Quickest Search and Learning Over Correlated Sequences: Theory and Application

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

***Abstract*—Consider a set of information sources, each generating a sequence of independent and identically distributed random variables over time. Each information source generates its data according to one of the two possible distributions F_0 or F_1. Due to potential physical couplings that govern the information sources, the underlying distribution of each sequence depends on those of the rest. Hence, the underlying distributions form a dependence kernel. Due to uncertainties in the physical models, however, the dependence kernel is not fully known. The objective is to design a sequential decision-making procedure that identifies a sequence generated according to F_1 with the fewest number of measurements. Earlier analyses of such search problems have demonstrated that the optimal design of the sequential rules strongly hinges on knowing the dependence kernel precisely. Motivated by the premise that the dependence kernel is not known, this paper designs a sequential inference mechanism that forms two intertwined inferential decisions for identifying a sequence of interest and learning the parameters of the dependence kernel. This paper devises three strategies that place different levels of emphasis on each of these inference goals. Optimal decision rules are characterized, and their performance is evaluated analytically. Also, the application of the proposed framework to wideband spectrum sensing is discussed. Finally, numerical evaluations are provided to compare the performance of the framework to those of the relevant existing literature.**

***Index Terms*—Model uncertainty, quickest detection, quickest search, sequential estimation, sequential sampling.**

I. Introduction

Searching over a set of data streams for the purpose of identifying data streams that exhibit desired statistical features arises in a wide range of applications. Categorically, identifying the streams of interest can represent identifying opportunities (arising, e.g., in wideband spectrum sensing) or mitigating risks (arising, e.g., in network intrusion or fraud detection). Identifying such data streams of interest, especially in large sets of data streams, is often very time-sensitive due to the transient nature of the opportunities that are attractive only when detected quickly, or due to the substantial costs that risks can incur if not responded to rapidly (c.f. [1] for more extended overview). The significance of searching over collections of data streams is expected to grow well into the future due to the advances in various technological, social, and economic domains, in which large-scale and complex data is routinely generated and processed for various inferential purposes.

Quickest search over a set of information streams aims to identify the desired streams in real-time and in the *quickest* fashion. Quickest search strikes a balance between the quality and the agility of the search process as two opposing figures of merit. Specifically, forming more reliable decisions necessitates collecting more data, which in turn delays the decision process, and subsequently penalizes the agility. The problem of quickest search over multiple data streams was first formalized and analyzed in [2] as an extension of sequential binary hypothesis testing [3] and [4]. The study in [2] considers a set of data streams, where each stream is generated according to one of the two known distributions F_0 and F_1 *independently* of the rest. The ultimate goal of quickest search in this context is identifying one sequence generated according to the desired distribution F_1 with the *fewest* number of measurements.

In network settings, however, in which there exists physical coupling among the mechanisms that govern the information sources and generate the data streams, the assumption that different data streams are generated according to one of the possible distributions independently of each other is not necessarily valid. Motivated by such network settings, [5] treats the quickest search over data streams that are generated according to a *known* linear dependency kernel, where it is shown that the structure of the optimal decision rules for performing quickest search strongly depends on the parameters of the dependency kernel. In reality, however, the dependency kernel is not necessarily known, e.g., when F_1 captures anomalous behaviors that have unpredictable patterns or causes.

In this paper we formalize a framework for quickest search problem over correlated data streams when the parameters of the dependency kernel are unknown. The framework is constructed in three stages, where each stage has its specific set of technical challenges. First we consider a *purely sequential detection* problem in which the objective is to identify one sequence generated according to F_1 with minimal delay. In this stage we are not concerned about learning the model, and all the uncertainties are dispensed with as nuance parameters. In the second stage, we consider a *purely sequential estimation* problem, in which the objective is to form reliable estimates for the parameters of

Manuscript received June 7, 2018; revised October 12, 2018; accepted November 4, 2018. Date of publication November 28, 2018; date of current version December 18, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hongbin Li. This research was supported in part by the U.S. National Science Foundation under Grant the CAREER Award ECCS-1554482 and Grant DMS-1737976. This paper was presented in part at the International Symposium on Information Theory, Aachen, Germany, Jun. 2017. *(Corresponding author: Ali Tajer.)*

J. Heydari was with the ECSE Department, Rensselaer Polytechnic Institute, Troy, NY 12180 USA. He is now with the LG Electronics, Santa Clara, CA 95054 USA (e-mail: javad.heydari@lge.com).

A. Tajer is with the ECSE Department, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: tajer@ecse.rpi.edu).

Digital Object Identifier 10.1109/TSP.2018.2883014

the dependency kernel with the fewest number of measurements made across the sequences. This stage does not identify any data stream and its focus is placed entirely on learning the model. Finally, we consider a sequential setting that pursues both inference goals. Specifically, a sequence generated according to F_1 is identified, and also reliable estimates about the parameters are produced. In analyzing all these three problems, the sequential framework and the associated sampling strategy are specified by two data-driven components. The first is the *stochastic* stopping time of the process, which is the time that sampling is terminated and reliable decisions are produced. The second component is the dynamic selection of the data streams over time for collecting data. Clearly, not all data streams are equally informative about the inference objectives, and optimally selecting and sampling the data streams over time have a pivotal role for ensuring the optimality of the coupled data-acquisition and decision-making processes.

Forming such dynamic decisions that pertain to data collection is closely related to the notion of *controlled sensing*, originally developed by Chernoff for binary composite hypothesis testing in order to dynamically decide about running one of the two possible experiments at each time [6]. Chernoff shows that performing the experiment with the best *immediate* return according to proper information measures achieves optimal performance in the asymptote of diminishing rate of erroneous decisions. Extensions of the Chernoff's rule to accommodate infinite number of available experiments and infinite number of hypotheses are studied in [7] and [8], respectively. Recent advances on controlled sensing for hypothesis testing include [9]–[12]. The existing relevant studies on sequential estimation include the pioneering work on sequentially estimating one parameter by observing one data stream of independent and identically distributed (i.i.d.) random variables in [13]. Such a sequential estimation routine is further extended in [14] and [15] to a setting in which multiple data streams are available. In such settings, a fully sequential strategy, besides the stopping rule and the final estimators, includes a selection rule that dynamically identifies the most informative data streams over time and gathers measurements from those data streams. In [14] multiple unknown parameters are available and each data stream depends only on one of the unknown parameters, while [15] generalizes the results to the setting in which data streams have common unknown parameters.

Other existing studies on quickest search that are relevant to the scope of this paper include scanning problems studied in [16]–[22]. The studies in [16]–[19] consider a finite set of data streams, with *exactly one* stream generated according to the desired distribution, and the objective is identifying the desired sequence. On the other hand, the studies in [20]–[22] consider a finite set of data streams that contains *multiple* data streams with the desired distribution, and the objective is identifying *all* of the data streams of interest. The search objective of detecting only *a fraction* of the sequences with the desired feature is studied in [23]–[25]. Fractional recovery of such sequences allows for missing some of them during the search process which ultimately leads to a faster process. Specifically, [23] characterizes some of the existing search procedures and their sampling complexity under different assumptions with the goal of finding one desired data stream among a large number of data streams, and [24] and [25] propose a data-adaptive search under a hard constraint on the sampling resources to detect a fraction of the desired data streams.

As a direct application domain of the theory developed, we consider the problem of identifying spectrum opportunities (holes) in the wideband spectrum. In such settings, the spectrum band is divided into multiple narrow-band channels, each of which being used by certain users for communication. Hence, at a given instant, the active users communicate over some of these channels, which we refer to as busy channels, and underutilize the rest, which we call spectrum holes. A user who seeks to initiate a communication session, scans the spectrum via sequentially tuning its receiver's filters to different sub-channels and collecting data. It examines the channels sequentially until it identifies a channel as a spectrum hole with sufficient confidence. The sequence of measurements collected from a busy channel are modeled by F_0, and F_1 models the measurements from a spectrum hole. The level of utilization of a spectrum band in an area varies over time, and the historical data are leveraged to only obtain certain stochastic occupancy patterns for each narrow-band channel. Since the occupancy statuses of the spectrum holes vary rapidly and the spectrum holes might not remain free for a long duration, it is of paramount importance to devise mechanisms 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. Also, we may seek to identify the level of under-utilization of the spectrum and the occupancy pattern by estimating the dependency kernel of the busy channels. Therefore, the spectrum sensing problem can be considered as an instance of the problem studied in this paper.

The remainder of the paper is organized as follows. Section II provides the data and sampling model, and formalizes the quickest search and learning problem of interest. The optimal inference rules are characterized in Section III, and the sampling procedures and the associated stopping rules for purely search and purely estimation routines are characterized in Section IV and Section V, respectively. The combined inference procedure is provided and analyzed in Section VI. Section VII discusses the application in spectrum sensing and provides the simulation results, and Section VIII concludes the paper.

II. Model and Formulation

A. Data Model

Consider an *ordered* set of n sequences $\{\mathcal{X}^i : i \in \{1, \dots, n\}\}$, where each sequence consists of i.i.d. real-valued observations $\mathcal{X}^i \triangleq \{X^i_j : j \in \mathbb{N}\}$. Each sequence is generated according to one of the two possible distributions, hence, obeying the following dichotomous model:

$$\begin{array}{lll} \mathsf{H}_0: & X^i_j \sim F_0 & \\ \mathsf{H}_1: & X^i_j \sim F_1 & \end{array} \quad \text{for } j \in \mathbb{N}\,, \tag{1}$$

where F_0 and F_1 denote cumulative distribution functions (cdfs) and are assumed to be distinct and known. Distribution F_0 captures the statistical behavior of the normal sequences, and the distribution of the abnormal sequences is F_1. We further assume that well-defined probability density functions (pdfs) corresponding to F_0 and F_1 exist, which are denoted by f_0 and f_1, respectively.

Each sequence is generated by F_0 or F_1 according to a dependency kernel among the *ordered* set $\{\mathcal{X}^i : i \in \{1,\ldots,n\}\}$. Specifically, we assume that sequence $\mathcal{X}^1$ is generated according to F_1 with prior probability ϵ_0 and for the subsequent sequences, the prior probability of each sequence being generated by F_1 is controlled by the distribution of its preceding sequence. More specifically, by denoting the *true* model underlying sequence $\mathcal{X}^i$ by T_i, we have

$$\mathbb{P}(\mathsf{T}_1 = \mathsf{H}_1) \;=\; \epsilon_0 , \tag{2}$$

and for $i \in \{2,\ldots,n\}$ we have

$$\mathbb{P}(\mathsf{T}_i = \mathsf{H}_1 \mid \mathsf{T}_{i-1} = \mathsf{H}_j) = \epsilon_j , \quad \text{for } j \in \{0,1\}. \tag{3}$$

In general, $\epsilon_0 \neq \epsilon_1$, and the setting $\epsilon_0 = \epsilon_1 = \epsilon$ subsumes the quickest search over multiple independent sequences [2]. Furthermore, we assume that $\epsilon_0, \epsilon_1 \in [0,1]$ are *unknown* random variables with continuous pdfs denoted by g_0 and g_1, respectively. Accordingly, we define $\boldsymbol{\epsilon} \triangleq [\epsilon_0 \;\; \epsilon_1]$ as the vector of unknown parameters that specify the dependency kernel parameters, which we need to estimate.

B. Sampling Model

The objective of the search process is to identify *one* abnormal sequence with the *fewest* number of measurements that can be collected from the entire set of sequences. To minimize the number of measurements, the sampling procedure sequentially examines the sequences by taking one measurement at-a-time until a sequence generated according to F_1 can be identified with sufficient reliability. We denote the index of the observed sequence at time t by $s_t \in \{1,\ldots,n\}$, and define Y_t as the measurement taken at time t. Hence, we can abstract the information accumulated sequentially by the filtration $\{\mathcal{F}_t \;:\; t \in \mathbb{N}\}$, where

$$\mathcal{F}_t \triangleq \sigma(Y_1, s_1, ,\ldots, Y_t, s_t) \; . \tag{4}$$

Without loss of generality, we assume that the sampling process starts from the first sequence, i.e., $s_1 = 1$, and at time t and based on the information accumulated up to time t, i.e., $\mathcal{F}_t$, it takes one of the following three possible actions:

A_1) *Decision:* stops taking more samples and declares one of the sequences observed up to time t as an abnormal one. We remark that as shown in [5], when ϵ_0 and ϵ_1 are known precisely, the optimal decision rules strongly depend on the actual values of ϵ_0 and ϵ_1. Hence, the quickest search objective is strongly coupled with also concurrently forming reliable estimates for ϵ_0 and ϵ_1. Hence, the decision also involves forming reliable estimates for ϵ_0 and ϵ_1.

