Quickest Wideband Spectrum Sensing Over Correlated Channels

Ali Tajer, *Senior Member, IEEE*, and Javad Heydari

***Abstract*—Quickest spectrum sensing seeks to optimize a balance between two opposing performance measures, one being the delay in identifying spectrum opportunities, and the other being the quality of the decision. The existing spectrum sensing approaches formed based on quickest detection theory rely on the assumption that the occupancy states of different spectrum bands over a wideband spectrum are statistically independent. This is an assumption that cannot be met in practice, especially in broadband communication schemes in which radio channels are dynamically grouped in bundles and allocated to different users based on the users' traffic needs. As a result of such channel grouping and allocation, the occupancy states of the channels, especially the adjacent ones, are correlated. This paper, in contrast to the existing literature on quickest spectrum sensing, considers a wideband spectrum in which the occupancy states of different channels follow a pre-specified dependency kernel. The objective is to design the *quickest* spectrum sensing approach for identifying spectrum holes, which aims to minimize the average delay in identifying spectrum opportunities while assuring, in parallel, certain guarantees on the quality of the decision. The closed-form characterization of the optimal sensing scheme is delineated and it is shown that this optimal scheme has low computational complexity.**

***Index Terms*—Quickest detection, correlated channels, wideband spectrum, spectrum sensing.**

I. Introduction

EVER increasing demands for data traffic necessitates for more efficient utilization of the spectrum and leads to reduced availability of under-utilized spectrum bands, which is expected to become even more severe as wireless networks grow in size and in their traffic loads. Spectrum opportunities, as a result, become more *scarce* and will be dynamically scattered throughout the entire frequency spectrum. Besides scarcity, the occupancy statuses of the spectrum holes also vary rapidly and the spectrum holes might not remain unoccupied for a long duration. Therefore, it is of paramount importance to devise mechanism that can identify the spectrum holes *quickly*, as any delay in identifying the spectrum holes leads to under-utilization of the spectrum and reduced spectrum efficiency. Quickest spectrum sensing aims to devise a sensing mechanism that strikes an optimal balance between sensing *agility* and *reliability*, as two important performance measures, where improving one penalizes the other one.

Categorically, there exist two research directions addressing the interplay between agility and reliability in spectrum sensing. In one direction the objective is to form an estimate about the occupancy state of the *entire* wideband spectrum via estimating the power spectral density of a the wideband channel [1]–[7]. In this category, it is assumed that the wideband channel is heavily under-utilized, and its sparsity structure is leveraged to construct compressed sensing-based machinery for estimating the power spectral density (PSD) of the wideband channel. Exploiting the sparsity empowers the spectrum sensing procedure to sample the signal activity over the wideband spectrum at a sub-Nyquist rate, which expedites the process of estimating the PSD.

In a different direction, in contrast, the objective is to identify only *one* (or a few) vacant channel(s). Under this objective, estimating the entire spectrum might be redundant and penalizes the agility in identifying one spectrum opportunity. In this direction, data-adaptive and sequential sampling strategies are deployed. Specifically, in the quickest sequential search approach of [8] and [9], the wideband spectrum is split into smaller narrowband channels and the users scan these channels sequentially one at-a-time. Upon scanning and accumulating enough information about each channel a wireless user decides whether the channel is vacant or occupied. If the channel is determined to be a vacant one, the search is terminated and otherwise the process is carried on until a vacant channel is detected. In another relevant direction, data-adaptive spectrum sensing procedures are devised in which a thresholding-based approach is proposed. In these approaches the channels with measurements that do not satisfy a threshold criteria are discarded recursively. The threshold is designed based on the statistical distributions of the measurements from the vacant and occupied channels [10] and [11].

This paper focuses on the second direction, in which the objective is to identify only *one* (or a few) channel(s) with one major distinction in the structure of occupancy statuses. The existing approaches on *quickest* spectrum sensing, irrespective of their discrepancies in model and objective, all conform in the assumption that different narrowband channels have statistically *independent* occupancy states [9]–[11]. This assumption cannot be met in practice, especially in broadband communication systems, in which the occupancy states of the narrowband channels, especially the adjacent ones, are co-dependent. In this paper, in contrast, we focus on such scenarios that there exists a dependency structure among the occupancy states of different channels. This is particularly of interest in broadband

Manuscript received November 23, 2014; revised March 1, 2015 and April 28, 2015; accepted June 20, 2015. Date of publication June 30, 2015; date of current version September 3, 2015. This research was supported in part by the U. S. National Science Foundation under Grant ECCS-1455228. The associate editor coordinating the review of this paper and approving it for publication was H. Li.

The authors are with the Department of Electrical, Computer, and System Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: tajer@ecse.rpi.edu).

Digital Object Identifier 10.1109/TCOMM.2015.2451101

communication schemes, such as orthogonal frequency division multiplexing (OFDM), in which a user gets access to a number of channels simultaneously. In such systems the channels are bundled together and assigned to different users dynamically by the network operator based on the instantaneous traffic needs of different users. Hence, the occupancy state of each channel is *not independent* of the rest of the channels. In other words, a channel being deemed as occupied provides some side information about the occupancy states of its adjacent channels.

Current studies on spectrum sensing in cognitive radio networks are primarily focused on the settings in which the occupancy states of the channels are independent (c.f. [12], [13]) with more recent developments on modeling and incorporating correlation structures among occupancy states in wideband channels [14]–[18]. The existing approaches address different sensing objectives (e.g., network throughput and sensing duration) with different correlation structures (e.g., Bayesian network and regression models). However, irrespective of their discrepancies in objectives and approaches, they all conform in the assumption that each channel (or spectrum segment) is sensed for a pre-specified *fixed* amount of time such that the frequencies of erroneous decisions (i.e., miss-detection and false alarm) are controlled below target levels. This is in sharp contrast with the search method proposed in this paper, in which the amount of time spent on each channel is data-adaptive, and are designed to minimize the duration of the search process.

II. PRELIMINARIES

A. Spectrum Model

Consider a wideband spectrum to be shared by multiple users with an interference-avoiding spectrum access policy. At any given time a group of users are actively accessing the spectrum. We assume that the available spectrum consists of n non-overlapping channels, indexed by $\{1,\ldots,n\}$. At a given instance, the active users communicate over some of these channels, which we refer to as *occupied channels*, and underutilize the rest, which we call *spectrum holes*. We adopt a probabilistic model for the occupancy status of the spectrum. Let the Bernoulli random variable Z_i, for $i \in \{1,\ldots,n\}$, indicate the occupancy state of the i^{th} channel, where $Z_i = 1$ means that the i^{th} channel is occupied and $Z_i = 0$ states that the i^{th} channel is a spectrum hole. Let us also denote the set of the indices of the spectrum holes and occupied channels by

$$\mathcal{H}_0 \overset{\triangle}{=} \{i \in \{1,\ldots,n\} : Z_i = 0\}, \tag{1}$$

$$\mathcal{H}_1 \overset{\triangle}{=} \{i \in \{1,\ldots,n\} : Z_i = 1\}. \tag{2}$$

In multi-tone and broadband communication schemes, e.g., OFDM, users access a bundle of channels. Therefore, the occupancy states of adjacent channels are *not* necessarily independent as there is a chance that they are bundled and assigned to certain users. Hence, we assume that the there exists a correlation structure in the occupancy states of adjacent channels.

