides a lower bound on the minimum detection time. We observe that $\hat{\pi}_N$ achieves near optimal performance and offers significant improvement over the single-channel strategy.

VII. CONCLUSION

In this paper, we have formulated and studied a new form of quickest change detection: quickest detection of an idle period in multiple ON–OFF processes. A Bayesian formulation of the problem is developed within a decision-theoretic framework based on the theory of POMDP. The simple threshold structure of the optimal decision rules are established.

Future work includes investigating the optimality of the proposed low-complexity policy $\hat{\pi}_N$ in the finite regime and the asymptotic $(\zeta \longrightarrow 0)$ optimality of the threshold policy π^*_∞ for infinite ON–OFF processes with arbitrarily distributed busy and idle times. In this paper, we have focused on the Bayesian formulation of quickest detection in multiple ON–OFF processes. Whether a minimax type of formulation of the problem can be developed is open for discussion and interpretation. Note that in the classic quickest change detection that deals with a single change point, one can consider the worst-case detection delay over all possible values of the change point. This formulation does not directly extend to quickest detection in multiple ON–OFF processes with an infinite number of change points in each process. One possibility is to consider the minimization of the worst-case detection delay over all possible *distributions* of the busy and idle times. Whether this leads to a well formulated problem with tractable solutions is open for exploration.

APPENDIX A
PROOF OF LEMMA 1

It follows directly from (11) and (12) that $V_S(\lambda_t) = V_C(\lambda_0)$ and is independent of λ_t. The linearity of $V_D(\lambda_t)$ is obvious from (13).

The concavity and monotonicity of $V_C(\lambda_t)$ is proven by considering the finite horizon problem of length K, i.e., the user needs to declare within K units of time. Let $V^K(\lambda_t), V^K_C(\lambda_t), V^K_S(\lambda_t)$ denote the corresponding value functions. We have

$$\begin{aligned} V^K(\lambda_t) &= \min\{V^K_C(\lambda_t), V^K_S(\lambda_t), V_D(\lambda_t)\},\ K > 0 \\ V^0(\lambda_t) &= V_D(\lambda_t) = \gamma(1-\lambda_t). \end{aligned} \tag{28}$$

Furthermore, it is easy to see that $\forall \lambda_t \in [0,1]$,

$$V_C(\lambda_t) = \lim_{K\rightarrow\infty} V^K_C(\lambda_t).$$

It thus suffices to show that $V^K_C(\lambda_t)$ is concave and monotone for all K, which is proven by induction.

Consider first the concavity of $V^K_C(\lambda_t)$. The initial condition holds as shown below.

$$\begin{aligned} V^1_C(\lambda_t) &= 1 + \int_{-\infty}^{\infty} P(x;\lambda_t) V^0(\mathcal{T}(\lambda_t|x))dx \\ &= 1 + \int_{-\infty}^{\infty} P(x;\lambda_t)\gamma(1-\mathcal{T}(\lambda_t|x))dx \\ &= 1 + \int_{-\infty}^{\infty} \gamma[\bar{p}_B - (1-p_B-p_I)\lambda_t] f_0(x)dx \\ &= [1+\gamma\bar{p}_B] - \gamma(1-p_B-p_I)\lambda_t \end{aligned} \tag{29}$$

which is linear, thus concave. As a consequence, $V^1(\lambda_t)$ is the minimum of three linear functions [see (28)]. It is thus also concave.

Assume that $V^K_C(\lambda_t)$ and $V^K(\lambda_t)$ are concave. Then $V^K(\lambda_t)$ can be written as the minimum of infinitely many (potentially uncountable) linear functions of λ_t. With an abuse of notation, we index these linear functions by $i \in \mathbb{R}^+$, i.e., $\exists\ a_i, b_i \in \mathbb{R}$ such that