A_2) *Observation:* due to lack of sufficient confidence to make a decision or form reliable estimates for ϵ_0 and ϵ_1, one more sample is taken from the same sequence, i.e., $s_{t+1} = s_t$.

A_3) *Exploration:* sequence s_t is discarded and the sampling procedure switches to the next sequence and takes one sample from the new sequence, 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 instant at which action A_1 (decision) is performed. We denote the estimate formed for $\boldsymbol{\epsilon}$ by $\hat{\boldsymbol{\epsilon}}(\tau) \triangleq [\hat{\epsilon}_0(\tau) \;\; \hat{\epsilon}_1(\tau)]$. For the detection rule, due to the dynamics of the system, we assume that the decision-maker can only declare one of the sequences observed during the last T time instants as an abnormal sequence. We denote the set of sequences observed during the interval $[\tau - T + 1, \ldots, \tau]$ by $\mathcal{S}(\tau, T)$, i.e.,

$$\mathcal{S}(\tau, T) \triangleq \{s_{\tau - T+1}, \ldots, s_\tau\} \; . \tag{5}$$

There is full flexibility in choosing T, and it can range from 1 to τ. We denote the detection rule at the stopping time by $\delta(\tau) \in \mathcal{S}(\tau, T)$. To characterize dynamic switching between observation and exploration actions we define the binary function

$$\psi : \{1, \ldots, \tau - 1\} \rightarrow \{0, 1\} \; , \tag{6}$$

such that at time $t \in \{1,\ldots,\tau-1\}$ if the decision is in favor of performing observation (A_2) we set $\psi(t) = 0$, while $\psi(t) = 1$ indicates a decision in favor of exploration (A_3). Hence, $\forall t \in \{1,\ldots,\tau-1\}$:

$$\psi(t) = \begin{cases} 0 & \text{action } \mathrm{A}_2 \\ 1 & \text{action } \mathrm{A}_3 \end{cases} . \tag{7}$$

A sequential decision is completely characterized by the combination

$$\Phi \triangleq (\tau, \hat{\boldsymbol{\epsilon}}(\tau), \delta(\tau), \psi(1), \ldots, \psi(\tau-1)) \; . \tag{8}$$

For a given $\boldsymbol{\epsilon}$, we denote the nominal probability measure governing the collected samples and the expectation with respect to this measure by $\mathbb{P}_{\boldsymbol{\epsilon}}$ and $\mathbb{E}_{\boldsymbol{\epsilon}}$, respectively. We assume that $\mathbb{P}_{\boldsymbol{\epsilon}}$ has a well-defined probability density function denoted by $p_{\boldsymbol{\epsilon}}$.

C. Problem Formulation

The optimal search procedure can be found by determining Φ. The natural performance measures for evaluating the efficiency of any sequential sampling strategy Φ include the quality of the final decisions and the agility of the process. The quality is captured by (i) the frequency of the erroneous decisions, and (ii) an estimation cost capturing the fidelity of the estimates. For this purpose, we define the terminal decision error probability as

$$\mathsf{P}_\Phi(\tau) \triangleq \mathbb{P}(\mathsf{T}_{\delta(\tau)} = \mathsf{H}_0 \mid \mathcal{F}_\tau) \; , \tag{9}$$

and adopt the mean-squared error as the estimation cost, and define the terminal estimation cost as

$$\mathsf{R}_\Phi(\tau) \triangleq \mathbb{E}\{\|\boldsymbol{\epsilon} - \hat{\boldsymbol{\epsilon}}(\tau)\|^2 \mid \mathcal{F}_\tau\} \; . \tag{10}$$

The agility is captured by the stochastic delay in reaching a decision, i.e., τ. By integrating these figures of merit into one cost function, the stochastic aggregate cost function for a given Φ at time t is given by

$$\mathscr{C}_\Phi(t) \triangleq c_{\rm d} \cdot \mathsf{P}_\Phi(t) + c_{\rm e} \cdot \mathsf{R}_\Phi(t) + c_{\rm s} \cdot t \ , \quad (11)$$

where $c_{\rm d}$, $c_{\rm e}$, and $c_{\rm s}$ are positive constants that balance the quality and the agility of the search process. It is noteworthy that the cost function defined in (11) is stochastic as it is a function of the random stopping time and the observations, which are random variables, through filtration $\mathcal{F}_t$. In the following sections, under different settings we characterize the stopping time, the switching rules, the final decision rules, and the associated performance guarantees.

D. Assumptions

Optimizing the cost function of (11) requires a set of assumptions to satisfy some regularity conditions. To this end, we define

$$\lambda(\boldsymbol{\epsilon}; Y^t) \triangleq \log p_{\boldsymbol{\epsilon}}(Y_1, \dots, Y_t) \ , \quad (12)$$

as the log-likelihood of $\boldsymbol{\epsilon}$ given the measurements taken up to time t. Furthermore, we denote the maximum likelihood estimator (MLE) of $\boldsymbol{\epsilon}$ based on the measurements up to time t by

$$\hat{\boldsymbol{\epsilon}}_{\rm ML}(t) \triangleq \underset{\boldsymbol{\epsilon}\in[0,1]^2}{\arg\max} \ \lambda(\boldsymbol{\epsilon}; Y^t) \ . \quad (13)$$

Throughout the analysis we make the following assumptions, the motivations for which are explained in the sequel.

1) For $j \in \{0, 1\}$,

$$\mathbb{E}_{\boldsymbol{\epsilon}} \left\{ \frac{\partial \lambda(\boldsymbol{\epsilon}; Y^t)}{\partial \epsilon_j} \right\} = 0 \ . \quad (14)$$

2) The normalized log-likelihood function is finite, i.e.,

$$\frac{1}{t}\mathbb{E}_{\boldsymbol{\epsilon}}\{|\lambda(\boldsymbol{\epsilon}; Y^t)|\} < \infty, \ \forall \boldsymbol{\epsilon} \ . \quad (15)$$

3) The log-likelihood function is continuous in $\boldsymbol{\epsilon}$.
4) The Fisher information matrix for $\boldsymbol{\epsilon}$ is positive-definite, continuous in $\boldsymbol{\epsilon}$, and all its entries are finite.
5) For $j \in \{0, 1\}$, the prior density $g_j(\epsilon_j)$ is continuous and strictly positive over $[0, 1]$.
6) For any real $\Delta > 0$

$$\mathbb{E}_{\boldsymbol{\epsilon}}\{\sup\{\lambda(\boldsymbol{\epsilon}; Y^t) - \lambda(\boldsymbol{\epsilon}'; Y^t) \ : \ \|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}'\| \geq \Delta\}\} < 0 \ . \quad (16)$$

Assumption 1 is the regularity conditions for the Cramer-Rao lower bound [26], assumptions 2–4 ensure that a strongly consistent MLE exists, assumption 5 guarantees the consistency of the Bayesian estimator, and assumption 6 ensures that the process does not degenerate, i.e., for sufficiently distinct values of $\boldsymbol{\epsilon}$ the likelihoods $\lambda(\boldsymbol{\epsilon}; Y^t)$ are distinguishable.

III. Optimal Inference Rules

In this section, we characterize the detection rule $\delta(\tau)$ and the estimation rule $\hat{\boldsymbol{\epsilon}}(\tau)$ for any given stopping time τ and any switching sequence $\{\psi(t) : t \in \{1, \dots, \tau - 1\}\}$. Then, in the following sections we focus on two special cases of the search problem. One case places the emphasis on the detection subproblem by setting $c_{\rm e} = 0$ and nulling the contribution of the estimation cost to the aggregate cost defined in (11), and the other case focuses on the estimation subproblem by setting $c_{\rm d} = 0$. Based on the insights from these special cases, we finally treat the problem in its general form.

It is shown in [27] that in a sequential setting, for any given sampling strategy and stopping rule, there exists a fixed set of final decisions that are optimal. Furthermore, the quadratic estimation cost $\mathsf{R}_\Phi(t)$ is independent of the detection rule, and also the detection cost is independent of the estimate of $\boldsymbol{\epsilon}$. Hence, for any given stopping time τ and sequence of switching functions $\{\psi(t) : t \in \{1, \dots, \tau - 1\}\}$, detection and estimation decision rules can be decoupled. The following theorem formalizes these results.

Theorem 1 [27]: The detection and estimation rules at the stopping time for optimizing the cost function (11) are independent of the stopping time and the switching rule.

Proof: The proof, the details of which can be found in [27], involves showing that optimizing the cost function in (11) can be carried out by first, inserting the optimal decision rules (detection and estimation decisions), then computing the total cost, and finally decising whether the sampling process should terminate or further measurements are required to make better decisions. ■

This theorem facilitates characterizing the detection and estimation decisions. To proceed, we define π_t^i as the posterior probability that the sequence $\mathcal{X}^i$ is abnormal given the information up to time t, i.e.,

$$\pi_t^i \triangleq \mathbb{P}(\mathsf{T}_i = \mathsf{H}_1 \,|\, \mathcal{F}_t) \ . \quad (17)$$

By defining

$$\kappa_t^i(\boldsymbol{\epsilon}) \triangleq \mathbb{P}(\mathsf{T}_i = \mathsf{H}_1 \,|\, \mathcal{F}_t, \boldsymbol{\epsilon}) \ , \quad (18)$$

it can be readily verified that

$$\kappa_{t+1}^i(\boldsymbol{\epsilon}) = \begin{cases} \Theta(\kappa_t^i(\boldsymbol{\epsilon}), Y_{t+1}) & \text{if } s_t = i \\ \kappa_t^i(\boldsymbol{\epsilon}) & \text{if } s_t \neq i \end{cases} , \quad (19)$$

where we have defined

$$\Theta(\kappa(\boldsymbol{\epsilon}), y) \triangleq \frac{\kappa f_1(y)}{\kappa f_1(y) + (1-\kappa) f_0(y)} \cdot \mathbb{1}_{\{\psi(t)=0\}} + \frac{\bar{\kappa} f_1(y)}{\bar{\kappa} f_1(y) + (1-\bar{\kappa}) f_0(y)} \cdot \mathbb{1}_{\{\psi(t)=1\}}, \quad (20)$$

where $\bar{\kappa} \triangleq (\epsilon_1 - \epsilon_0)\kappa + \epsilon_0$. Based on this, we have

$$\pi_t^i = \int \kappa_t^i(\boldsymbol{\epsilon}) g(\boldsymbol{\epsilon} \,|\, \mathcal{F}_t) \, d\epsilon_0 d\epsilon_1, \quad (21)$$

where $g(\boldsymbol{\epsilon} \,|\, \mathcal{F}_t)$ denotes the posterior pdf of $\boldsymbol{\epsilon}$. Based on these definitions, the next theorem provides the optimal decision rules for any given stopping time and switching sequence.