A user seeking spectrum opportunity scans the spectrum via sequentially tuning its receivers filters to different channels. It examines the channels sequentially until it identifies a channel as a spectrum hole with sufficient confidence (formal formulation is provided in Section IV). We consider an *ordered* set of the channels, and without loss of generality consider the following dependency kernel between the occupancy states of adjacent channels. For any $i \in \{2,\ldots,n\}$:

$$\mathbb{P}(Z_i = 0 | Z_{i-1} = j) = \epsilon_j \quad \text{for} \quad j \in \{0,1\}, \tag{3}$$

where $\epsilon_0, \epsilon_1 \in (0,1)$. In other words, if a channel is vacant (busy), the probability that the following adjacent channel is vacant is ϵ_0 (ϵ_1). When channels are allocated in bundles, compared to a vacant channel, a busy channel is more likely to have an adjacent busy channel. By noting that the probability that a busy channel is followed by another busy one is $1-\epsilon_1$ and the probability that a vacant channel is followed by a busy one is $1-\epsilon_0$, the setting of interest for this dependency kernel is $\epsilon_1 < \epsilon_0$.

We remark that setting $\epsilon_0 = \epsilon_1$ reduces the setting to one in which the occupancy states of different channels are statistically independent studied in [8]. It is noteworthy that the correlation model in (3) is adopted to focus the attention on the pivotal assumption that occupancy states are correlated. The analysis provided can be generalized to address any arbitrary correlation structure, albeit at the cost of higher dimensions for parameterization. Finally we assume that channel 1 is a vacant channel with prior probability ϵ, i.e.,

$$\mathbb{P}(Z_1 = 0) = \epsilon. \tag{4}$$

The correlation structure is studied in [15] and [16]. In particular [15] investigates the correlation structures among the occupancy states of different channels over time, frequency, and space in a more general form. Specifically, it adopts a Bayesian random graphical model in which the nodes capture the occupancy states of channels (over time, frequency, and space) and the edges represent the mutual effects of the occupancy states of different channels on each other. Based on this adopted model and by leveraging the historical data about the spectral activities of the channels accumulated over time, the parameters that describe the transition probabilities among different nodes are estimated. The correlation model adopted in our setting is a special case of the more general correlation structure analyzed in [15], and as a result, the correlation parameters ϵ_0 and ϵ_1 can be estimated by following the guidelines in [15].

B. Measurement Model

A spectrum-seeking user forms decisions about the occupancy states of different channels by capitalizing on the information collected via channel measurements. Sequence of measurements from channel i is denoted by $\mathcal{X}^i \overset{\triangle}{=} \{X_1^i, X_2^i, \ldots\}$. Each sequence $\mathcal{X}^i$ (i.e., measurements from channel i) consists of independent and identically distributed (i.i.d.) observations $\mathcal{X}^i \overset{\triangle}{=} \{X_1^i, X_2^i, \ldots\}$ taking real values endowed with a σ-field of observations that obey the following dichotomous hypothesis model:

$$\begin{aligned} \mathsf{H}_0 &: X_j^i \sim F_0, \quad j = 1,2,\ldots \\ \mathsf{H}_1 &: X_j^i \sim F_1, \quad j = 1,2,\ldots \end{aligned} \tag{5}$$

where F_0 and F_1 denote the cumulative distribution functions (cdfs) of two distinct distributions on $(\mathbb{R}, \Omega)$. Distribution F_0 captures the statistical behavior of the measurements from vacant channels and F_1 models the statistical behavior of the measurements from busy channels. The probability density functions (pdfs) corresponding to F_0 and F_1 are denoted by f_0 and f_1, respectively.

III. Quickest Spectrum Sensing

A. Sensing Model

With the ultimate objective of identifying *one* vacant channel (its measurements generated by F_0), the sampling procedure examines the ordered sequences $\{\mathcal{X}^1, \dots, \mathcal{X}^n\}$ sequentially by taking one measurement at-a-time from one sequence until a reliable decision can be formed. Specifically, by defining s_t as the index of the sequence observed at time $t = 1, 2, \dots$ and denoting the sample taken at time t by Y_t, the information accumulated sequentially can be abstracted by the filtration $\{\mathcal{F}_t | t = 1, 2, \dots\}$, where

$$\mathcal{F}_t \stackrel{\triangle}{=} \sigma(Y_1, \dots, Y_t). \tag{6}$$

By initiating the sampling procedure by setting $s_1 = 1$ and based on the information accumulated up to time t, i.e., $\mathcal{F}_t$, the sampling procedure at time t takes one of the following actions.

- A_1) *Detection*: stops further sampling and identifies channel s_t as a vacant channel for which the measurements are generated according to F_0;
- A_2) *Observation*: due to lack of sufficient confidence making any decision is deferred and one more observation is taken from the same channel, i.e., $s_{t+1} = s_t$; or
- A_3) *Exploration*: channel s_t is deemed to be busy (measurements generated by F_1) and it is discarded; the sampling procedure switches to the next channel and takes one observation from the new channel, i.e., $s_{t+1} = s_t + 1$.

In order to formalize the sampling procedure we define τ as the stopping time of the procedure, that is the time after which no further measurement is taken and detection (action A_1) is performed. Furthermore, for modeling the dynamic decisions between observation and exploration actions we define the switching function $\psi : \{1, \dots, \tau - 1\} \to \{0, 1\}$. At time $t \in \{1, \dots, \tau - 1\}$, if the decision is in favor of performing observation (A_2) we have $\psi(t) = 0$, while $\psi(t) = 1$ indicates a decision in favor of exploration (A_3), i.e., $\forall t \in \{1, \dots, \tau - 1\}$:

$$\psi(t) = \begin{cases} 0 & \text{action } A_2 \text{ and } s_{t+1} = s_t \\ 1 & \text{action } A_3 \text{ and } s_{t+1} = s_t + 1. \end{cases} \tag{7}$$

Finally, we remark that in this paper we assume that the number of sequences $n \to \infty$ to emphasize that the spectrum is wideband, which in turn ensures that one vacant channel can be identified almost surely.

B. Problem Formulation

Designing sequential sampling procedures involves a tension between two opposing performance measures, one being the delay in reaching a decision and the other one being the frequency of erroneous decisions. Clearly improving one penalizes the other one and designing an optimal strategy rests on striking an appropriate balance between them. By defining $\mathcal{T}$ as the set of all possible stopping times and $\psi_\tau \stackrel{\triangle}{=} \{\psi(1), \dots, \psi(\tau - 1)\}$ as the sequence of switching functions up to the stopping time τ, the optimal sequential sampling strategy can be crafted as the solution to the following optimization problem, which minimizes the *average* delay in reaching a decision subject to a controlled quality for the decision, i.e.,

$$\begin{array}{ll} \inf_{\tau \in \mathcal{T}, \psi_\tau} & \mathbb{E}[\tau] \\ \text{s.t.} & \mathbb{P}\left(Z_{s_\tau} = 1\right) \leq \beta. \end{array} \tag{8}$$

where β is the maximum tolerable probability of error in the final decision. We remark that this approach tolerates missing some of the spectrum holes throughout the search process in favor of identifying one spectrum hole in the quickest fashion. Hence, the formulation in (8) focuses only on the agility and quality of the decision about the channel identified as a spectrum hole and does not incorporate the qualities of the decisions about the channels that have been sensed and discarded prior to the stopping time. It is noteworthy that when the objective is to form a decision for every single channel with controlled miss-detection and false-alarm rates, the optimal search process with the smallest average delay consists of performing sequential probability ratio test (SPRT) on each of the channels. Nevertheless, SPRT is too conservative for our purposes and results in an average delay longer than the optimal process found by solving (8).