$$V^K(\lambda_t) = \min_{i\in\mathbb{R}^+}\{a_i + b_i\lambda_t\}. \tag{30}$$

We thus have

$$\begin{aligned} &V^{K+1}_C(\lambda_t) \\ &= 1 + \int_{-\infty}^{\infty} P(x;\lambda_t)V^K(\mathcal{T}(\lambda_t|x))dx \\ &= 1 + \int_{-\infty}^{\infty} P(x;\lambda_t)\min\{a_i + b_i\mathcal{T}(\lambda_t|x)\}dx \\ &= 1 + \int_{-\infty}^{\infty} \min\{a_iP(x;\lambda_t) + b_i[p_B + (1-p_B-p_I)\lambda_t]f_1(x)\}dx \\ &= 1 + \int_{-\infty}^{\infty} \min\{[(a_i+b_i)f_1(x)p_B + f_0(x)\bar{p}_B] \\ &\quad + [(a_i+b_i)f_1(x)(1-p_I-p_B) - f_0(x)(1-p_B-p_I)]\lambda_t\}dx. \end{aligned} \tag{31}$$

We can see that for any given x, the integrand in (31) is the minimum of linear functions of λ_t, thus concave in λ_t. It thus follows that $V^{K+1}_C(\lambda_t)$ is concave (this can be easily shown by the definition of concavity). Consequently, $V^{K+1}(\lambda_t)$ is concave from (28).

Next, we show the monotonicity of $V^K_C(\lambda_t)$ by induction. From (29) we know that $V^1_C(\lambda_t)$ is monotonically decreasing with λ_t when $p_B + p_I \leq 1$. The monotonicity of $V^1(\lambda_t)$ thus follows from (28) by noticing that $V_D(\lambda_t)$ is decreasing and $V^1_S(\lambda_t) = V^1_C(\lambda_0)$ is constant in λ_t.

Assume that $V^K_C(\lambda_t)$ and $V^K(\lambda_t)$ are monotonically decreasing with λ_t when $p_B + p_I \leq 1$. We now show the monotonicity of $V^{K+1}_C(\lambda_t)$ by considering its derivative.

$$\begin{aligned} &\frac{d}{d\lambda_t}V^{K+1}_C(\lambda_t) \\ &\quad = \frac{d}{d\lambda_t}\int_{-\infty}^{\infty} P(x;\lambda_t)V^K(\mathcal{T}(\lambda_t|x))dx && (32) \\ &\quad = \int_{-\infty}^{\infty} \frac{d}{d\lambda_t}[P(x;\lambda_t)V^K(\mathcal{T}(\lambda_t|x))]dx && (33) \\ &\quad = \int_{-\infty}^{\infty} \frac{d}{d\lambda_t}[P(x;\lambda_t)]V^K(\mathcal{T}(\lambda_t|x))dx \\ &\qquad + \int_{-\infty}^{\infty} P(x;\lambda_t)\frac{d}{d\lambda_t}[V^K(\mathcal{T}(\lambda_t|x))]dx && (34) \end{aligned}$$

where the exchange of derivative and integral in (33) follows from [30, Theorem 2.27].

Next, we show that both terms in (34) are non-positive, which leads to the monotonicity of $V_C^{K+1}(\lambda_t)$. Consider first the second term:

$$\frac{d}{d\lambda_t}V^K(\mathcal{T}(\lambda_t|x)) = \frac{d}{d\mathcal{T}(\lambda_t|x)}V^K(\mathcal{T}(\lambda_t|x))\frac{d}{d\lambda_t}\mathcal{T}(\lambda_t|x).$$

It can be shown from (8) that $d/d\lambda_t\mathcal{T}(\lambda_t|x) \geq 0$ when $p_B + p_I \leq 1$. By the induction assumption, $d/d\mathcal{T}(\lambda_t|x)V^K(\mathcal{T}(\lambda_t|x)) \leq 0$. The second term is thus non-positive.

To show that the first term is non-positive, we first prove the following lemma.

Lemma 3:

$$\inf_{x:f_1(x)\leq f_0(x)} V^K(\mathcal{T}(\lambda_t|x)) \geq \sup_{x:f_1(x)>f_0(x)} V^K(\mathcal{T}(\lambda_t|x)). \tag{35}$$

Proof: By the induction assumption, $V^K(\mathcal{T}(\lambda_t|x))$ is monotonically decreasing with $\mathcal{T}(\lambda_t|x)$. To prove (35), it is equivalent to prove that

$$\sup_{x:f_1(x)\leq f_0(x)} \mathcal{T}(\lambda_t|x) \leq \inf_{x:f_1(x)>f_0(x)} \mathcal{T}(\lambda_t|x)$$

which is equivalent to prove that

$$\inf_{x:f_1(x)\leq f_0(x)} \frac{1}{\mathcal{T}(\lambda_t|x)} \geq \sup_{x:f_1(x)>f_0(x)} \frac{1}{\mathcal{T}(\lambda_t|x)}. \tag{36}$$