Theorem 2 (Decision Rules): For a given stopping time τ and switching sequence:

1) The optimal detection rule, which minimizes $\mathscr{C}_\Phi(\tau)$, is

$$\delta(\tau) = \underset{i\in\mathcal{S}(\tau,T)}{\arg\max} \ \pi_\tau^i \ . \quad (22)$$

2) The optimal estimation rule, which minimizes $\mathscr{C}_\Phi(\tau)$, is

$$\hat{\boldsymbol{\epsilon}}(\tau) = \mathbb{E}\{\boldsymbol{\epsilon} \,|\, \mathcal{F}_\tau\} \; . \tag{23}$$

Proof: The proof results from Theorem 1 and leveraging the fact that the detection and estimation costs are independent. The rest follows the standard proof for showing that, in Bayesian settings, maximum a posterior probability (MAP) is the optimal detection rule and posterior mean is the optimal estimation rule. For the detection rule, a randomized detection rule is considered, the detection cost of which is computed and is shown to be minimized by a pure strategy that detects the sequence with the maximum posterior probability. The optimal estimator is obtained by expanding the estimation cost function and taking derivative with respect to the estimator. ■

Remark 1: It is noteworthy that, since the previously discarded sequences will not be revisited in the future and the final detection rule only selects one sequence among the ones observed during the last T time instants, at each time t and after observing Y_t, we only need to update the posterior probabilities π_t^i for sequences $i \in \mathcal{S}(t,T)$. For larger values of T, we need to update the posterior probabilities of more sequences. Specifically, in two extreme cases, if $T = t$ we should update the posterior probabilities of all the observed sequences up to time t and select one of them at the stopping time, while if $T = 1$ only the sequence under observation is considered as a candidate for being generated according to F_1.

Remark 2: The other quantity required for decision-making at each time is the posterior distribution of $\boldsymbol{\epsilon}$, i.e., $g(\boldsymbol{\epsilon} \,|\, \mathcal{F}_t)$. To this end, the decision maker is required to remember only some statistics of the collected data for each sequence, no all of the measurements. Specifically, at each time instant t and for any sequence $s \leq \psi(t)$ and $j \in \{0,1\}$ if we denote the likelihood that the sequence s is drawn from F_j, given all the measurements from sequence s, by $L_j^t(s)$, i.e.,

$$L_j^t(s) = \prod_{\{t:\psi(t)=s\}} f_j(Y_t) \; ,$$

the decision-maker is only required to remember the set $\{L_j^t(s) \;:\; j \in \{0,1\}, \;\; s \leq \psi(t)\}$. Hence, for the exact computation of the posterior distribution, the decision-maker is only required to remember twice the number of visited sequences. Then, when only one sequence has been observed, we have

$$\begin{aligned} g(\boldsymbol{\epsilon} \,|\, \mathcal{F}_t) &= \frac{[L_1^t(1) + (1-\epsilon_0)L_0^t(1)]g_0(\epsilon_0)g_1(\epsilon_1)}{\int [\epsilon_0 L_1^t(1) + (1-\epsilon_0)L_0^t(1)]g_0(\epsilon_0)g_1(\epsilon_1)d\epsilon_0 d\epsilon_1} \\ &= \left[\frac{[\epsilon_0 L_1^t(1) + (1-\epsilon_0)L_0^t(1)]g_0(\epsilon_0)}{\int [\epsilon_0 L_0^t(1) + (1-\epsilon_0)L_1^t(1)]g_0(\epsilon_0)d\epsilon_0}\right] g_1(\epsilon_1) \; . \end{aligned}$$

It is observed that observing the first sequence does not change the marginal posterior distribution of ϵ_1 and the best estimate for ϵ_1 remains the prior mean. However, the marginal posterior distribution of ϵ_0 changes and the posterior mean estimator for ϵ_0 becomes

$$\mathbb{E}\{\epsilon_0 | Y_1\} = \frac{\mathbb{E}\{\epsilon_0^2\}[L_1^t(1) - L_0^t(1)] + \mathbb{E}\{\epsilon_0\}L_0^t(1)}{\mathbb{E}\{\epsilon_0\}[L_1^t(1) - L_0^t(1)] + L_0^t(1)},$$

where all the expectations on the right hand side are with respect to the prior distribution. Now, for any sequence $k \leq \psi(t)$ by defining $\mathcal{Y}_k^t$ as the set of samples taken from sequence k and $\mathcal{Y}_{-k}^t$ as the set of samples taken from sequences other than k we obtain

$$g(\boldsymbol{\epsilon} \,|\, \mathcal{F}_t) = \frac{\mathbb{P}(\epsilon_0, \epsilon_1 | \mathcal{Y}_{-k}^t)\mathbb{P}(\mathcal{Y}_k^t | \boldsymbol{\epsilon}, \mathcal{Y}_{-k}^t)}{\int \mathbb{P}(\boldsymbol{\epsilon} | \mathcal{Y}_{-k}^t)\mathbb{P}(\mathcal{Y}_k^t | \boldsymbol{\epsilon}, \mathcal{Y}_{-k}^t) d\epsilon_0 d\epsilon_1} \; .$$

The first term in the nominator is computed by using the samples from the first $k-1$ sequences, while the second term is obtained from

$$\begin{aligned} &\mathbb{P}(\mathcal{Y}_k^t | \boldsymbol{\epsilon}, \mathcal{Y}_{-k}^t) \\ &\quad = \frac{L_0^t(k) + [L_1^t(k) - L_0^t(k)]p(\boldsymbol{\epsilon}, k, t)}{\int [L_0^t(k) + [L_1^t(k) - L_0^t(k)]p(\boldsymbol{\epsilon}, k, t)] d\epsilon_0 d\epsilon_1} \; , \end{aligned}$$

where we have defined

$$p(\boldsymbol{\epsilon}, k, t) \stackrel{\triangle}{=} \epsilon_0 + (\epsilon_1 - \epsilon_0)\mathbb{P}(\mathsf{T}_{k-1} = \mathsf{H}_1 | \mathcal{Y}_{-k}^t, \boldsymbol{\epsilon}) \; ,$$

and $\mathbb{P}(\mathsf{T}_{k-1} = \mathsf{H}_1 | \mathcal{Y}_{-k}^t, \boldsymbol{\epsilon})$ also follows a recursive equation as follows:

$$\begin{aligned} &\mathbb{P}(\mathsf{T}_{k-1} = \mathsf{H}_1 | \mathcal{Y}_{-k}^t, \boldsymbol{\epsilon}) \\ &= \frac{L_1^t(k-1)p(\boldsymbol{\epsilon}, k-1, t)}{L_1^t(k-1)p(\boldsymbol{\epsilon}, k-1, t) + L_0^t(k-1)(1 - p(\boldsymbol{\epsilon}, k-1, t))} \; . \end{aligned}$$

In practice, since we are calculating the integrals approximately, instead of all the likelihood values, we can save the value of $\mathbb{P}(\mathsf{T}_k = \mathsf{H}_1 | \mathcal{Y}_k^t, \boldsymbol{\epsilon})$ for all the combinations of ϵ_0 and ϵ_1. Since these two quantities are bounded between 0 and 1, the memory required for this purpose depends on the resolution with which the decision-maker computes the integrals. One disadvantage of this approximation is that the approximation errors propagate over time and may lead to large errors when the number of visited sequences increases.

Given the decision rules in Theorem 2, we next characterize the optimal stopping time τ, and the associated optimal switching rules $\{\psi(t) : t \in \{1, \ldots, \tau - 1\}\}$. For this purpose, we first consider the sequential detection and sequential estimation problems separately in two different settings. By leveraging the insight gained from these special cases we solve the joint detection and estimation problem in Section VI. In the sequel, we consider the problem of optimizing the cost formulated in (11) in the asymptote of large sample-size setting, or equivalently $c_s \rightarrow 0$. In this setting, since the samples, while being conditionally independent, are not identically distributed we focus on a weak version of asymptotic optimality [15].

IV. Emphasis on Detection ($c_{\rm e} = 0$)

We first consider a purely sequential search setting, in which the estimation quality is unintegrated by setting $c_{\rm e} = 0$. This problem under the assumption that $\boldsymbol{\epsilon}$ is known and the objective is to minimize the delay is studied in [5], where the analyses provide the optimal stopping time and switching rules. In this section, we provide stopping and switching rules that admit certain optimality guarantees and facilitate generalizing the results to the general case of $c_{\rm e} \neq 0$. To this end, we propose a stopping time and switching rules, the combination of which accepts asymptotic optimality guarantees. Specifically, we define

$$\tau_{\rm d}^* \triangleq \inf \left\{ t \ : \ \max_{i\in\mathcal{S}(t,T)} \pi_t^i \geq 1 - c_{\rm sd} \right\} , \tag{24}$$

where $c_{\rm sd} \triangleq \frac{c_{\rm s}}{c_{\rm d}}$. According to this stopping rule, the sampling process continues until one is confident enough that one of the sequences observed in the window of past T time instants is generated by F_1.

Next, we characterize the switching rule prior to the stopping time in order to dynamically decide between the *exploration* and *observation* actions. To proceed, at any time $t \in \{1,\dots,\tau_{\rm d}^* - 1\}$ we form the maximum a posterior (MAP) estimate for $\boldsymbol{\epsilon}$, denoted by

$$\tilde{\boldsymbol{\epsilon}}(t) \triangleq \arg\max_{\boldsymbol{\epsilon}\in[0,1]^2} \ g(\boldsymbol{\epsilon} \mid \mathcal{F}_t) \ . \tag{25}$$

Then, at any time $t \in \{1,\dots,\tau_{\rm d}^* - 1\}$ we set the switching rule to discard sequence s_t and switch to sequence (s_t+1) when $\pi_t^{s_t}$ falls below a data-driven threshold. Specifically,

$$\psi_{\rm d}(t) \triangleq \begin{cases} 1 & \text{if } \pi_t^{s_t} < [\tilde{\epsilon}_1(t) - \tilde{\epsilon}_0(t)]\pi_t^{s_t} + \tilde{\epsilon}_0(t) \\ 0 & \text{if } \pi_t^{s_t} \geq [\tilde{\epsilon}_1(t) - \tilde{\epsilon}_0(t)]\pi_t^{s_t} + \tilde{\epsilon}_0(t) \end{cases} . \tag{26}$$

This switching rule, when combined with the stopping time given in (24), achieves the weakly asymptotic *pointwise* optimality [15], formalized in the following theorem.

Theorem 3: Consider a sequential strategy $\Phi_{\rm d}$ with the detection and estimation rules specified in (22) and (23), the stopping time defined in (24) and the switching rule given in (26). For any other sampling strategy $\hat{\Phi}$ with the stopping time $\hat{\tau}$, we have

$$\lim_{c_{\rm sd}\to 0} \mathbb{P}\left\{ \frac{\mathscr{C}_{\Phi_{\rm d}}(\tau_{\rm d}^*)}{\mathscr{C}_{\hat{\Phi}}(\hat{\tau})} \leq 1 + \Delta \right\} = 1 \ , \quad \forall \Delta > 0 \ . \tag{27}$$

Proof: The proof involves showing that

$$\frac{1}{\tau_{\rm d}^*} \log \mathsf{P}_\Phi(\tau_{\rm d}^*) \xrightarrow{c_{\rm sd}\to 0} -D_{\rm KL}(f_1 \| f_0) \ , \quad \text{a.s.}$$

This is carried out by showing that the number of collected samples is dominated by the samples taken from the sequence being declared as an abnormal sequence, and using the law of large number of convergence of log-likelihood ratio. Then, the assumptions of Theorem 2.1 (ii) in [13] is satisfied and the proof is concluded. See Appendix B for the details of the proof. ■

This theorem indicates that a weakly asymptotically pointwise optimal sequential strategy, defined by its stopping time and switching rule, can be reduced to comparing the posterior probability terms with two thresholds at each time t.

V. Emphasis on Estimation ($c_{\rm d} = 0$)

The problem of sequential estimation from one sequence is studied in [13], where an unknown parameter is estimated by sequentially observing a sequence of i.i.d. random variables until a reliable estimate can be formed. The results have been extended to a setting in which multiple data streams are available and only one of them can be observed at each time instant [14] and [15]. In such settings, a dynamic selection rule should be designed in order to identify the most informative sequence at each time to collect its data. In this section, the dynamic selection rule is the switching rule $\psi(t)$ and should decide between the observation and exploration actions.