IV. Optimal Sampling Procedure

The sampling strategy is captured by the stopping time τ and the switching sequence ψ_τ, and determining their optimal values delineates the optimal sampling strategy. To proceed, we define π_t as the posterior probability that channel s_t is generated according to F_0. Hence, by this definition

$$\pi_t \stackrel{\triangle}{=} \mathbb{P}(Z_{s_t} = 0 | \mathcal{F}_t). \tag{9}$$

It can be readily shown that

$$\pi_1 = \frac{\epsilon f_0(Y_1)}{\epsilon f_0(Y_1) + (1 - \epsilon) f_1(Y_1)}. \tag{10}$$

Also, by invoking the statistical independent among the measurements *within* a sequence and the correlation structure between *consecutive* sequences, as delineated in (3), we can obtain the following recursive connection between π_{t+1} and π_t for $t \in \mathbb{N}$:

$$\begin{aligned} \pi_{t+1} = & \frac{\pi_t f_0(Y_{t+1})}{\pi_t f_0(Y_{t+1}) + (1 - \pi_t) f_1(Y_{t+1})} \cdot \mathbb{1}\left(\psi(t) = 0\right) \\ & + \frac{\bar{\pi}_t f_0(Y_{t+1})}{\bar{\pi}_t f_0(Y_{t+1}) + (1 - \bar{\pi}_t) f_1(Y_{t+1})} \cdot \mathbb{1}\left(\psi(t) = 1\right), \end{aligned} \tag{11}$$

where we have defined

$$\bar{\pi}_t \stackrel{\triangle}{=} \pi_t(\epsilon_0 - \epsilon_1) + \epsilon_1, \tag{12}$$

and $\mathbb{1}(\cdot)$ denotes the indicator function. As discussed in [19] and [20], the canonical quickest search optimization problem in (8) can be solved equivalently by solving the following Bayesian formulation:

$$\inf_{\tau\in\mathcal{T},\psi_\tau} \left[\mathbb{P}(Z_{s_\tau}=1)+c_\beta\mathbb{E}[\tau]\right], \tag{13}$$

where c_β is a constant that can be found uniquely as a function of β. The primary motivation for such transformation is that the problem cast in the variational form in (8) is not mathematically tractable, while transforming it in the Bayesian form above facilitates the solution. This Bayesian formulation incorporates the costs associated with the average delay and decision accuracy into one cost function. In this Bayesian formulation, making one more measurement penalizes the sampling cost ($c_\beta\mathbb{E}[\tau]$) by c_β, on one hand, and when the additional sample is used judiciously, it is expected to increase the quality of decision by reducing $\mathbb{P}(Z_{s_\tau}=1)$. The sequential sampling process is terminated when the additional cost of taking one more sample exceeds the gain in the expected reduction in error probability. The problem formulated in (13) can be solved recursively by using the dynamic programming argument. Accordingly, we can solve this problem step by step, and minimize the whole cost by minimizing the cost at each step. Therefore, we take into consideration all the actions that can be made at each time and determine the cost that each action incurs to the total Bayesian cost.

Based on the cost function in (13), we characterize the expected cost of deciding in favor of each of the detection, observation, and exploration actions at time t given the information accumulated up to time t. Specifically, given $\mathcal{F}_t$, the minimal cost-to-go at time t, denoted by $\tilde{G}_t(\mathcal{F}_t)$, is related to the costs associated with actions $\{\mathrm{A}_i\}_{i=1}^3$, denoted by $\{\tilde{J}_{t;i}(\mathcal{F}_t)\}_{i=1}^3$, according to:

$$\tilde{G}_t(\mathcal{F}_t)=\min\left\{\tilde{J}_{t;1}(\mathcal{F}_t),\, c_\beta+\min_{i\in\{2,3\}}\tilde{J}_{t;i}(\mathcal{F}_t)\right\}, \tag{14}$$

where, $\tilde{J}_{t;1}(\mathcal{F}_t)$ is the cost associated with the stopping at time t. Since at the stopping time we have a detection action, the associated cost at the stopping time is $1-\pi_t$. The second cost in (14) measures the cost of continuing sampling at time t. Therefore, this cost is the minimum of two costs in the form of observation and exploration costs, and we have the following recursive relationship between the minimal cost-to-go and action costs:

$$\mathrm{A}_1:\quad \tilde{J}_{t;1}(\mathcal{F}_t)=1-\pi_t \tag{15}$$

$$\mathrm{A}_2:\quad \tilde{J}_{t;2}(\mathcal{F}_t)=\mathbb{E}\left[\tilde{G}_{t+1}(\mathcal{F}_{t+1})|\mathcal{F}_t,\psi(t)=0\right] \tag{16}$$

$$\mathrm{A}_3:\quad \tilde{J}_{t;3}(\mathcal{F}_t)=\mathbb{E}\left[\tilde{G}_{t+1}(\mathcal{F}_{t+1})|\mathcal{F}_t,\psi(t)=1\right]. \tag{17}$$

Specifically, $\tilde{J}_{t;1}(\mathcal{F}_t)$ is the cost of stopping taking any further measurements and declaring channel s_t vacant (generated by F_0); $(c_\beta+\tilde{J}_{t;2}(\mathcal{F}_t))$ is the expected cost associated with taking one more sample from the same channel; and $(c_\beta+\tilde{J}_{t;3}(\mathcal{F}_t))$ is the expected cost pertinent to switching to a new channel. The next lemma establishes that the optimal decision rule for stopping and switching at time t is related to $\mathcal{F}_t$ only through π_t.

Lemma 1: Cost functions $\{\tilde{J}_{t;i}(\mathcal{F}_t)\}_{i=1}^3$ and $\tilde{G}_t(\mathcal{F}_t)$ depend on $\mathcal{F}_t$ only through π_t and can be cast as functions of π_t denoted by $\{J_{t;i}(\pi_t)\}_{i=1}^3$ and $G_t(\pi_t)$, respectively. Furthermore, the decision rules for stopping or switching at time t can be uniquely determined as functions of π_t.

Proof: See [21, Lemma 1]. □

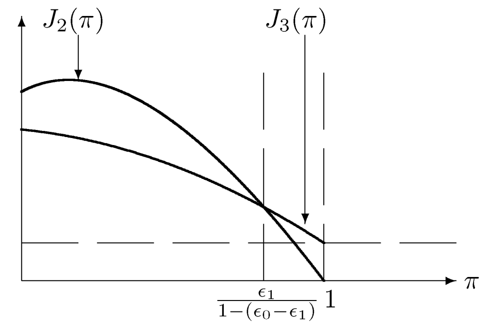


Fig. 1. $J_2(\pi)$ and $J_3(\pi)$ versus π when $\epsilon_0>\epsilon_1$.

The result in lemma above indicates that the information required for forming a decision at time t is embedded in the probability term π_t and the cost associated with taking any of the three possible actions can be uniquely determined by π_t. Hence, if for two different time instances $t\neq t'$ the sampling procedure encounters $\pi_t=\pi_{t'}=\pi$, the cost-to-go terms at these instances will become identical, i.e., $J_{t;i}(\pi)=J_{t';i}(\pi)$. In order to emphasize this equivalence, in the remainder of the paper we use the shorthand terms $J_i(\pi)$ and $G(\pi)$ for referring to $J_{t;i}(\pi)$ and $G_t(\pi)$, respectively.

The next lemma sheds light on the structures of functions $\{J_i(\pi)\}_{i=2}^3$, and is instrumental to delineating the optimal stopping time.

Lemma 2: Functions $J_2(\pi)$ and $J_3(\pi)$ are non-negative concave functions of π over (0,1).

Proof: See [21, Lemma 2]. □

Lemma 3: For functions $J_2(\pi)$ and $J_3(\pi)$ we have $J_2(0)>J_3(0)$ and $J_2(1)<J_3(1)$.

Proof: See Appendix A. □

Based on the results of lemmas 1, 2, and 3 the optimal stopping time and switching rules are characterized in the next subsections.