From (8), it is easy to see that

$$\begin{aligned}\inf_{x:f_1(x)\leq f_0(x)} \frac{1}{\mathcal{T}(\lambda_t|x)} &= \inf_{x:f_1(x)\leq f_0(x)} \left\{1 + C\frac{f_0(x)}{f_1(x)}\right\}\\ &\geq 1 + C\\ \sup_{x:f_1(x)>f_0(x)} \frac{1}{\mathcal{T}(\lambda_t|x)} &= \sup_{x:f_1(x)>f_0(x)} \left\{1 + C\frac{f_0(x)}{f_1(x)}\right\}\\ &< 1 + C,\end{aligned}$$

where

$$C = \frac{\lambda_t p_I + \bar{\lambda}_t \bar{p}_B}{\lambda_t \bar{p}_I + \bar{\lambda}_t p_B}.$$

Thus, (36) follows, and we arrive at Lemma 3. ■

Since

$$\frac{d}{d\lambda}[P(x;\lambda)] = (1 - p_B - p_I)(f_1(x) - f_0(x)),$$

the first term of (34) can be written as, after omitting the positive constant $1 - p_B - p_I$,

$$\begin{aligned}&\int_{-\infty}^{\infty}(f_1(x) - f_0(x))V^K(\mathcal{T}(\lambda_t|x))dx\\ &= \int_{x:f_1(x)\leq f_0(x)}(f_1(x) - f_0(x))V^K(\mathcal{T}(\lambda_t|x))dx\\ &\quad + \int_{x:f_1(x)>f_0(x)}(f_1(x) - f_0(x))V^K(\mathcal{T}(\lambda_t|x))dx\\ &\leq \inf_{x:f_1(x)\leq f_0(x)} V^K(\mathcal{T}(\lambda_t|x))\\ &\quad \times \int_{x:f_1(x)\leq f_0(x)}(f_1(x) - f_0(x))dx\\ &\quad + \sup_{x:f_1(x)>f_0(x)} V^K(\mathcal{T}(\lambda_t|x))\\ &\quad \times \int_{x:f_1(x)>f_0(x)}(f_1(x) - f_0(x))dx\\ &\leq \inf_{x:f_1(x)\leq f_0(x)} V^K(\mathcal{T}(\lambda_t|x)) \int_{-\infty}^{\infty}(f_1(x) - f_0(x))dx = 0\end{aligned} \tag{37}$$

where (37) follows from Lemma 3.

The monotonicity of $V_C^{K+1}(\lambda_t)$ thus follows. The monotonicity of $V^{K+1}(\lambda_t)$ follows from (28). This completes the proof of Lemma 1.

Appendix B
Proof of Lemma 2

L2.2 follows from the homogeneity of the channels and the sufficiency of the belief statistics. L2.3 follows directly from (20). Similar to the proof for Lemma 1, L2.1 is proven by considering the finite horizon problem with length K. Specifically, we show by induction that when $p_B + p_I \leq 1$, $V_{C_i}^K(\lambda_1, \lambda_2)$ are concave and monotonically decreasing with λ_i for all i.

Consider first the concavity of $V_{C_i}^K(\lambda_1, \lambda_2)$. Without loss of generality, we consider $V_{C_1}^K(\lambda_1, \lambda_2)$.

When $K = 1$, we have

$$\begin{aligned}&V_{C_1}^1(\lambda_1, \lambda_2)\\ &= 1 + \int_{-\infty}^{\infty} P(x;\lambda_1)V^0(\mathcal{T}(\lambda_1|x), \mathcal{T}(\lambda_2))dx\\ &= 1 + \int_{-\infty}^{\infty} P(x;\lambda_1)\min\{\gamma(1 - \mathcal{T}(\lambda_1|x)), \gamma(1 - \mathcal{T}(\lambda_2))\}dx\\ &= 1 + \int_{-\infty}^{\infty} \min\{\gamma(1 - \mathcal{T}(\lambda_1))f_0(x), \gamma P(x;\lambda_1)(1 - \mathcal{T}(\lambda_2))\}dx.\end{aligned} \tag{38}$$