To formalize the decision rules, we define the Fisher information matrices $\mathcal{I}^0(\boldsymbol{\epsilon}) \in \mathbb{R}^{2\times 2}$ and $\mathcal{I}^1(\boldsymbol{\epsilon}) \in \mathbb{R}^{2\times 2}$ such that

$$[\mathcal{I}^0(\boldsymbol{\epsilon})]_{ij} \triangleq -\mathbb{E}\left\{ \frac{\partial}{\partial \epsilon_{i-1}} \frac{\partial}{\partial \epsilon_{j-1}} \log \mathbb{P}(Y_{t+1} \mid \mathcal{F}_t, \boldsymbol{\epsilon}, \psi(t) = 0) \right\}, \tag{28}$$

is the (i,j) entry of $\mathcal{I}^0(\boldsymbol{\epsilon})$, and

$$[\mathcal{I}^1(\boldsymbol{\epsilon})]_{ij} \triangleq -\mathbb{E}\left\{ \frac{\partial}{\partial \epsilon_{i-1}} \frac{\partial}{\partial \epsilon_{j-1}} \log \mathbb{P}(Y_{t+1} \mid \mathcal{F}_t, \boldsymbol{\epsilon}, \psi(t) = 1) \right\}, \tag{29}$$

is the (i,j) entry of $\mathcal{I}^1(\boldsymbol{\epsilon})$. Based on these definitions, we characterize a stopping time and a switching rule that achieve asymptotic optimality when $c_{\rm se} \triangleq \frac{c_{\rm s}}{c_{\rm e}}$ tends to zero. Specifically, we define the stopping time as

$$\tau_{\rm e}^* \triangleq \inf \left\{ t \ : \ \mathsf{R}_\Phi(t) \leq (t+1)\cdot c_{\rm se} \right\} . \tag{30}$$

According to this stopping rule, when the estimation cost $\mathsf{R}_\Phi(t)$ falls below the total sampling cost $(t+1)\cdot c_{\rm se}$, the sampling process terminates. Furthermore, the switching rule should select the action between *observation* and *exploration* such that it minimizes the estimation variance. According to the Cramer-Rao bound, the estimation cost defined in (10) is lower bounded by the trace of the inverse of the Fisher information matrix associated with the data and the unknown parameters. In the sequel, we will show that this lower bound is achievable in the asymptote of large sample size. Hence, we select the action that minimizes this lower bound. Therefore, we first compute $\hat{\boldsymbol{\epsilon}}_{\rm ML}(t)$ as the maximum likelihood (ML) estimate of $\boldsymbol{\epsilon}$ at time t based on the observations up to time t, i.e.,

$$\hat{\boldsymbol{\epsilon}}_{\rm ML}(t) \triangleq \arg\max_{\boldsymbol{\epsilon}\in[0,1]^2} \ \mathbb{P}(Y_1,\dots,Y_t; s_1,\dots,s_t \mid \boldsymbol{\epsilon}) \ . \tag{31}$$

Then, the switching rule $\psi_{\rm e}(t)$ is set to be a randomized rule such that it is set to 0 or 1 randomly according to a Bernoulli random variable with parameter $p^*(\hat{\boldsymbol{\epsilon}}_{\rm ML}(t))$, which is the solution to

$$p^*(\boldsymbol{\epsilon}) \triangleq \arg\min_{p\in[0,1]} \ \sum_{i=1}^{2} \left[\left(p\cdot \mathcal{I}^1(\boldsymbol{\epsilon}) + (1-p)\cdot \mathcal{I}^0(\boldsymbol{\epsilon}) \right)^{-1} \right]_{ii} . \tag{32}$$

This switching rule ensures that the sampling process takes the action that minimizes the variance of estimation. While ignoring the impact of the current decision on the future ones, it can be shown that in the large sample regimes, it is asymptotically optimal.

Theorem 4: Let $\Phi_{\rm e}$ be the sampling strategy characterized by the stopping time and the switching rule given in (30) and (32), respectively. Then, when $c_{\rm se}$ approaches zero, for any other sampling strategy $\hat{\Phi}$ we have

$$\lim_{c_{\rm se}\to 0} \mathbb{P}\left\{ \frac{\mathscr{C}_{\Phi_{\rm e}}(\tau_{\rm e}^*)}{\mathscr{C}_{\hat{\Phi}}(\hat{\tau})} \leq 1 + \Delta \right\} = 1 \ , \quad \forall \Delta > 0 \ . \tag{33}$$

Proof: The main idea of the proof is to show that the posterior distribution of $\boldsymbol{\epsilon}$ is Gaussian. Since for Gaussian distributions the Bayesian estimator is efficient, the Fisher information matrix is exactly the amount of information obtained from a new sample. Therefore, it provides all the information needed for decision-making. For complete proof, see Appendix C. ■

This switching rule ensures that a sufficient number of samples are taken from the current sequence so that its distribution is distinguishable before switching to the next sequence, and the stopping rule guarantees that enough number of sequences are observed such that a reliable estimate of $\boldsymbol{\epsilon}$ can be formed.

VI. Combined Search and Estimation

With the insights gained from the previous two sections, in this section we treat the quickest search problem of interest in its general form, which involves forming reliable decisions for both estimation and detection routines. We first characterize a stopping time by noting that in the search problem the sampling process terminates when $(1 - \max_{i \in \mathcal{S}(t,T)} \pi_t^i)$ falls below $c_{\rm sd}$, i.e., the relative cost of one new sample, while in the estimation problem, it stops when the normalized estimation cost $\left(\frac{\mathsf{R}_\Phi(t)}{t+1}\right)$ is smaller than the relative cost of one new sample $c_{\rm se}$. Hence, for the general quickest search problem we define the stopping time as

$$\tau^* \stackrel{\triangle}{=} \inf \left\{ t \ : \ c_{\rm d} \left(1 - \max_{i \in \mathcal{S}(t,T)} \pi_t^i \right) + c_{\rm e} \frac{\mathsf{R}_\Phi(t)}{t+1} \leq c_{\rm s} \right\} . \tag{34}$$

While it is a combination of the stopping rules in the previous settings, we will show that it can be also obtained directly by optimizing the total Bayesian cost given in (11).

Theorem 5: Let $\bar{\Phi}$ be the sampling strategy with the stopping time given in (34) and the optimal switching sequence $\{\psi^*(t) \ : \ t \in \{1, \ldots, \tau^* - 1\}\}$, and let $\hat{\Phi}$ be any arbitrary sampling strategy with the same switching rule and any other stopping time $\hat{\tau}$. For all $\hat{\Phi}$ and $\hat{\tau}$ we have

$$\lim_{c_{\rm sd}, c_{\rm se} \rightarrow 0} \mathbb{P} \left\{ \frac{\mathscr{C}_{\bar{\Phi}}(\tau^*)}{\mathscr{C}_{\hat{\Phi}}(\hat{\tau})} \leq 1 + \Delta \right\} = 1 \,, \quad \forall \Delta > 0 \,. \tag{35}$$

Proof: We use the convergence properties of estimation and detection cost functions in combination with the stopping rule in (34) and show that the ratio between the cost of the stopping rule given in (34) and cost of any other stopping rule is less than 1. See Appendix D for more details. ■

Next, we characterize the switching rules for dynamically deciding between *exploration* and *observation* actions. First, we note that, taking a new sample from any sequence reduces the average estimation cost, while the detection error probability depends on the number of samples taken from the sequence declared as the abnormal sequence. Also, taking many samples from one sequence cannot improve the estimation cost significantly. Hence, for the switching rule, at the beginning we apply the rule given in the purely estimation setting (32). When the estimation cost becomes sufficiently small, we apply the switching rule given in the purely detection setting (26). Based on this, we set the switching rule as follows:

$$\psi(t) = \begin{cases} \psi_{\rm e}(t) & \text{if } \frac{\mathsf{R}_\Phi(t)}{t+1} > c_{\rm se} \\ \psi_{\rm d}(t) & \text{if } \frac{\mathsf{R}_\Phi(t)}{t+1} \leq c_{\rm se} \end{cases} . \tag{36}$$

This switching rule, at the beginning of the sampling process, is more focused on forming a reliable estimate for $\boldsymbol{\epsilon}$. When the estimation cost is sufficiently small, it gradually shifts the focus to forming a reliable detection decision. The following theorem formalizes the asymptotic optimality properties of the sampling strategy characterized by the stopping rule in (34) and the switching rule in (36).

Theorem 6: Let Φ^* be the sampling strategy with the stopping time given in (34) and a switching rule given in (36), and let $\hat{\Phi}$ be any arbitrary sampling strategy with the stopping time $\hat{\tau}$. Then, in the asymptote of $c_{\rm se} \rightarrow 0$, $c_{\rm sd} \rightarrow 0$, and when $c_{\rm se} = O(c_{\rm sd})$ we have

$$\lim_{c_{\rm sd}, c_{\rm se} \rightarrow 0} \mathbb{P} \left\{ \frac{\mathscr{C}_{\Phi^*}(\tau^*)}{\mathscr{C}_{\hat{\Phi}}(\hat{\tau})} \leq 1 + \Delta \right\} = 1 \,, \quad \forall \Delta > 0 \,. \tag{37}$$

Proof: We provide a constructive proof. We show that the switching rule given in (36) leads to the convergence properties that are the basis of the proof for Theorem 5. See Appendix E for the details of the proof. ■

Based on the results of Theorem 1–Theorem 6, the steps involved in the optimal sampling strategy and the decision rules involved are summarized in Algorithm 1.

Algorithm 1: Quickest Search and Learning Algorithm.

1 Set $t = 1$ and $s_1 = 1$
2 Take one sample from sequence s_t
3 Update π_t^i for $i \in \mathcal{S}(t,T)$ and find $\tilde{\boldsymbol{\epsilon}}(t)$
4 **If** $c_{\rm d}(1 - \max_{i \in \mathcal{S}(t,T)} \pi_t^i) + c_{\rm e} \frac{\mathsf{R}_\Phi(t)}{t+1} \leq c_{\rm s}$
4   **If** $c_{\rm e} \mathsf{R}_\Phi(t) > c_{\rm s}(t+1)$
5   **If** $\mathcal{I}^1(\hat{\boldsymbol{\epsilon}}_{\rm ML}(t)) > \mathcal{I}^0(\hat{\boldsymbol{\epsilon}}_{\rm ML}(t))$ **Then** Set $s_{t+1} = s_t$
6       $t \leftarrow t + 1$
7   Go to Step 2
8   **Else**
9   **If** $\pi_t^{s_t} < \frac{\tilde{\epsilon}_0(t)}{1 - \tilde{\epsilon}_1(t) + \tilde{\epsilon}_0(t)}$ **Then** Set $s_{t+1} = s_t + 1$
10       $t \leftarrow t + 1$
11   Go to Step 2
12   **End** if
13 **End** if
14 Set $\tau = t$
15 Declare sequence $\left(\arg\max_{i \in \mathcal{S}(\tau,T)} \pi_\tau^i\right)$ as an outlier
16 Return $\mathbb{E}\{\boldsymbol{\epsilon} \mid \mathcal{F}_\tau\}$ as the estimate of $\boldsymbol{\epsilon}$

VII. Application to Spectrum Sensing

In this section we apply the developed theory to spectrum sensing, and specifically the quickest search for identifying spectrum opportunities in a wideband spectrum, while the occupancy states of the narrowband channels are co-dependent with an unknown dependency kernel. Quick search for spectrum access is motivated by the fact that the demand for communication bandwidth is constantly increasing. This necessitates more efficient utilization of the available spectrum band. Since the spectrum band is sometimes under-utilized, identifying the unused spectrum bands (i.e., spectrum holes) for opportunistic access by radios seeking spectrum alleviates the scarcity issue. However, the occupancy statuses of the spectrum holes vary rapidly. Hence, the spectrum holes might not remain free for long duration, as a result of which it is of paramount importance to devise mechanisms that can identify the spectrum holes quickly. Essentially, any delay in identifying the spectrum holes leads to under-utilization of the spectrum and reduced spectrum efficiency. In the following subsections, we show how the spectrum sensing problem can be mapped into the quickest search problem studied in this paper.