A. Optimal Switching Rule

The switching function defined in (7), which occurs prior to the stopping time, dynamically decides between performing the observation or exploration actions. For $t\in\{1,\dots,\tau-1\}$ such binary decision can be cast as

$$\psi(t)=\begin{cases}0 & \text{if}\quad J_3(\pi_t)\geq J_2(\pi_t)\\ 1 & \text{if}\quad J_3(\pi_t)<J_2(\pi_t),\end{cases} \tag{18}$$

where $\psi(t)=0$ models observation with decision cost $c_\beta+J_2(\pi_t)$ and $\psi(t)=1$ models exploration with the associated cost $c_\beta+J_3(\pi_t)$. Based on the continuity and concavity of functions $J_2(\pi)$ and $J_3(\pi)$ in π, and the fact that $J_2(0)>J_3(0)$ and $J_2(1)<J_3(1)$, it can be readily shown that $J_2(\pi)$ and $J_3(\pi)$ in the interval $\pi\in[0,1]$ intersect at exactly one point as depicted in Fig. 1. The intersection point and the optimal switching rule are determined by the following theorem.

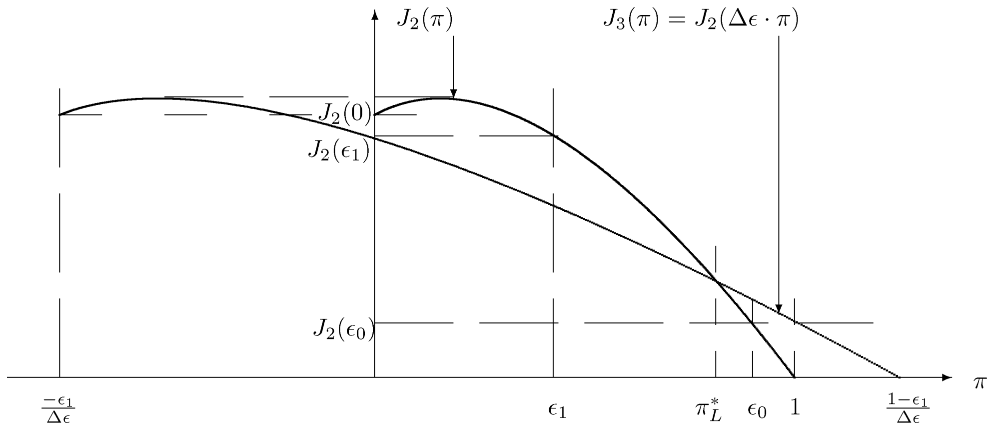


Fig. 2. $J_3(\pi)$ as an extended $J_2(\pi)$ from ϵ_1 to ϵ_0.

Theorem 1 (Switching Rule): For $t \in \{1, \dots, \tau - 1\}$, the cost functions $J_2(\pi)$ and $J_3(\pi)$ intersect at *exactly one* point, which is given by

$$\pi_L^* \triangleq \frac{\epsilon_1}{1 - (\epsilon_0 - \epsilon_1)}, \tag{19}$$

and the optimal switching rule at time t is

$$\psi(t) = \begin{cases} 0 & \pi_t \geq \pi_L^* \\ 1 & \pi_t < \pi_L^*. \end{cases} \tag{20}$$

Proof: The definitions of $J_2(\pi)$ and $J_3(\pi)$ given in (16) and (17) respectively, establish that $J_3(\pi) = J_2(\Delta\epsilon \cdot \pi + \epsilon_1)$ where we have defined $\Delta\epsilon \triangleq \epsilon_0 - \epsilon_1$. $J_2(\pi)$ and $J_3(\pi)$ clearly intersect when

$$\Delta\epsilon \cdot \pi + \epsilon_1 = \pi, \tag{21}$$

or equivalently when $\pi = \frac{\epsilon_1}{1-(\epsilon_0-\epsilon_1)}$. Next, we show that this intersection is unique. Fig. 2 demonstrates the connection between $J_2(\pi)$ and $J_3(\pi)$, where it can be easily verified that $J_2(\epsilon_1) = J_3(0) < J_2(0)$. Since $J_2(\pi)$ is concave and $J_2(\epsilon_1) < J_2(0)$, then $J_2(\pi)$ is monotonically decreasing for $\pi \in [\epsilon_1, 1]$. Therefore, $J_3(\pi)$ is monotonically decreasing in [0, 1]. From concavity, for every $\pi \in [0, \epsilon_1]$, we have

$$\begin{aligned} J_2(\pi) &= J_2\left(\frac{\epsilon_1 - \pi}{\epsilon_1} \cdot 0 + \frac{\pi}{\epsilon_1} \cdot \epsilon_1\right) \\ &\geq \frac{\epsilon_1 - \pi}{\epsilon_1} J_2(0) + \frac{\pi}{\epsilon_1} J_2(\epsilon_1) \\ &\overset{(a)}{>} J_2(\epsilon_1) \\ &= J_3(0) \\ &> J_3(\pi), \end{aligned} \tag{22}$$

where (a) holds since $J_2(0) > J_2(\epsilon_1)$. Hence according to (22) $J_2(\pi)$ and $J_3(\pi)$ do not intersect for $\pi \in [0, \epsilon_1]$.

Suppose that there is one intersection in $(\epsilon_1, 1]$ which we denote it by π_r. It implies that $J_2(\pi_r) = J_3(\pi_r) = J_2(\Delta\epsilon \cdot \pi_r + \epsilon_1)$, which means that there should be one $\pi_r^* = \Delta\epsilon \cdot \pi_r + \epsilon_1 \neq \pi_r$ such that $J_2(\pi_r) = J_2(\pi_r^*)$, but $J_2(\pi)$ is monotonic in $[\epsilon_1, 1]$, and for $\pi \in [0, \epsilon_1]$ we have $J_2(\pi) > J_2(\epsilon_1) > J_2(\pi_r)$. Consequently there cannot be such π_r and there is just one intersection at π_L^*. Hence, from Lemma 3 and the fact that $J_2(\pi)$ and $J_3(\pi)$ intersect only at π_L^* we have

$$\begin{cases} J_2(\pi) > J_3(\pi) & \text{if } \pi < \pi_L^* \\ J_2(\pi) < J_3(\pi) & \text{if } \pi > \pi_L^*, \end{cases} \tag{23}$$

and the optimal switching rule follows according to (18). □

This theorem demonstrates that the dynamic comparisons of the cost functions associated with observation and exploration actions can be reduced to comparing the posterior probability at each time with a constant pre-specified threshold, which is determined by the correlation structure. In other words, when the posterior probability π_t falls below π_L^*, the sampling strategy discards the current channel permanently and switches to the following one. This result subsumes the switching rule provided in [9] for the settings in which every channel is a vacant channel independently of the rest, i.e., $\epsilon_0 = \epsilon_1 = \epsilon$.

Corollary 1: When all channels are vacant independently of each other and with prior probabilities ϵ, the optimal switching rule at time $t \in \{1, \dots, \tau - 1\}$ is

$$\psi(t) = \begin{cases} 0 & \pi_t \geq \epsilon \\ 1 & \pi_t < \epsilon. \end{cases} \tag{24}$$

Given this switching sequence which characterizes the sampling strategy up to the stopping time, in the next step we identify the optimal stopping time, at which the sampling process is terminated and the last channel under scrutiny is declared to be a vacant channel.