It is easy to see that for any given x, the integrand of (38) is the minimum of two linear functions of λ_1 when λ_2 is given (and vise versa). The concavity of $V_{C_1}^1(\lambda_1, \lambda_2)$ thus follows. Consequently, $V^1(\lambda_1, \lambda_2) = \min\{V_{D_1}(\lambda_1), V_{D_2}(\lambda_2), V_{C_1}(\lambda_1, \lambda_2), V_{C_2}(\lambda_1, \lambda_2)\}$ is concave in λ_i for all i.

Assume that $V_{C_1}^K(\lambda_1, \lambda_2)$ and $V^K(\lambda_1, \lambda_2)$ are concave in λ_i for all i. Similar to the proof in Lemma 1, the concavity of $V_{C_1}^{K+1}(\lambda_1, \lambda_2)$ can be shown by writing $V^K(\lambda_1, \lambda_2)$ as the minimum of infinitely many linear functions of λ_i [see (30)]. Consequently, $V^{K+1}(\lambda_1, \lambda_2)$ is concave in λ_i for all i.

Consider now the monotonicity of $V_{C_1}^K(\lambda_1, \lambda_2)$. When $K = 1$, we have

$$V_{C_1}^1(\lambda_1, \lambda_2) = 1 + \int_{-\infty}^{\infty} P(x;\lambda_1)V^0(\mathcal{T}(\lambda_1|x), \mathcal{T}(\lambda_2)) \tag{39}$$

where $V^0(\lambda_1,\lambda_2)=\min\{\gamma(1-\lambda_1),\gamma(1-\lambda_2)\}$, is monotonically decreasing with λ_1 for any given λ_2.

Since $P(x;\lambda_1)$ and $\mathcal{T}(\lambda_1|x)$ in (39) have the same form as $P(x;\lambda_t)$ and $\mathcal{T}(\lambda_t|x)$ in the case with infinite processes, the monotonicity of $V^1_{C_1}(\lambda_1,\lambda_2)$ in λ_1 can be proved in the same way as given in Lemma 1 [see (32)] by checking the derivative of $V^1_{C_1}(\lambda_1,\lambda_2)$ with respect to λ_1.

To show the monotonicity of $V^1_{C_1}(\lambda_1,\lambda_2)$ in λ_2, consider

$$\begin{aligned}\frac{\partial V^1_{C_1}(\lambda_1,\lambda_2)}{\partial\lambda_2}&=\frac{\partial}{\partial\lambda_2}\int_{-\infty}^{\infty}P(x;\lambda_1)V^0(\mathcal{T}(\lambda_1|x),\mathcal{T}(\lambda_2))dx\\&=\int_{-\infty}^{\infty}P(x;\lambda_1)\frac{\partial}{\partial\lambda_2}(V^0(\mathcal{T}(\lambda_1|x),\mathcal{T}(\lambda_2)))dx\\&=\int_{-\infty}^{\infty}P(x;\lambda_1)\frac{\partial V^0(\mathcal{T}(\lambda_1|x),\mathcal{T}(\lambda_2))}{\partial\mathcal{T}(\lambda_2)}\frac{\partial\mathcal{T}(\lambda_2)}{\partial\lambda_2}\\&\le 0\end{aligned}$$

where the last inequality follows from the monotonicity of $V^0(\lambda_1,\lambda_2)$ in λ_2 and $\partial\mathcal{T}(\lambda_2)/\partial\lambda_2\ge 0$. Consequently, we have the monotonicity of $V^1(\lambda_1,\lambda_2)$ in λ_i for all i.

Assume that $V^K_{C_1}(\lambda_1,\lambda_2)$ and $V^K(\lambda_1,\lambda_2)$ are monotonically decreasing in λ_i for all i. Then we can prove the statement for $K+1$ by following the same procedure as in proving the case of $K=1$.

APPENDIX C
PROOF OF THEOREM 2

The symmetric structure of the optimal policy follows directly from L2.2. The concavity and monotonicity of $V_{C_i}(\lambda_1,\lambda_2)$ given in L2.1 and the linearity of $V_{D_i}(\lambda_1,\lambda_2)$ in L2.3 lead to the unique interface between actions C and D given in (21) and illustrated in Fig. 6.