A. Related Work

Applications of sequential sampling in spectrum sensing are studied rather extensively [28]–[36]. Specifically, the studies in [28] and [29] focus on *only one* narrowband channel, and the former performs a binary hypothesis test while the latter applies quickest *change-point* detection to identify free bands. The study in [30] considers the optimal order of sensing in wideband spectrum sensing. A heuristic sequential sensing algorithm based on a chi-squared test is proposed in [31], which enjoys low-complexity and simple implementation. The problem of quickest search for one spectrum hole when the prior probability of occupancy of channels is known is studied in [32], and the results are extended to a two-stage mixed-observation setting in [33]. Cooperative sequential sensing mechanisms are studied in [35] and [36].

All the aforementioned studies conform in the assumption that the occupancy states of different narrowband channels are statistically independent. However, this assumption might be violated in practice, especially in broadband communication systems, in which the channels are bundled together and assigned to different users. The study in [34] considers co-dependent among the occupancy states, and leverages this dependency structure to design more efficient spectrum sensing algorithms. This paper, similarly to [34], assumes co-dependence of the occupancy states, with the distinction that the parameters of the dependency kernels are unknown. Since the algorithm designed in [34] strongly hinges on these parameters, in this section we aim to design alternative strategies for finding spectrum holes.

B. Spectrum Model

A wideband spectrum consisting of n *ordered* and non-overlapping narrowband channel is considered. The narrowband channels are allocated to different users based on the scheduling and interference management policies. We denote the true occupancy state of channel i by T_i, where $\mathsf{T}_i = \mathsf{H}_0$ indicates that channel i is busy and $\mathsf{T}_i = \mathsf{H}_1$ indicates that the channel is idle. By accounting for the uncertainties in the usage of each narrowband channel, we model the occupancy state T_i for $i \in \{1, \dots, n\}$ by a binary random variable and assume that the occupancy states of different narrowband channels are not necessarily independent. This could be due to the fact that in broadband communication systems, such as orthogonal frequency division multiplexing (OFDM), a user may get access to a number of channels simultaneously based on its traffic need. The network operator bundles adjacent narrowband channels and assigns them to such users. Therefore, a channel being deemed as busy provides some side information about the occupancy states of its adjacent channels. Motivated by such underlying coupling among the occupancy states of the adjacent narrowband channels, we consider the following dependency kernel between the occupancy states of adjacent channels. For any $i \in \{2, \dots, n\}$ we have

$$\mathbb{P}(\mathsf{T}_i = \mathsf{H}_0 \mid \mathsf{T}_{i-1} = \mathsf{H}_j) \;=\; \epsilon_j \quad \text{for} \quad j \in \{0, 1\}\,, \tag{38}$$

where $\epsilon_0, \epsilon_1 \in [0, 1]$ control the dependency level. The actual values of ϵ_0 and ϵ_1 depend on the variations of the traffic and scheduling patterns of different narrowband channels over time, which are not necessarily known. Hence, we assume that they are random quantities. Finally we assume that channel 1 is a idle channel with prior probability ϵ_0.

C. Sensing Model

A communication radio seeking spectrum opportunities scans the spectrum via sequentially tuning its receiver's filters to different channels and collecting information via channel measurements. Sequence of measurements collected from channel i is denoted by $\mathcal{X}^i = \{X_1^i, X_2^i, \dots\}$. We assume that the occupancy status of a channel remains the same while it is under observation. Then, if the samples collected from an occupied channel and a vacant channel are modeled by cdfs F_0 and F_1, respectively, the measurements obey the following dichotomous hypothesis model

$$\begin{array}{lll} \mathsf{H}_0: & X_j^i \sim F_0 & \\ \mathsf{H}_1: & X_j^i \sim F_1 & \end{array} \quad \text{for} \;\; j \in \mathbb{N}\,, \tag{39}$$

which is exactly the same as (1). The ultimate objective of spectrum sensing is to identify one idle channel. In practice, it is more convenient for a radio to sweep the spectrum band by applying a linear search as it requires to tune its receiver filter to the channel under scrutiny. Also, without loss of generality, we assume that a secondary user collects one measurement at-a-time[1]. Therefore, at each time instant the spectrum seeking radio should decide among detection, observation, and exploration actions, as defined in Section II-B. This search procedure, is specified by its stopping time τ and a switching rule $\psi(t)$. The objective of the spectrum sensing may be to identify one spectrum hole, or to estimate the occupancy pattern parameters $\boldsymbol{\epsilon}$, or both. Since the communication network is a dynamic system, the occupancy pattern changes rapidly. Hence, it is of paramount importance to detect the holes as quickly as possible. By integrating all these figures of merit into one cost function, the quickest spectrum sensing and parameter learning strategy can be obtained by optimizing the cost function in (11).

D. Numerical Evaluations

In this subsection, we apply the proposed framework strategies to wideband spectrum sensing. We consider a wideband

[1]Generalization to taking ℓ samples at-a-time is straightforward.

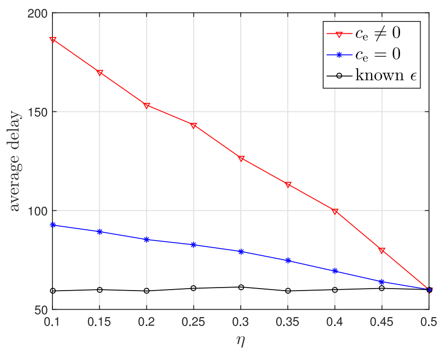


Fig. 1. Average delay versus uncertainty level of prior probability η.

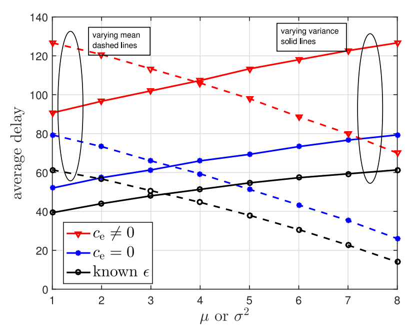


Fig. 2. Average delay versus noise power σ^2 and signal power μ.

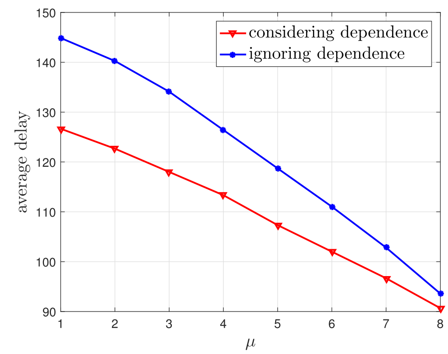


Fig. 3. Comparing average delays.

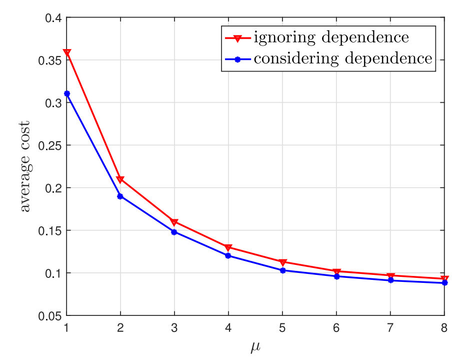


Fig. 4. Comparing average cost.

spectrum consisting of $n = 1000$ narrowband channels. The measurements from the idle channels are assumed to be drawn from $\mathcal{N}(0, \sigma^2)$[2], and the measurements from a busy channel are generated by $\mathcal{N}(\mu, \sigma^2)$, where $\mu > 0$. We also assume that both dependency kernel parameters ϵ_0 and ϵ_1 have uniform distributions over $[\eta, 1-\eta]$ for some $\eta \in [0, 0.5]$. We also set the sliding detection window interval to $T = 1$. Finally, we set the weights associated with the costs of sampling, detection, and estimation to $c_{\rm s} = 0.01$, $c_{\rm d} = 1$, and $c_{\rm e} = 1$, respectively. In Figs. 1 and 2 we evaluate the variations of the delay of the search process τ versus varying values of η, σ^2, and μ. In these figures the delay is compared with that of the purely detection setting in which $c_{\rm e} = 0$, when both have the same rate of erroneous detection decisions. As a benchmark, we also include the delay of the process in which $\boldsymbol{\epsilon}$ is perfectly known [5].

First, we set $\mu = 1$ and $\sigma^2 = 1$, and evaluate the delay for $\eta \in [0.1, 0.5)$. Fig. 1 shows that by increasing the uncertainty of the dependency kernel, more samples are required even in the purely detection setting in which we are not concerned about estimating $\boldsymbol{\epsilon}$. This is due to the fact that the detection quality implicitly depends on forming sufficiently reliable estimates for $\boldsymbol{\epsilon}$. Also, it is observed that when $c_{\rm e} \neq 0$, for forming reliable estimates for $\boldsymbol{\epsilon}$, we require to take more measurements compared to the case that estimation cost is unintegrated from the total cost ($c_{\rm e} = 0$). As expected, for the setting with known $\boldsymbol{\epsilon}$, the performance remains unchanged for varying η.

Next, we evaluate the impact of signal to noise ratio of the measurements on the average delay. To this end, in Fig. 2 we set $\eta = 0.3$ and compare the average delay of different settings for varying signal power μ and noise power σ^2. It is observed that the average delay increases as the measurements become more noisy (higher noise power σ^2), which is due to the fact that the presence of more noise renders the task of distinguishing between a busy channel and an idle channel more difficult. Also, it is observed that for larger signal power the three sequential methods reach the decisions faster.

In Fig. 3, in the given setting and for $c_{\rm s} = 0.001$ we evaluate the performance of the proposed rule and compare the results with the setting that ignores the existence of the dependency kernel in the generation of sequences. In this setting, it is assumed that each sequence is generated according to F_1 independently of the rest of the sequences and with prior probability ϵ, where ϵ is unknown. We assume that the goal is the same and the only difference is that since we only have one unknown parameter to estimate, we set $c_{\rm e} = 2$ in contrast to the setting of our problem where $c_{\rm e} = 1$. It is observed that ignoring the dependence among the sequences incurs more delay for detection and estimation processes, and our improvement is specially more significant when then two distributions are less-distinguishable (more similar), which is the more challenging case. The reason for larger delay is that by ignoring the dependence, we discard all the information about the desired sequence collected from each sequence after discarding it. Finally, in Fig. 4 we compare the average aggregate cost for the proposed rule and the setting that ignores the existence of dependence. It is observed that the average cost behaves similarly to the average delay.

VIII. Conclusion

We have analyzed the problem of quickest search over multiple correlated sequences, in which each sequence is

[2] $\mathcal{N}(a, b)$ denotes Gaussian distribution with mean a and variance b.

generated according to one of the two possible distributions F_0 and F_1. Generations of the sequences are not independent, and they follow a dependency kernel that is unknown. The main objective is to identify one sequence generated according to F_1 with the fewest number of measurements. Motivated by the fact that designing the detection rules strongly hinges on knowing the kernel, achieving the detection objective also necessitates producing reliable estimates for the dependency kernel parameters. For this purpose, we have also integrated a learning routine in the search process, the purpose of which is estimating the parameters of the dependency kernel. We have characterized the optimal detection and estimation rules, and have designed asymptotically optimal sequential mechanisms that at each time dynamically decide which sequence should be sampled. The decision rules are characterized in three stages. First, we have considered a purely detection setting in which the estimation of the parameters is not a concern, and we have shown that the decision rules reduce to a thresholding policy. In the next setting, we have focused on the reliable estimation of the unknown parameters and have shown that the optimal procedure consists of a random switching rule that switches to the following sequence with a parameter that minimizes a function of Fisher information matrices, and stops when the cost of estimation falls below the total cost of sampling. Finally, we have combined the results of the first two settings to characterize the sampling strategy for the quickest search problem in its general form. We have also investigated the application of the quickest search framework in wideband spectrum sensing.