Remark 1: One important observation from the structure of the optimal switching rule given in (20) is that this decision rule is independent of c_β. The significance of this observation is that we solved the quickest search problem (8) through solving an equivalent Bayesian formulation given in (13). These two equivalent forms of the quickest search problem are related through parameters β, which controls the decision quality in the variational form and c_β (as a function of β), which determines how the decision quality and decision delay should be integrated into one cost function in the Bayesian form. Based on the discussions in [19] the exact connection between β and c_β is unknown and only approximate connections can be established. The switching decisions being independent of c_β makes the design of the optimal sampling strategy independent of the unknown exact connection between β and c_β.

B. Optimal Stopping Time

According to the discussion in Remark 1, there are two approaches for characterizing the optimal stopping time by using the variational and the Bayesian formulations appropriate for scenarios in which we are interested in performance qualities with and without (respectively) randomized tests.

1) Bayesian Formulation: Based on the Bayesian cost function in (13), the sampling process terminates when the cost-to-go associated with stopping (detection action) falls below those associated with observation and exploration actions. In other words, the stopping time is the first time instance at which

$$J_1(\pi_t) < c_\beta + \min\{J_2(\pi_t), J_3(\pi_t)\}, \tag{25}$$

or equivalently, the optimal stopping time is

$$\tau^* = \inf\left\{t : J_1(\pi_t) < c_\beta + \min_{i\in\{2,3\}} J_i(\pi_t)\right\}. \tag{26}$$

By recalling that $J_1(\pi) = 1 - \pi$, and following the same line of argument as in [20, Th. 3.7], the optimal stopping time is given by the next theorem.

Theorem 2 (Stopping Time With Randomized Test): The optimal stopping time τ^* of the sampling process is

$$\tau^* = \inf\left\{t : \pi_t \geq \pi_U^*\right\}. \tag{27}$$

where π_U is the solution of

$$\pi_U^* = 1 - c_\beta - \min_{i\in\{2,3\}} J_i\left(\pi_U^*\right). \tag{28}$$

Proof: According to (14), the procedure stops taking further samples when the cost associated with terminating the procedure, $J_1(\pi_t) = 1 - \pi_t$, falls below the one associated with the observation action, $c_\beta + J_2(\pi_t)$, and the one associated with the exploration action, $c_\beta + J_3(\pi_t)$. In other words, from (14), the optimal stopping time is obtained by

$$\tau^* = \inf\left\{t : \pi_t \geq 1 - c_\beta - \min_{i=2,3} J_i(\pi_t)\right\}, \tag{29}$$

from which it is seen that the stopping rule can be equivalently cast as comparing the probability term π_t with a threshold. By following the same lines of argument as in [9] and [19] this threshold can be found as the solution of:

$$\pi_U^* = 1 - c_\beta - \min_{i\in\{2,3\}} J_i\left(\pi_U^*\right).$$

□

This result indicates that the stopping rule boils down to comparing the posterior probability π_t with a pre-specified threshold π_U^*, such that the sampling procedure is terminated when π_t exceeds this pre-specified threshold. The result of Theorem 1 in conjunction with that of Theorem 2 establishes that the optimal sampling strategy consists of dynamically comparing the posterior probability π_t, which is the probability that channel s_t is vacant, with known and constant lower and upper thresholds, and performing the detection and exploration actions when π_t exceeds, or falls below these upper and lower thresholds, respectively.

Characterizing π_U^* through solving (28) requires the exact knowledge of c_β, which as discussed earlier is unknown. By using the information about optimal structure of π_U^*, in the next subsection we characterize π_U^*.

2) Variational Formulation: Obtaining π_U^* through the Bayesian formulation ensures that the *average* error probability $\mathbb{P}(Z_{s_\tau} = 1)$, which is averaged over all ensembles of channel realizations, does not exceed the target tolerable error rate β. In certain scenarios, however, the network operator might be interested in ensuring a more stringent requirement on the error probability by enforcing that the error probability is kept below β for *every* single channel realization. This is especially of significance when avoiding co-channel interference is strictly enforced, in which case the network operator is interested in enforcing this for every channel realization. In such scenarios, the appropriate error probability requirement is $\mathbb{P}(Z_{s_\tau} = 1|\mathcal{F}_t)$.

By leveraging the insight gained from Theorem 2, which is the optimal stopping rule has a thresholding structure, we next provide the solution for π_U^* defined in (28) such that the error probability $\mathbb{P}(Z_{s_\tau} = 1|\mathcal{F}_t)$ is kept below β. The pivotal fact that we use is the opposing tension between the decision quality and the decision delay. Specifically, irrespective of what the exact value of π_U^* is, one can ascertain that higher values of π_U^*, which ensures higher decision quality, leads to increased decision delay. Hence, based on the threshold structure of the stopping rule, solving the functional formulation of the quickest search problem, which aims at minimizing the decision delay subject to decision quality, can be equivalently cast as the problem of obtaining the minimum value of π_U, for which the decision quality of interest is guaranteed. Hence, the problem (8) can be equivalently cast as

$$\pi_U^* = \begin{cases} \inf_{\tau\in\mathcal{T},\psi_\tau} & \pi_U \\ \text{s.t.} & \mathbb{P}\left(Z_{s_\tau} = 1|\mathcal{F}_t\right) \leq \beta. \end{cases} \tag{30}$$

On the other hand, by the definition of stopping rule, at the stopping time τ, for any hypothetical π_U we have

$$\pi_\tau \geq \pi_U. \tag{31}$$

Furthermore, according to the definitions of $\mathbb{P}(Z_{s_\tau} = 1|\mathcal{F}_t)$ and π_τ:

$$\mathbb{P}\left(Z_{s_\tau} = 1|\mathcal{F}_t\right) = 1 - \pi_\tau. \tag{32}$$

Equations (31) and (32) establish that

$$\pi_U^* \leq \begin{cases} \inf_{\tau\in\mathcal{T},\psi_\tau} & 1 - \mathbb{P}\left(Z_{s_\tau} = 1|\mathcal{F}_t\right) \\ \text{s.t.} & \mathbb{P}\left(Z_{s_\tau} = 1|\mathcal{F}_t\right) \leq \beta, \end{cases} \tag{33}$$

or equivalently $\pi_U^* \leq 1 - \beta$. Furthermore, it can be verified that π_U^* cannot be strictly smaller than $(1 - \beta)$ as otherwise in the unlikely but not impossible circumstances that π_t coincides with π_U^*, i.e., $\pi_t = \pi_U^* < 1 - \beta$, on one hand according to the stopping rule, the sampling process terminates at time t, and on the other hand

$$\mathbb{P}\left(Z_{s_\tau} = 1|\mathcal{F}_t\right) = 1 - \pi_t > \beta, \tag{34}$$

which violates the decision quality constraint. Hence, $\pi_U^* = 1 - \beta$, as formalized in the following theorem.

TABLE I
QUICKEST SEARCH ALGORITHM

1	set $t = 1$ and $s_1 = 1$
2	Take one sample from channel s_t
3	Update π_t according to (10) or (11).
4	If $\pi_L^* \leq \pi_t < \pi_U^*$
5	set $s_{t+1} = s_t$
6	$t \leftarrow t + 1$
7	Go to Step 2
8	Else if $\pi_L^* > \pi_t$
9	set $s_{t+1} = s_t + 1$
10	$t \leftarrow t + 1$
11	Go to Step 2
12	Else if $\pi_U^* \leq \pi_t$
13	set $\tau = t$
14	Declare channel s_τ vacant
15	End if

Theorem 3 (Stopping Time Without Randomized Test): The optimal stopping time τ^* of the sampling process is

$$\tau^* = \inf\left\{t : \pi_t \geq \pi_U^* = 1 - \beta\right\}. \tag{35}$$

Based on the results of Theorem 1 and Theorem 3, the steps involved in the optimal sampling strategy are summarized in Table I.