To see the monotonicity of $\eta_d(\lambda)$, we take $\eta_d(\lambda_2)$ for example and prove its monotonicity by contradiction.

Let $\eta_i(\lambda_2)$ $(i=1,2)$ be the intersecting curve between $V_{C_i}(\lambda_1,\lambda_2)$ and $V_{D_1}(\lambda_1)$. We thus have

$$\eta_d(\lambda_2)=\max\{\eta_1(\lambda_2),\ \eta_2(\lambda_2)\}.$$

Therefore, it is sufficient to prove that both $\eta_1(\lambda_2)$ and $\eta_2(\lambda_2)$ are monotonically increasing in λ_2. We prove this property for $\eta_1(\lambda_2)$; the proof for $\eta_2(\lambda_2)$ is similar.

Assume that $\eta_1(\lambda_2)$ is not monotonically increasing in λ_2. Then there exist two points (λ_1',λ_2') and $(\lambda_1'',\lambda_2'')$ on $\eta_1(\lambda_2)$ such that $\lambda_2''>\lambda_2'$ and $\lambda_1''<\lambda_1'$. Since $\eta_1(\lambda_2)$ is the intersecting curve, we have

$$V_{C_1}(\lambda_1',\lambda_2')=V_{D_1}(\lambda_1') \quad (40)$$
$$V_{C_1}(\lambda_1'',\lambda_2'')=V_{D_1}(\lambda_1''). \quad (41)$$

Consider the point (λ_1'',λ_2'). Recall that (λ_1',λ_2') is a point on the intersecting curve $\eta_1(\lambda_2)$. Thus, for the fixed λ_2', the concave function $V_{C_1}(\lambda_1,\lambda_2')$ of λ_1 and the linear function $V_{D_1}(\lambda_1)$ have a unique intersecting point at $\lambda_1=\lambda_1'$. Since the end point $V_{C_1}(0,\lambda_2')$ is smaller than the end point $V_{D_1}(0)$, we know that at $\lambda_1=\lambda_1''$ which is smaller than the intersecting point λ_1', $V_{C_1}(\lambda_1,\lambda_2')$ lies below $V_{D_1}(\lambda_1)$, i.e.,

$$V_{C_1}(\lambda_1'',\lambda_2')<V_{D_1}(\lambda_1'') \quad (42)$$

which contradicts with the monotonicity of $V_{C_1}(\lambda_1,\lambda_2)$ stating that

$$V_{C_1}(\lambda_1'',\lambda_2')>V_{C_1}(\lambda_1'',\lambda_2'')=V_{D_1}(\lambda_1'').$$

APPENDIX D
PROOF OF THEOREM 3

When $p_B+p_I\le 1$, the underlying two-state Markov chain governing each ON–OFF process is positively correlated. For any $\lambda<\lambda_0$, we have

$$\lambda<\mathcal{T}(\lambda)<\lambda_0,$$

i.e., if a process with belief smaller than λ_0 is not observed, then its belief will increase monotonically but can never exceed λ_0 (see Lemma 1 in [31] for a detailed proof with the following notation exchanges: $\lambda_0=\omega_o$, $\lambda=\omega$, $p_B=p_{01}$, $p_I=p_{10}$). Under $\hat{\pi}_N$, since the initial belief of all processes is λ_0, the user abandons the current process when its belief drops below λ_0. When $N\longrightarrow\infty$, the user will not revisit an abandoned process since its belief is always smaller than that of a new process. The actions taken under $\hat{\pi}_N$ thus converge to that under π^*_∞ as $N\longrightarrow\infty$.

ACKNOWLEDGMENT

The authors would like to thank the associate editor and the anonymous reviewers for their invaluable comments and suggestions.