APPENDIX A
PROOF OF THEOREM 2

When the stopping time and the switching sequence are fixed, given the collected data up to time t, the detection cost and the estimation cost are independent. Hence, in order to minimize $\mathscr{C}_\Phi(\tau)$, we require to minimize the detection and estimation costs separately. In order to find the optimal decision rule that minimizes the detection cost $\mathsf{P}_\Phi(\tau)$, we consider a randomized decision rule $\tilde{\delta}(\tau)$ that decides in favor of sequence j from the set $\mathcal{S}(\tau,T)$ with probability p_j where $\sum_{j\in\mathcal{S}(\tau,T)} p_j = 1$. For this randomized decision rule, the detection error probability is lower bounded by

$$\begin{aligned}\mathsf{P}_\Phi(\tau) &= \sum_{j\in\mathcal{S}(\tau,T)} p_j \cdot \mathbb{P}(\mathsf{T}_j = \mathsf{H}_0 \mid \mathcal{F}_\tau) \\ &\overset{(17)}{=} \sum_{j\in\mathcal{S}(\tau,T)} p_j(1-\pi_t^j) \\ &\geq \sum_{j\in\mathcal{S}(\tau,T)} p_j\left(1-\max_{i\in\mathcal{S}(\tau,T)} \pi_t^i\right) \\ &= 1-\max_{i\in\mathcal{S}(\tau,T)} \pi_t^i \; . \end{aligned} \tag{40}$$

This lower bound is achieved by setting

$$p_j = \begin{cases} 1, & \text{for } j = \arg\max_{i\in\mathcal{S}(\tau,T)} \pi_t^i \\ 0, & \text{otherwise} \end{cases} . \tag{41}$$

Hence, the decision rule specified in (22) achieves this lower bound. Similarly, for the estimator we also have

$$\begin{aligned}\mathsf{R}_\Phi(\tau) &= \mathbb{E}\left\{\|\hat{\boldsymbol{\epsilon}}(\tau)-\boldsymbol{\epsilon}\|^2 \mid \mathcal{F}_\tau\right\} &(42)\\ &= \mathbb{E}\left\{\|\hat{\boldsymbol{\epsilon}}(\tau)-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}+\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}-\boldsymbol{\epsilon}\|^2 \mid \mathcal{F}_\tau\right\} &(43)\\ &= \mathbb{E}\left\{\|\hat{\boldsymbol{\epsilon}}(\tau)-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\|^2 \mid \mathcal{F}_\tau\right\} &(44)\\ &\quad + \mathbb{E}\left\{\|\boldsymbol{\epsilon}-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\|^2 \mid \mathcal{F}_\tau\right\} &(45)\\ &\quad + 2\cdot\mathbb{E}\left\{(\hat{\boldsymbol{\epsilon}}(\tau)-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\})^T(\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\right. \\ &\quad \left.-\boldsymbol{\epsilon})\mid\mathcal{F}_\tau\right\} &(46)\\ &= \mathbb{E}\left\{\|\hat{\boldsymbol{\epsilon}}(\tau)-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\|^2 \mid \mathcal{F}_\tau\right\} &(47)\\ &\quad + \mathbb{E}\left\{\|\boldsymbol{\epsilon}-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\|^2 \mid \mathcal{F}_\tau\right\} &(48)\\ &\quad + 2\cdot\mathbb{E}\left\{(\hat{\boldsymbol{\epsilon}}(\tau)-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\})^T\right\} &(49)\\ &\quad \times \underbrace{\mathbb{E}\left\{(\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}-\boldsymbol{\epsilon})\mid\mathcal{F}_\tau\right\}}_{0} &(50)\\ &= \mathbb{E}\left\{\|\hat{\boldsymbol{\epsilon}}(\tau)-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\|^2 \mid \mathcal{F}_\tau\right\} &(51)\\ &\quad + \mathbb{E}\left\{\|\boldsymbol{\epsilon}-\mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\}\|^2 \mid \mathcal{F}_\tau\right\} . &(52)\end{aligned}$$

The term in (52) is independent of the estimator. Hence, in order to minimize the estimation cost, we only need to minimize (51). The minimum value for (51) is zero and is achieved by setting

$$\hat{\boldsymbol{\epsilon}}(\tau) = \mathbb{E}\{\boldsymbol{\epsilon}\mid\mathcal{F}_\tau\} \; . \tag{53}$$

We observe that the detection rule that minimizes the detection cost and the estimation rule that minimizes the estimation cost are independent. Therefore, they are decoupled and the proof is established.

APPENDIX B
PROOF OF THEOREM 3

The cost function in (11) in a pure detection setting, i.e., $c_{\rm e} = 0$, becomes

$$\mathscr{C}_\Phi(t) = c_{\rm d}\cdot\mathsf{P}_\Phi(t) + c_{\rm s}\cdot t \; .$$

If we show that

$$\frac{1}{\tau_{\rm d}^*}\log\mathsf{P}_\Phi(\tau_{\rm d}^*) \xrightarrow{c_{\rm sd}\to 0} -D_{\rm KL}(f_1\|f_0) \; , \quad \text{a.s.} \; , \tag{54}$$

then, the assumptions of Theorem 2.1 (ii) in [13] will be satisfied and the proof is concluded. The proof of (54) is carried out in two steps.

Step 1: In the first step, we assume that L sequences are visited during the sampling process, and τ_ℓ, for $\ell\in\mathcal{L}\overset{\triangle}{=}\{1,\dots,L\}$, is the number of samples taken from sequence ℓ. Hence,

$$\tau_{\rm d}^* = \sum_{\ell=1}^{L}\tau_\ell \; .$$

We denote the index of the sequence declared as an abnormal sequence by k. We show that the values of τ_ℓ, for all $\ell\in\mathcal{L}\setminus\{k\}$, are exponentially bounded and the delay is dominated by the number of samples taken from the sequence that will be declared as abnormal. For this purpose, we define

$$\mathcal{A}_\ell = \sum_{i=1}^{\ell-1} \tau_i, \tag{55}$$

and accordingly we define Λ_ν^ℓ as the likelihood ratio when the we draw ν samples from sequence ℓ, i.e.,

$$\Lambda_\nu^\ell \triangleq \prod_{u=\mathcal{A}_\ell+1}^{\mathcal{A}_\ell+\nu} \frac{f_1(Y_u)}{f_0(Y_u)}. \tag{56}$$

It can be readily verified that thresholding the posterior probability with two thresholds is equivalent to thresholding the likelihood ratio Λ_ν^ℓ with thresholds [5], which we denote by γ_L^ℓ and γ_U^ℓ. Hence, for any $\ell \in \mathcal{L}$ we have

$$\begin{aligned}
\mathbb{P}(\tau_\ell > u) &= \mathbb{P}\big(\Lambda_\nu^\ell \in (\gamma_L^\ell, \gamma_U^\ell),\ \forall \nu \in \{1, \dots, u\}\big) && (57)\\
&\leq \mathbb{P}(\Lambda_u^\ell \in (\gamma_L^\ell\ ,\ \gamma_U^\ell)) && (58)\\
&= \mathbb{P}(\Lambda_u^\ell > \gamma_L^\ell\ ,\ \Lambda_u^\ell < \gamma_U^\ell) && (59)\\
&= q_\ell \cdot \mathbb{P}_1(\Lambda_u^\ell > \gamma_L^\ell\ ,\ \Lambda_u^\ell < \gamma_U^\ell) && (60)\\
&\quad + (1-q_\ell)\mathbb{P}_0(\Lambda_u^\ell > \gamma_L^\ell\ ,\ \Lambda_u^\ell < \gamma_U^\ell) && (61)\\
&\leq \mathbb{P}_1(\Lambda_u^\ell < \gamma_U^\ell) + \mathbb{P}_0(\Lambda_u^\ell > \gamma_L^\ell)\ , && (62)
\end{aligned}$$

where q_ℓ is the prior probability of sequence ℓ being drawn from F_1, and $\mathbb{P}_j$ is the probability measure corresponding to cdf F_j. On the other hand,

$$\begin{aligned}
\mathbb{P}_0(\Lambda_u^\ell > \gamma_L^\ell) &= \mathbb{P}_0\left(\sqrt{\Lambda_u^\ell} > \sqrt{\gamma_L^\ell}\right) && (63)\\
&\leq \frac{1}{\sqrt{\gamma_L^\ell}} \mathbb{E}_0^u\left[\sqrt{\Lambda_1^\ell}\right] && (64)\\
&= \frac{1}{\sqrt{\gamma_L^\ell}} \rho^u\ , && (65)
\end{aligned}$$

where (64) is due to the Markov inequality, and the set of likelihood ratios being i.i.d., and in (65) we have defined

$$\begin{aligned}
\rho &\triangleq \mathbb{E}_0\left[\sqrt{\Lambda_1^\ell}\right] && (66)\\
&= \int \sqrt{\frac{f_1(y)}{f_0(y)}} f_0(y)\, dy && (67)\\
&= \int \sqrt{f_1(y) f_0(y)}\, dy && (68)\\
&< \sqrt{\int f_1(y)\, dy} \sqrt{\int f_0(y)\, dy} && (69)\\
&= 1\ , && (70)
\end{aligned}$$

where (69) holds owing to the Cauchy-Schwartz inequality. By following the same procedure, we have

$$\begin{aligned}
\mathbb{P}_1(\Lambda_u^\ell < \gamma_U^\ell) &= \mathbb{P}_1([\Lambda_u^\ell]^{-\frac{1}{2}} > [\gamma_U^\ell]^{-\frac{1}{2}}) && (71)\\
&\leq \sqrt{\gamma_U^\ell} \cdot \mathbb{E}_1^u\{[\Lambda_1^\ell]^{-\frac{1}{2}}\} && (72)\\
&= \sqrt{\gamma_U^\ell} \cdot \rho^u\ . && (73)
\end{aligned}$$

Based on (62), (65), and (73), it is observed that

$$\mathbb{P}(\tau_\ell > u) < \left(\frac{1}{\sqrt{\gamma_L^\ell}} + \sqrt{\gamma_U^\ell}\right) \rho^u\ , \tag{74}$$

which shows that $\mathbb{P}(\tau_\ell > u)$ is exponentially-bounded and finite.

Step 2: In the second step we show that we have the following convergence almost surely (a.s.):

$$\frac{1}{\tau_{\rm d}^*} \log \mathsf{P}_\Phi(\tau_{\rm d}^*) \xrightarrow{c_{\rm sd} \to 0} -D_{\rm KL}(f_1 \| f_0)\ , \quad \text{a.s.} \tag{75}$$

This can be proved by directly applying [13, Theorem 4.2]. Here, we provide a simpler proof. First, we note that since τ_ℓ for $\ell \in \mathcal{L} \setminus \{k\}$ is exponentially bounded and it satisfies

$$\mathbb{P}\Big(\lim_{c_{\rm sd} \to 0} \tau_{\rm d}^* = \infty\Big) = 1\ , \tag{76}$$

then we have

$$\frac{\tau_k}{\tau_{\rm d}^*} \xrightarrow{c_{\rm sd} \to 0} 1\ , \quad \text{a.s.} \tag{77}$$

On the other hand, based on the following bounds on $\pi_{\tau_{\rm d}^*}^k$

$$\min_{\boldsymbol{\epsilon}} \kappa_{\tau_{\rm d}^*}^k(\boldsymbol{\epsilon}) \leq \pi_{\tau_{\rm d}^*}^k \leq \max_{\boldsymbol{\epsilon}} \kappa_{\tau_{\rm d}^*}^k(\boldsymbol{\epsilon})\ , \tag{78}$$

for some $\boldsymbol{\epsilon}^* \in [0,1]^2$ we have $\pi_{\tau_{\rm d}^*}^k = \kappa_{\tau_{\rm d}^*}^k(\boldsymbol{\epsilon}^*)$. Therefore, $\pi_{\tau_{\rm d}^*}^k$ can be written as

$$\pi_{\tau_{\rm d}^*}^k = \frac{\theta \Lambda_{\tau_{\rm d}^*}^k}{(1-\theta) + \theta \Lambda_{\tau_{\rm d}^*}^k}, \tag{79}$$

for some $\theta \in (0,1)$. Also, as $c_{\rm sd}$ approaches zero, the value of $\pi_{\tau_{\rm d}^*}^k$ tends to 1. This indicates that the value of $\Lambda_{\tau_{\rm d}^*}^k$ tends to infinity. Hence, we have

$$\frac{1}{\tau_k} \log(1 - \pi_{\tau_{\rm d}^*}^k) \xrightarrow{c_{\rm sd} \to 0} -\frac{1}{\tau_k} \log \Lambda_{\tau_{\rm d}^*}^k\ , \quad \text{a.s.} \tag{80}$$