V. SIMULATIONS

A. Parameter Estimation

To the best of our knowledge, there exist two different approaches to determine the correlation parameters of the occupancy states of the channels, both hinging on the availability of the historical data. In the first approach, collected measurements from past are leveraged to compute empirical probabilities for dependency parameters or duty cycles, c.f., [22]–[25]. In this line of approaches, the data is fit into different models, such as Markov chains and autoregression in order to compute the parameters of the correlation structure. Implementing this approach for estimating correlation parameters can be done as follows. Bernoulli random variable Z_i stands for the occupancy state of channel i, and $Z_i = 0$ and $Z_i = 1$ correspond to the spectrum holes and occupied channels respectively. Assuming that for Z_i we have N_0 unoccupied channels and N_1 occupied ones for the case that $Z_{i-1} = 0$, and M_0 unoccupied channels and M_1 occupied ones when $Z_{i-1} = 1$, we can estimate ϵ_0 and ϵ_1 as

$$\epsilon_0 = \hat{\mathbb{P}}[Z_i = 0 | Z_{i-1} = 0] = \frac{N_0}{N_0 + N_1},$$
$$\epsilon_1 = \hat{\mathbb{P}}[Z_i = 0 | Z_{i-1} = 1] = \frac{M_0}{M_0 + M_1}. \tag{36}$$

The second direction deploys a statistical approach for estimating correlation parameters. In this approach, the correlation among occupancy states of the channel across time, space, and frequency is modeled by Bayesian network [15]. In such a Bayesian network finding the conditional probability densities especifies the model completely. In order to estimate the parameters of this model, Bayesian estimation can be used [15]. Specifically, a Beta distribution with parameters α and β as the prior distribution for both ϵ_0 and ϵ_1 is adopted. Having the same observations for Z_i as before, we have

$$\epsilon_0 = \frac{\alpha + N_0}{\alpha + \beta + N_0 + N_1},$$
$$\text{and} \qquad \epsilon_1 = \frac{\alpha + M_0}{\alpha + \beta + M_0 + M_1}, \tag{37}$$

where parameters α and β can be obtained from the historical data. As the size of observations increases, the quality of Bayesian estimates improves and converges to the empirical one.

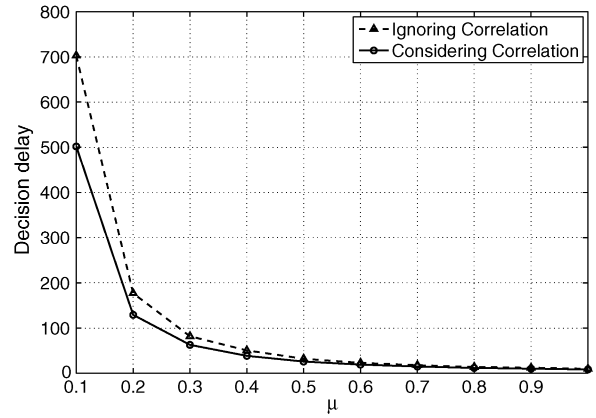


Fig. 3. Decision delay versus μ for $c = 0.001$.

B. Simulation Results

In this section, we assess the performance of the proposed spectrum sensing approach via simulations. The primary focus is placed on comparing the obtained optimal quickest search approach with the existing one in the literature which does not incorporate the correlation in the channel occupancy structure. Throughout the simulations we assume that $\epsilon = 0.2$, $\epsilon_0 = 0.6$ and $\epsilon_1 = 0.2$ as defined in (3) and (4). For selecting the appropriate models for distributions F_0 and F_1, we focus on two widely used settings. In the first setting we assume that $F_0 \sim N(0, \sigma^2)$ and $F_1 \sim N(\mu, \sigma^2)$ where $\mu > 0$ denotes the presence of an active user accessing the channel and $\sigma^2 = 1$.

Fig. 3 compares the decision delay of those two different approaches for a range of distances between F_0 and F_1. It is observed that using the information about the correlation between adjacent narrowband channels leads to a faster search, and the gains become more significant when the distance between F_0 and F_1 becomes smaller, which represents the situations in which spectrum holes are less distinguishable. Fig. 4 focuses on the same setting and compares the Bayesian costs. The results follow the same trend observed for the delay.

In Fig. 5, the trade-off between detection delay and the probability of a false alarm is shown. Demanding a more reliable detection of channel holes leads to increased number of observations needed. For a fixed target detection error, the test that considers the correlation structure between adjacent channels needs less observations of the channel, and with the same number of observations, that test has a more accurate decision on the unoccupied narrowband channels.

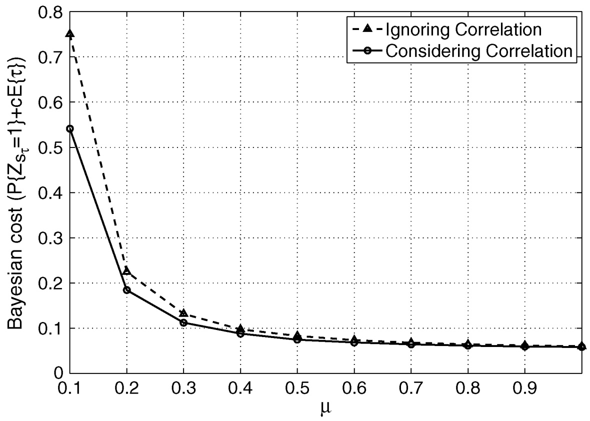


Fig. 4. Total Bayesian cost of decision versus μ for $c = 0.001$.

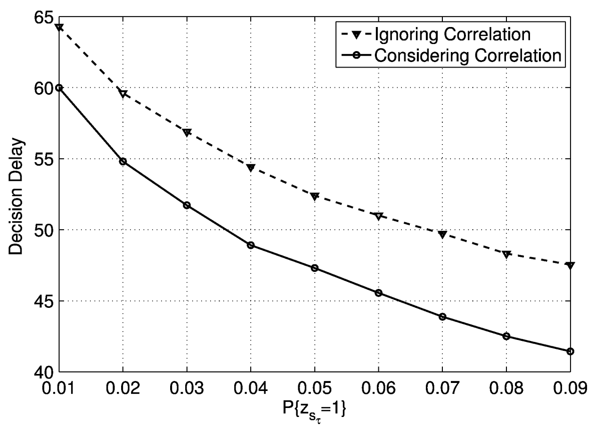


Fig. 5. Decision delay versus probability of error in detection.

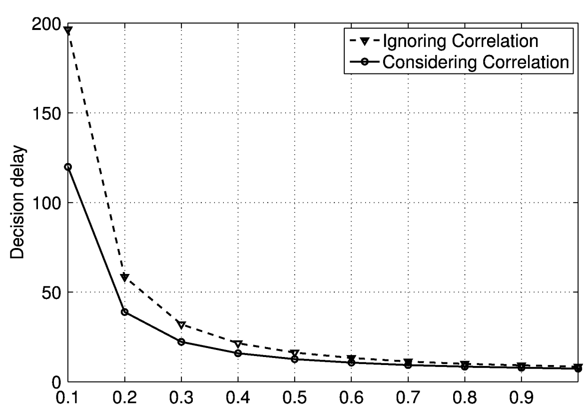


Fig. 6. Decision delay versus γ for $c = 0.01$.