REFERENCES

[1] W. Shewhart, *Economic Control of Quality of Manufactured Product*. New York: Van Nostrand, 1931.
[2] A. N. Shiryaev, "The problem of quickest detection of a violation of stationary behavior," *Dokl. Akad. Nauk SSSR*, vol. 138, pp. 1039–1042, 1961.
[3] A. N. Shiryaev, "The problem of the most rapid detection of a disturbance in a stationary process," *Sov. Math., Dokl.*, vol. 2, pp. 795–799, 1961.
[4] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory Prob. Appl.*, vol. 8, pp. 22–46, Jun. 1963.
[5] S. W. Roberts, "A comparison of some control chart procedures," *Technometrics*, vol. 8, pp. 411–430, 1966.
[6] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic Bayesian theory of quickest change detection," *Theory Prob. Appl.*, vol. 49, no. 3, pp. 458–497, 2005.
[7] A. A. Borovkov, "Asymptotically optimal solutions in the change-point problem," *Theory Prob. Appl.*, vol. 43, no. 4, pp. 539–561, Oct. 1999.
[8] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, pp. 1897–1908, 1971.
[9] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.
[10] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Statist.*, vol. 14, no. 4, pp. 1379–1387, 1986.
[11] M. Pollak, "Optimal detection of a change in distribution," *Ann. Statist.*, vol. 13, no. 1, pp. 206–227, 1985.
[12] A. S. Polunchenko and A. G. Tartakovsky, "On optimality of the Shiryaev–Roberts procedure for detecting a change in distribution," *Ann. Statist.*, unpublished.

[13] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, pp. 79–89, May 2007.

[14] C. H. Papadimitrou and J. N. Tsitsiklis, "The complexity of Markov decision processes," *Math. Oper. Res.*, vol. 12, no. 3, pp. 441–450, 1987.

[15] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

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

[17] A. G. Tartakovsky and V. V. Veeravalli, "Asymptotically optimal quickest change detection in distributed sensor systems," *Sequent. Anal.*, vol. 27, no. 4, pp. 441–475, 2008.

[18] T. He, S. Ben-David, and L. Tong, "Nonparametric change detection and estimation in large-scale sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1204–1217, Apr. 2006.

[19] A. Isaac, P. Willett, and Y. Bar-Shalom, "Quickest detection and tracking of spawning targets using monopulse radar channel signals," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1302–1308, Mar. 2008.

[20] A. Betran-Martinez, O. Simeone, and Y. Bar-Ness, "Detecting primary transmitters via cooperation and memory in cognitive radio," in *Proc. CISS*, Mar. 2007, pp. 369–369.

[21] H. Li, C. Li, and H. Dai, "Quickest spectrum sensing in cognitive radio," in *Proc. CISS*, Mar. 2008, pp. 203–208.

[22] A. N. Shiryaev, "On detecting of disorders in industrial processes: I," *Theory Probab. Appl.*, vol. 8, no. 3, pp. 264–281, 1963.

[23] A. N. Shiryaev, "On detecting of disorders in industrial processes: II," *Theory Probab. Appl.*, vol. 8, no. 4, 1963.

[24] J. R. Norris, *Markov Chains*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[25] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Oper. Res.*, vol. 21, pp. 1071–1088, 1973.

[26] J. D. Isom, S. P. Meyn, and R. D. Braatz, "Piecewise linear dynamic programming for constrained POMDPs," in *Proc. AAAI Conf. Artificial Intelligence*, 2008.

[27] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic," in *Proc. AXM SIGCOMM'93*, 1993, pp. 183–193.

[28] K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation*. New York: Wiley, 2000.

[29] I. Mitrani, *Probabilistic Modelling*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[30] G. Folland, *Real Analysis: Modern Techniques and Their Applications*. New York: Wiley-Interscience, 1999.

[31] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, Nov. 2010.

Qing Zhao (S'97–M'02–SM'08) received the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, in 2001.

In August 2004, she joined the Department of Electrical and Computer Engineering at the University of California at Davis (UC Davis), where she is currently an Associate Professor. Her research interests are in the general area of dynamic systems and communication networks.

Dr. Zhao holds the title of UC Davis Chancellor's Fellow. She received the 2000 Young Author Best Paper Award from IEEE Signal Processing Society and the 2008 Outstanding Junior Faculty Award from the UC Davis College of Engineering.

Jia Ye received the B.S. degree in electrical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2004 and the M.S. degree in electrical engineering from Fudan University, Shanghai, China, in 2007. She is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, University of California at Davis.

Her research interests are in signal processing in wireless networks and stochastic optimization in dynamic systems.