Next, by leveraging (77) and (80), as $c_{\rm sd} \to 0$ we have

$$\begin{aligned}
\frac{\log \mathsf{P}_\Phi(\tau_{\rm d}^*)}{\tau_{\rm d}^*} &= \frac{\log \mathsf{P}_\Phi(\tau_{\rm d}^*)}{\tau_k} \cdot \frac{\tau_k}{\tau_{\rm d}^*} && (81)\\
&\to \frac{\log \mathsf{P}_\Phi(\tau_{\rm d}^*)}{\tau_k}, \quad \text{a.s.} && (82)\\
&= \frac{1}{\tau_k} \log(1 - \pi_{\tau_{\rm d}^*}^k) && (83)\\
&\to -\frac{1}{\tau_k} \log \Lambda_{\tau_{\rm d}^*}^k\ , \quad \text{a.s.} && (84)\\
&\to -D_{\rm KL}(f_1 \| f_0)\ , \quad \text{a.s.,} && (85)
\end{aligned}$$

where the almost sure convergence in (82) holds because of (76) and (77), the equality in (83) holds due to the definition of π_t^i in (9) and (17), the convergence in (80) leads to (84), and the last convergence (85) is due to the law of large numbers. The convergence of

$$\frac{\log \mathsf{P}_\Phi(\tau_{\rm d}^*)}{\tau_{\rm d}^*}$$

in (85) ensures that the conditions of [13, Theorem 2.1 (ii)] hold, and therefore, the optimality of the proposed rule is established.

Appendix C
Proof of Theorem 4

In a pure estimation setting, i.e., $c_{\rm e} = 0$, the cost function in (11) becomes

$$\mathscr{C}_\Phi(t) = c_{\rm e} \cdot \mathsf{R}_\Phi(t) + c_{\rm s} \cdot t\ .$$

If we show that $t \cdot \mathsf{R}_\Phi(t)$ for the given stopping time and switching rule convegrges to a positive random variable as $c_{\rm se} \rightarrow 0$, then the assumptions of Theorem 2.1 (i) in [13] will be satisfied and the proof is concluded. To prove the convergence of $t \cdot \mathsf{R}_\Phi(t)$, we define

$$V(\boldsymbol{\epsilon}) \triangleq \inf_{p\in[0,1]} \left[p \cdot \mathcal{I}^1(\boldsymbol{\epsilon}) + (1-p)\cdot \mathcal{I}^0(\boldsymbol{\epsilon})\right]^{-1} \ . \tag{86}$$

Then, from the Cramer-Rao lower bound [37], $\forall \Delta > 0$ we have

$$\mathbb{P}_{\boldsymbol{\epsilon}}\left(t \cdot \inf_\Phi \ \mathsf{R}_\Phi(t) \geq \text{trace}(V(\boldsymbol{\epsilon})) - \Delta\right) \xrightarrow{t\rightarrow\infty} 1\ . \tag{87}$$

Next, we show that for the proposed switching sequence $\{\psi_{\rm e}(t)\ :\ t \in \mathbb{N}\}$ and any $\Delta > 0$ we have

$$\lim_{t\rightarrow\infty} \mathbb{P}_{\boldsymbol{\epsilon}}\left(|t \cdot \mathsf{R}_\Phi(t) - \text{trace}(V(\boldsymbol{\epsilon}))| \geq \Delta\right) = 0\ . \tag{88}$$

For this purpose, note that due to the consistency of the MLE we have

$$\mathcal{I}^\ell(\hat{\boldsymbol{\epsilon}}_{\rm ML}(t)) \xrightarrow{t\rightarrow\infty} \mathcal{I}^\ell(\boldsymbol{\epsilon})\ , \quad \text{a.s.} \tag{89}$$

On the other hand, by specifying r_t as the frequency of switching between sequences in the first t instants, due to the consistency of MLE, and also the continuity of the likelihood function $\Lambda_{\boldsymbol{\epsilon}}(\cdot)$ over $\boldsymbol{\epsilon}$ we have

$$r_t \xrightarrow{t\rightarrow\infty} p^*(\boldsymbol{\epsilon})\ , \quad \text{a.s.} \tag{90}$$

Therefore, for a sufficiently large number of samples and observed sequences, the frequency of using the correct switching rule is arbitrarily close to the optimal switching rule. Hence, it only remains to establish the asymptotic efficiency of the proposed rules [15], i.e.,

$$t \cdot \mathsf{R}_\Phi(t) \xrightarrow{t\rightarrow\infty} \text{trace}(V(\boldsymbol{\epsilon}))\ , \quad \mathbb{P}_{\boldsymbol{\epsilon}}, \tag{91}$$

where the convergence is in probability under measure $\mathbb{P}_{\boldsymbol{\epsilon}}$. To prove (91), we first note that, due to a weak version of Bernstein-von Mises Theorem [38, Theorem 20.2], the posterior distribution of the estimator converges to a normal distribution, i.e.,

$$g\left(\sqrt{t}(\hat{\boldsymbol{\epsilon}}(t) - \boldsymbol{\epsilon}) \mid \mathcal{F}_t\right) \xrightarrow{t\rightarrow\infty} \mathcal{N}\left(0, (\mathcal{I}^*(\boldsymbol{\epsilon}))^{-1}\right)\ , \quad \mathbb{P}_{\boldsymbol{\epsilon}}, \tag{92}$$

where

$$\mathcal{I}^*(\boldsymbol{\epsilon}) \triangleq p^*(\boldsymbol{\epsilon})\mathcal{I}^1(\boldsymbol{\epsilon}) + (1 - p^*(\boldsymbol{\epsilon}))\mathcal{I}^0(\boldsymbol{\epsilon})\ . \tag{93}$$

The convergence of the estimation risk in (91) is proved by establishing the finiteness of the moments of the posterior distribution of $\sqrt{t}(\hat{\boldsymbol{\epsilon}}(t) - \boldsymbol{\epsilon})$. Specifically, for the purpose of this theorem, it suffices to show that the first two moments of the posterior distribution converge to that of a zero-mean normal Gaussian distribution with covariance matrix $(\mathcal{I}^*(\boldsymbol{\epsilon}))^{-1}$, i.e., for $m \in \{1, 2\}$

$$\int (\sqrt{t}(\hat{\boldsymbol{\epsilon}}(t) - \boldsymbol{\epsilon}))^m \cdot g\left(\sqrt{t}(\hat{\boldsymbol{\epsilon}}(t) - \boldsymbol{\epsilon}) \mid \mathcal{F}_t\right) \xrightarrow{t\rightarrow\infty} \mathbb{E}\{(\sqrt{t}(\hat{\boldsymbol{\epsilon}}(t) - \boldsymbol{\epsilon}))^m\}\ , \quad \mathbb{P}_{\boldsymbol{\epsilon}}, \tag{94}$$

where the expectation in (94) is under distribution $\mathcal{N}\left(0, (\mathcal{I}^*(\boldsymbol{\epsilon}))^{-1}\right)$. This is proved by applying the Helly-Bray Theorem and following the same line of argument as in [13].

Appendix D
Proof of Theorem 5

The detection cost depends strongly on sequence k which is the one declared as the desired sequence, while the estimation cost relies on all the observed sequences. When the sampling cost is substantially smaller than the costs of estimation and detection decisions, from the proofs of Theorems 3 and 4, for the optimal switching rule we can conclude that

$$\frac{1}{t}\log \mathsf{P}_\Phi(t) \xrightarrow{t\rightarrow\infty} -\gamma D_{\rm KL}(f_1\|f_0)\ , \quad \text{a.s.}, \tag{95}$$

and

$$t \cdot \mathsf{R}_\Phi(t) \xrightarrow{t\rightarrow\infty} \hat{V}(\boldsymbol{\epsilon})\ , \quad \mathbb{P}_{\boldsymbol{\epsilon}}, \tag{96}$$

for some $0 < \hat{V}(\boldsymbol{\epsilon}) < \infty$ and $\gamma \in [0, 1]$, where γ depends on the relative values of $c_{\rm d}$ and $c_{\rm e}$. To prove (95), let us denote the first time instant at which the first sample is taken from sequence $\delta(\tau^*)$ by t_1. Then, from the proof of Theorem 3, as $\tau^* - t_1 \rightarrow \infty$ we have

$$\frac{1}{\tau^* - t_1}\log \mathsf{P}_\Phi(\tau^*) \rightarrow -D_{\rm KL}(f_1\|f_0)\ , \quad \text{a.s.} \tag{97}$$

Now, by defining

$$\gamma \triangleq 1 - \frac{t_1}{\tau^*}, \tag{98}$$

and noting that $\gamma \in [0, 1]$ the proof of (95) is concluded. If (95) and (96) are deterministic and hold for any t, and not only in the asymptotes of large t, the stopping time would be the first time t for which we have

$$c_{\rm d}\mathsf{P}_\Phi(t) + c_{\rm e}\mathsf{R}_\Phi(t) + c_{\rm s}(t) \leq c_{\rm d}\mathsf{P}_\Phi(t+1) + c_{\rm e}\mathsf{R}_\Phi(t+1) + c_{\rm s}\cdot(t+1)\ , \tag{99}$$

which by replacing P_Φ and R_Φ from (95) and (96) reduces to

$$c_{\rm d}\mathsf{P}_\Phi(t)(1 - \mathsf{P}_\Phi(t)^{\frac{1}{t}}) + c_{\rm e}\frac{\mathsf{R}_\Phi(t)}{t+1} \leq c_{\rm s}. \tag{100}$$

The remainder of the proof involves showing that for the asymptotic convergence of random sequences in (95) and (96) we have asymptotic optimality. To prove this, we define

$$W_t \triangleq c_{\rm d} \mathsf{P}_\Phi(t) \;, \tag{101}$$

$$\text{and} \qquad Z_t \triangleq c_{\rm e} \mathsf{R}_\Phi(t) \;, \tag{102}$$

for which, from (95) and (96) and some $W < 0$ and $Z > 0$ we have

$$\frac{1}{t} \log W_t \xrightarrow{t\to\infty} W \;, \quad \text{a.s.}, \tag{103}$$

$$\text{and} \quad tZ_t \xrightarrow{t\to\infty} Z \;, \quad \mathbb{P}_{\boldsymbol\epsilon} \;. \tag{104}$$

For convenience in notations, in the remainder of the proofs we suppress the superscript $*$ from τ^*. Hence, at the stopping time

$$W_\tau + \frac{Z_\tau}{\tau} \leq c_{\rm s} \;. \tag{105}$$

For any other strategy $\hat\Phi$ with stopping time ν we have

$$\frac{\mathscr{C}_{\Phi^*}(\tau)}{\mathscr{C}_{\hat\Phi}(\nu)} = \frac{W_\tau + Z_\tau + c_{\rm s}\tau}{W_\nu + Z_\nu + c_{\rm s}\nu} \tag{106}$$

$$= \frac{\frac{W_\tau}{c_{\rm s}\tau} + \frac{Z_\tau}{c_{\rm s}\tau} + 1}{\frac{W_\nu}{c_{\rm s}\tau} + \frac{Z_\nu}{c_{\rm s}\tau} + \frac{c_{\rm s}\nu}{c_{\rm s}\tau}} \tag{107}$$

$$\leq \frac{\frac{W_\tau}{c_{\rm s}} + \frac{Z_\tau}{c_{\rm s}\tau} + 1}{\frac{W_\nu}{c_{\rm s}\tau} + \frac{Z_\nu}{c_{\rm s}\tau} + \frac{c_{\rm s}\nu}{c_{\rm s}\tau}} \tag{108}$$

$$\leq \frac{2}{\frac{W_\nu}{c_{\rm s}\tau} + \frac{Z_\nu}{c_{\rm s}\tau} + \frac{\nu}{\tau}}, \tag{109}$$

where (108) is the result of dropping τ from the first term of the numerator, and (109) holds due to (105). If we show that the denominator of (108) is greater than 2 the proof is complete as the ratio between the total costs associated with the proposed rule and any other stopping rule becomes less than 1. Since τ is the stopping time, we have

$$W_\tau + \frac{Z_\tau}{\tau} \leq c_{\rm s} < W_{\tau-1} + \frac{Z_{\tau-1}}{\tau-1}, \tag{110}$$

or, equivalently, $\tau^2 c_{\rm s}$ can be bounded as

$$\tau^2 W_\tau + \tau Z_\tau \leq \tau^2 c_{\rm s} < \tau^2 W_{\tau-1} + \frac{\tau^2}{(\tau-1)^2}(\tau-1) Z_{\tau-1} \;. \tag{111}$$