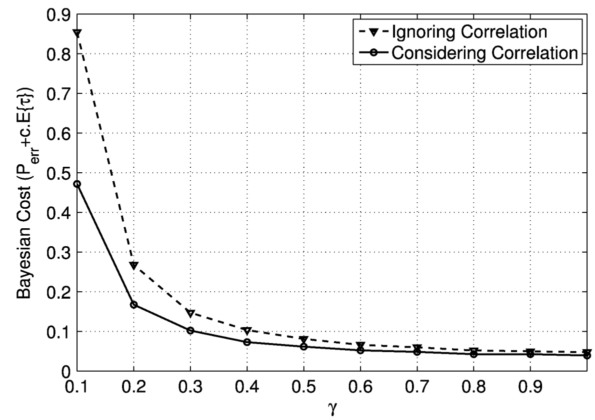


Fig. 7. Total Bayesian cost of decision versus γ for $c = 0.01$.

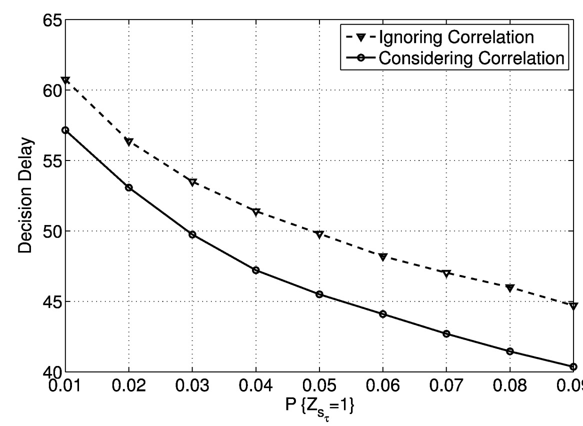


Fig. 8. Decision delay versus probability of error in detection.

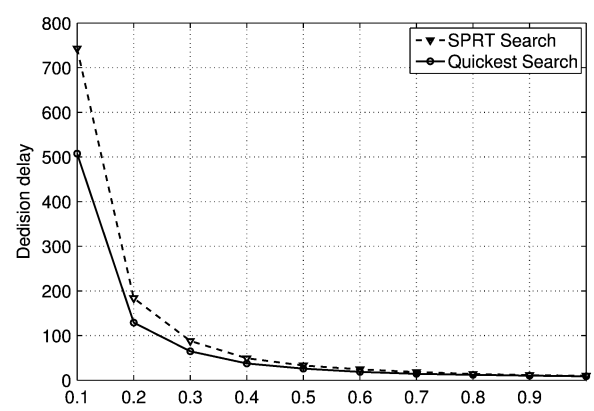


Fig. 9. Decision delay versus μ.

In the second setting we focus on the optimal maximum a posterior spectrum hole detection rules characterized in [10]. In this setting F_0 and F_1 are Gamma distributions with different scaling factors. By considering the power of all primary users to be γ, under H_1 the scaling factor would be $1+\gamma$, and it would be 1 under H_0.

Figs. 6–8 evaluate the performance of the proposed search method by applying the same assumptions about the priors and considering Gamma distributed observations. The results for this setting follow the same trend as that of Gaussian observations. When there are low power primary users, according to Fig. 6, the agility of the search will be improved by applying the dependency kernel of narrowband channels in the search. Bayesian cost which has been shown in Fig. 7 shows the same behavior as the number of observations. Again, the proposed test returns the spectrum holes with lower cost. Fig. 8 represents the trade-off between two performance measures.

In order to assess the gains of the optimal quickest search over performing sequential SPRTs on all narrow-band channels, Fig. 9 compares the decision delay for these two approaches for Gaussian observations. In the SPRT search we need to specify the maximum affordable probability of a miss to characterize the test completely; therefore we set $P_{\text{miss}} \leq 0.1$. The results show that quickest search leads to a shorter delay, especially when the Kullback-Leibler divergence between F_0 and F_1 is small (i.e, μ is small).

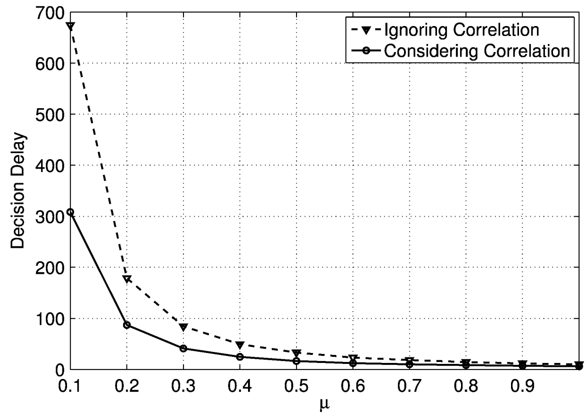


Fig. 10. Decision delay versus μ for $\beta = 0.05$.

Finally, in order to have an insight into the importance of correlation parameters, in Fig. 10, we adopted a new setting in which the values of ϵ_0 and ϵ_1 change across channels and in all cases $\epsilon_0 > \epsilon_1$. Distributions F_0 and F_1 are Gaussian with the same variance values ($\sigma^2 = 1$) and mean values 0 and μ, respectively, and $\beta = 0.05$. The results in this plot establish that the relevant performance of the two approaches compared follows the same observation in the setting in which the correlation parameters remain constant across different channels.

It is noteworthy that in the simulations we have not restricted the communication scheme deployed by the active users, and it can take any arbitrary form from narrowband (memory-less channels) to wideband (multipath fading). In the proposed sequential sensing framework, the spectrum seeking radio tunes its receiving filters to one of the narrowband channels, over which the energy level remains unchanged, and takes one measurement. If this channel is vacant the measurement will be primarily shaped by temporal noise. If, on the other hand, another user is utilizing the channel, we do not impose any constraint on the structure of the communication between this active user and the spectrum seeking radio. The only assumption made is that we know that this measurement when coming from noise follows distribution F_0 and when it comes from an active user is generated by F_1. The noise model and communication strategy govern F_0 and F_1.

VI. Conclusion

We have considered the problem of searching for spectrum opportunities in wideband channels, in which a channel being occupied or busy is not necessarily independent of the status of the rest of the channels. Specifically, it is assumed that the statuses of adjacent channels follow a known dependency kernel, representing a correlation among the distributions of measurements from the channels. The sequential sampling procedure dynamically decides to perform one of the three possible actions at each time, namely, the detection, observation, and exploration actions. The optimal decision rules for identifying the optimal action to be taken at each time are formulated and closed-form representations for these decision rules are characterized. Numerical evaluations are provided to demonstrate the gains of the optimal quickest sequential search approach over the alternative solutions that do not take into consideration the correlation structure among the distributions of the sequences.

Appendix A
Proof of Lemma 3

We prove $J_2(0) > J_3(0)$ by contradiction. Assume

$$J_2(0) \leq J_3(0). \tag{38}$$

From (14) and by setting $\pi_t = 0$ we obtain

$$\begin{aligned} G(0) &= \min\left\{1, c_\beta + \min\{J_2(0), J_3(0)\}\right\} \\ &\overset{(38)}{=} \min\left\{1, c_\beta + J_2(0)\right\}. \end{aligned} \tag{39}$$

On the other hand, from (16) we have

$$\begin{aligned} J_2(0) &= \mathbb{E}\left[G(\pi_{t+1})|\mathcal{F}_t, \psi(t) = 0\right] \\ &= \int G(0) f_1(Y_{t+1}) dY_{t+1} \\ &= G(0). \end{aligned} \tag{40}$$

From (39) and (40) we have $J_2(0) = \min\{1, c_\beta + J_2(0)\}$, which implies that $J_2(0) = 1$ and subsequently $G(0) = 1$. Furthermore, based on the definition of $G(\cdot)$ in (14) we have $G(1) = 0$ and $G(\pi) \leq 1 - \pi$. By noting that $G(\cdot)$ is also concave, the only possible choice for $G(\cdot)$ is $G(\pi) = 1 - \pi$. On the other hand,

$$\begin{aligned} J_3(0) &= \mathbb{E}\left[G(\pi_{t+1})|\mathcal{F}_t, \psi(t) = 1\right] \\ &= \int G\left(\frac{\epsilon_1 f_0(Y_{t+1})}{\epsilon_1 f_0(Y_{t+1}) + (1-\epsilon_1) f_1(Y_{t+1})}\right) \\ &\quad \times \left(\epsilon_1 f_0(Y_{t+1}) + (1-\epsilon_1) f_1(Y_{t+1})\right) dY_{t+1} \\ &= \int \left(1 - \frac{\epsilon_1 f_0(Y_{t+1})}{\epsilon_1 f_0(Y_{t+1}) + (1-\epsilon_1) f_1(Y_{t+1})}\right) \\ &\quad \times \left(\epsilon_1 f_1(Y_{t+1}) + (1-\epsilon_1) f_1(Y_{t+1})\right) dY_{t+1} \\ &= 1 - \epsilon_1, \end{aligned} \tag{41}$$

which contradicts the assumption $1 = J_2(0) \leq J_3(0) = 1 - \epsilon_1$ in (38). Next, we show that $J_2(1) < J_3(1)$. By using (16), we have

$$J_2(1) = \int G(1) f_0(Y_{t+1}) dY_{t+1} = G(1) = 0. \tag{42}$$

On the other hand, from (17) we have

$$\begin{aligned} J_3(1) &= \int G\left(\frac{\epsilon_0 f_0(Y_{t+1})}{\epsilon_0 f_0(Y_{t+1}) + (1-\epsilon_0) f_1(Y_{t+1})}\right) \\ &\quad \times \left(\epsilon_1 f_0(Y_{t+1}) + (1-\epsilon_1) f_1(Y_{t+1})\right) dY_{t+1} \\ &> \int G(0) \left(\epsilon_1 f_0(Y_{t+1}) + (1-\epsilon_1) f_1(Y_{t+1})\right) dY_{t+1} \\ &= 0 \end{aligned}$$

Hence, $J_3(1) > J_2(1)$, which concludes the proof.

References

[1] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Honolulu, HI, USA, Apr. 2007, pp. 1357–1360.

[2] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wideband spectrum sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 2337–2340.

[3] Y. Wang, Z. Tian, and C. Feng, "A two-step compressed spectrum sensing scheme for wideband cognitive radios," in *Proc. IEEE Global Commun. Conf.*, Miami, FL, USA, Dec. 2010, pp. 1–5.

[4] G. Vazquez-Vilar, R. R. López-Valcarce, C. Mosquera, and N. González-Prelcic, "Wideband spectral estimation from compressed measurements exploiting spectral a priori information in cognitive radio systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Dallas, TX, USA, Apr. 2010, pp. 2958–2961.

[5] S. Hong, "Multi-resolution bayesian compressive sensing for cognitive radio primary user detection," in *Proc. IEEE Global Commun. Conf.*, Miami, FL, USA, Dec. 2010, pp. 1–6.

[6] Z. Yu, S. Hoyos, and B. M. Sadler, "Mixed-signal parallel compressed sensing and reception for cognitive radio," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Las Vegas, NV, USA, Apr. 2008, pp. 3861–3864.

[7] Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," in *Proc. IEEE Global Commun. Conf.*, New Orleans, LA, USA, Dec. 2008, pp. 1–5.

[8] L. Lai, Y. Fan, and H. V. Poor, "Quickest detection in cognitive radio: A sequential change detection framework," in *Proc. IEEE Global Commun. Conf.*, New Orleans, LA, USA, Dec. 2008, pp. 1–5.

[9] L. Lai, H. V. Poor, Y. Xin, and G. Georgiadis, "Quickest search over multiple sequences," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5375–5386, Aug. 2011.

[10] A. Tajer, R. Castro, and X. Wang, "Adaptive sensing of congested spectrum bands," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 6110–6125, Sep. 2012.

[11] A. Tajer, R. Castro, and X. Wang, "Adaptive spectrum sensing for agile cognitive radios," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Dallas, TX, USA, Mar. 2010, pp. 2966–2969.

[12] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[13] Z. Quan, S. Cui, A. Sayed, and H. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.

[14] C.-H. Hwang, G.-L. Lai, and S.-C. Chen, "Spectrum sensing in wideband OFDM cognitive radios," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 709–719, Feb. 2010.

[15] H. Li and R. Qiu, "A graphical framework for spectrum modeling and decision making in cognitive radio networks," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2010, pp. 1–6.

[16] K. Hossain, B. Champagne, and A. Assra, "Cooperative multiband joint detection with correlated spectral occupancy in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2682–2687, May 2012.

[17] S. Monfared, A. Taherpour, and T. Khattab, "Compressed wideband spectrum sensing with correlated subband occupancy in multi-antenna cognitive radios," in *Proc. IEEE 23rd Int. Symp. Pers. Indoor Mobile Radio Commun.*, Sep. 2012, pp. 2195–2201.

[18] L. Luo and S. Roy, "Modelling and analysis of detection time trade-offs for channel searching in cognitive radio networks," *IET Commun.*, vol. 6, no. 8, pp. 819–827, May 2012.

[19] A. N. Shiryaev, *Optimal Stopping Rules*. New York, NY, USA: Springer-Verlag, vol. 8, Nov. 2007.

[20] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge, U.K.: Cambridge Univ. Press, Dec. 2009.

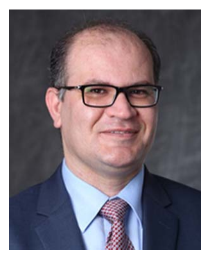
[21] S. Balaneshin, A. Tajer, and H. Poor, "Finite-horizon quickest search in correlated high-dimensional data," in *Proc. IEEE 6th Int. Symp. Commun., Control Signal Process.*, May 2014, pp. 222–225.

[22] M. Wellens, A. de Baynast, and P. Mahonen, "Exploiting historical spectrum occupancy information for adaptive spectrum sensing," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2008, pp. 717–722.

[23] M. Wellens, A. de Baynast, and P. Mahonen, "Performance of dynamic spectrum access based on spectrum occupancy statistics," *IET Commun.*, vol. 2, no. 6, pp. 772–782, Jul. 2008.

[24] A. Gorcin, H. Celebi, K. Qaraqe, and H. Arslan, "An autoregressive approach for spectrum occupancy modeling and prediction based on synchronous measurements," in *Proc. IEEE 22nd Int. Symp. Pers. Indoor Mobile Radio Commun.*, Sep. 2011, pp. 705–709.

[25] K. Patil, K. Skouby, R. Prasad, and S. Barge, "Spectrum occupancy information in support of adaptive spectrum sensing for cognitive radio," *Netw. Protocols Algorithms*, vol. 6, no. 1, pp. 76–86, Apr. 2014.



Ali Tajer (S'05–M'10–SM'15) received the B.Sc. and M.Sc. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2002 and 2004, respectively. He received the M.A. degree in statistics and the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA. He was with Princeton University, Princeton, NJ, USA during 2010–2012 as a postdoctoral research associate. He is currently an Assistant Professor of electrical, computer, and systems engineering at Rensselaer Polytechnic Institute. His research interests include estimation and detection theory, network information theory, wireless communications, and smart grids. Dr. Tajer is serving as Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON SMART GRID.

Javad Heydari received the B.Sc. and M.Sc. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively. He is currently pursuing his Ph.D. degree in the Department of Electrical, Computer, and System Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. During 2010–2012 he was with the Advanced Communication Research Institute at Sharif University of Technology where he worked on the applications of signal processing in communication systems. His research interests lie in the general areas of statistical signal processing, wireless communications, and statistical information theory.