Since W_τ decreases exponentially as $\tau \to \infty$, both upper bound and lower bound on $\tau^2 c_{\rm s}$ converge to Z. Since we have

$$\frac{Z_\nu}{c_{\rm s}\tau} = \frac{\tau}{\nu} \cdot \frac{\nu Z_\nu}{\tau^2 c_{\rm s}} \xrightarrow{\tau\to\infty} \frac{\tau}{\nu}, \quad \mathbb{P}_{\boldsymbol\epsilon}, \tag{112}$$

then for the denominator of (108) we have

$$\frac{W_\nu}{c_{\rm s}\tau} + \frac{Z_\nu}{c_{\rm s}\tau} + \frac{\nu}{\tau} \geq \frac{\tau}{\nu} + \frac{\nu}{\tau} \geq 2 \;, \quad \mathbb{P}_{\boldsymbol\epsilon}, \tag{113}$$

which concludes the proof.

Appendix E
Proof of Theorem 6

The optimality of the stopping time is proved in Theorem 5 and we only need to prove that the proposed switching rule is optimal. The proof is by construction. We note that in order to have a reliable estimate of the prior probability $\boldsymbol\epsilon$ we require to observe many sequences, and as $c_{\rm se}$ tends to zero the number of observed sequences approaches infinity. On the other hand, we have observed in the proof of Theorem 3 that for a purely detection problem, the stopping time is dominated by the number of samples taken from the sequence detected as the desired one. Hence, the number of samples taken for finding one abnormal sequence has negligible effect on the estimation quality. On the other hand, since T is finite and $c_{\rm sd}$ tends to zero, only the last sequence under observation can be declared as the desired one. As a result, the switching rule should be concerned about the estimation quality at the beginning of the sampling process. When the estimation cost is sufficiently small and comparable to the sampling cost, it has to turn to identifying one outlier sequence. To formalize the proof, we need to establish that for the given switching rule we have

$$\frac{1}{t} \log \mathsf{P}_\Phi(t) \xrightarrow{t\to\infty} -\gamma D_{\rm KL}(f_1 \| f_0) \;, \quad \text{a.s.}, \tag{114}$$

$$\text{and} \quad t \cdot \mathsf{R}_\Phi(t) \xrightarrow{t\to\infty} \hat V(\boldsymbol\epsilon) \;, \quad \mathbb{P}_{\boldsymbol\epsilon}, \tag{115}$$

for some $0 < \hat V(\boldsymbol\epsilon) < \infty$ and $\gamma \in (0, 1)$. To this end, let us denote the first time instant at which

$$c_e R_\Phi(t) \leq (t + 1) c_s \tag{116}$$

by t_1. Then, from the proof of Theorem 4 we have

$$t_1 \cdot \mathsf{R}_\Phi(t_1) \xrightarrow{t_1\to\infty} \text{trace}(V(\boldsymbol\epsilon)) \;, \quad \mathbb{P}_{\boldsymbol\epsilon} \;. \tag{117}$$

On the other hand, from the proof of Theorem 3 we have

$$\frac{1}{\tau^* - t_1} \log \mathsf{P}_\Phi(\tau^* - t_1) \xrightarrow{\tau^*-t_1\to\infty} -D_{\rm KL}(f_1 \| f_0) \;, \quad \text{a.s.} \tag{118}$$

Let us define

$$\eta \triangleq \frac{\tau^*}{t_1} \;. \tag{119}$$

We show that η is finite, then by defining

$$\hat V(\boldsymbol\epsilon) \triangleq \eta \cdot V(\boldsymbol\epsilon) \;, \text{ and } \; \gamma \triangleq 1 - \frac{1}{\eta}, \tag{120}$$

the proof is concluded. We prove the finiteness of η by contradiction. Assume that η can be arbitrarily large. Then, since we have

$$c_e \mathsf{R}_\Phi(t_1) \leq (t_1 + 1) c_s \;, \tag{121}$$

we get

$$\frac{c_{\rm e} \mathsf{R}_\Phi(\tau^*)}{\tau^* + 1} = o(c_{\rm s}) \;. \tag{122}$$

On the other hand, while the estimation cost decreases with the rate of t^{-1}, the detection cost decreases exponentially in t.

Hence, if $\eta \to \infty$ we have

$$c_{\rm d} P_{\Phi}(\tau^*) = o(c_{\rm s}) \; . \tag{123}$$

Combination of (122) and (123) contradicts the stopping rule in (34).

REFERENCES

[1] A. Tajer, V. V. Veeravalli, and H. V. Poor, "Outlying sequence detection in large data sets: A data-driven approach," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 44–56, Sep. 2014.

[2] 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.

[3] A. Wald, "Sequential tests of statistical hypotheses," *Ann. Math. Statist.*, vol. 16, no. 2, pp. 117–186, Jun. 1945.

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

[5] J. Heydari, A. Tajer, and H. V. Poor, "Quickest linear search over correlated sequences," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5786–5808, Oct. 2016.

[6] H. Chernoff, "Sequential design of experiments," *Ann. Math. Statist.*, vol. 30, no. 3, pp. 755–770, Sep. 1959.

[7] S. Bessler, "Theory and applications of the sequential design of experiments, k-actions and infinitely many experiments: Part I theory," Dept. Statist., Stanford Univ., Stanford, CA, USA, Tech. Rep. SOL ONR 55, Mar. 1960.

[8] A. E. Albert, "The sequential design of experiments for infinitely many states of nature," *Ann. Math. Statist.*, vol. 32, no. 3, pp. 774–799, Sep. 1961.

[9] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2451–2464, May 2013.

[10] S. Nitinawarat and V. V. Veeravalli, "Controlled sensing for sequential multihypothesis testing with controlled Markovian observations and non-uniform control cost," *Sequential Anal.*, vol. 34, no. 1, pp. 1–24, Feb. 2015.

[11] K. Cohen and Q. Zhao, "Active hypothesis testing for anomaly detection," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1432–1450, Mar. 2015.

[12] J. G. Ligo, G. K. Atia, and V. V. Veeravalli, "A controlled sensing approach to graph classification," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6468–6480, Dec. 2014.

[13] P. J. Bickel and J. A. Yahav, "Asymptotically pointwise optimal procedures in sequential analysis," in *Proc. 5th Berkeley Symp. Math. Statist. Probab., vol. 1, Statist.*, Berkeley, CA, USA, 1967, pp. 401–413.

[14] V. J. Yohai, "Asymptotically optimal Bayes sequential design of experiments for estimation," *Ann. Statist.*, vol. 1, no. 5, pp. 822–837, Sep. 1973.

[15] G. Atia and S. Aeron, "Asymptotic optimality results for controlled sequential estimation," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput.*, Oct. 2013, pp. 1098–1105.

[16] K. Zigangirov, "On a problem in optimal scanning," *Theory Probab. Appl.*, vol. 11, no. 2, pp. 294–298, Jan. 1966.

[17] V. Dragalin, "A simple and effective scanning rule for a multi-channel system," *Metrika*, vol. 43, no. 1, pp. 165–182, Dec. 1996.

[18] Y. Li, S. Nitinawarat, and V. V. Veeravalli, "Universal sequential outlier hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 3205–3209.

[19] S. Nitinawarat and V. V. Veeravalli, "Universal scheme for optimal search and stop," in *Proc. Inf. Theory Appl. Workshop*, Feb. 2015, pp. 322–328.

[20] K. Cohen, Q. Zhao, and A. Swami, "Optimal index policies for anomaly localization in resource-constrained cyber systems," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4224–4236, Aug. 2014.

[21] S. Zou, Y. Liang, H. V. Poor, and X. Shi, "Nonparametric detection of anomalous data streams," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5785–5797, Nov. 2017.

[22] Y. Li, S. Nitinawarat, and V. V. Veeravalli, "Universal outlier hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4066–4082, Jul. 2014.

[23] M. L. Malloy, G. Tang, and R. D. Nowak, "Quickest search for a rare distribution," in *Proc. 46th Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2012, pp. 1–6.

[24] A. Tajer and H. V. Poor, "Quick search for rare events," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4462–4481, Jul. 2013.

[25] A. Tajer and H. V. Poor, "Hypothesis Testing for Partial Sparse Recovery," in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2012, pp. 901–908.

[26] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed. New York, NY, USA: Springer, 1998.

[27] K. J. Arrow, D. Blackwell, and M. A. Girshick, "Bayes and minimax solutions of sequential decision problems," *Econometrica*, vol. 17, no. 3/4, pp. 213–244, 1949.

[28] R. Caromi, Y. Xin, and L. Lai, "Fast multiband spectrum scanning for cognitive radio systems," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 63–75, Jan. 2013.

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

[30] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297–307, Jan. 2009.

[31] Y. Xin, H. Zhang, and L. Lai, "A low-complexity sequential spectrum sensing algorithm for cognitive radio," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 387–399, Mar. 2014.

[32] R. Caromi, S. Mohan, and L. Lai, "Optimal sequential channel estimation and probing for multiband cognitive radio systems," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2696–2708, Aug. 2014.

[33] J. Geng, W. Xu, and L. Lai, "Quickest sequential multiband spectrum sensing with mixed observations," *IEEE Trans. Signal Process.*, vol. 64, no. 22, pp. 5861–5874, Nov. 2016.

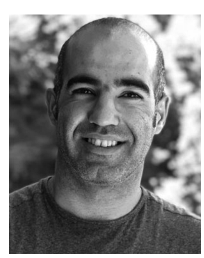
[34] A. Tajer and J. Heydari, "Quickest wideband spectrum sensing over correlated channels," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3082–3091, Sep. 2015.

[35] Y. Yilmaz, G. V. Moustakides, and X. Wang, "Cooperative sequential spectrum sensing based on level-triggered sampling," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4509–4524, Sep. 2012.

[36] Q. Zou, S. Zheng, and A. H. Sayed, "Cooperative sensing via sequential detection," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6266–6283, Dec. 2010.

[37] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer, 1994.

[38] A. DasGupta, *Asymptotic Theory of Statistics and Probability*. New York, NY, USA: Springer, 2008.



Javad Heydari (M'19) is currently a Senior AI/Machine Learning Research Scientist with LG Electronics, Santa Clara, CA, USA. During 2014–2018, he was with Rensselaer Polytechnic Institute, Troy, NY, USA, where he received the M.Sc. degree in mathematics and the Ph.D. degree in electrical engineering. His research interests lie in the general areas of statistical learning theory, machine learning, and information theory.

Ali Tajer (S'05–M'10–SM'15) received the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA, in 2010. He is currently an Assistant Professor of electrical, computer, and systems engineering with Rensselaer Polytechnic Institute, Troy, NY, USA. His research interests include mathematical statistics and network information theory, with applications in wireless communications and power grids. Dr. Tajer serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON SMART GRID. He was the recipient of an NSF CAREER award in 2016.