

# Searching for Anomalies over Composite Hypotheses

Bar Hemo, Tomer Gafni, Kobi Cohen, Qing Zhao

## Abstract

The problem of detecting anomalies in multiple processes is considered. We consider a composite hypothesis case, in which the measurements drawn when observing a process follow a common distribution with an unknown parameter (vector), whose value lies in normal or abnormal parameter spaces, depending on its state. The objective is a sequential search strategy that minimizes the expected detection time subject to an error probability constraint. We develop a deterministic search algorithm with the following desired properties. First, when no additional side information on the process states is known, the proposed algorithm is asymptotically optimal in terms of minimizing the detection delay as the error probability approaches zero. Second, when the parameter value under the null hypothesis is known and equal for all normal processes, the proposed algorithm is asymptotically optimal as well, with better detection time determined by the true null state. Third, when the parameter value under the null hypothesis is unknown, but is known to be equal for all normal processes, the proposed algorithm is consistent in terms of achieving error probability that decays to zero with the detection delay. Finally, an explicit upper bound on the error probability under the proposed algorithm is established for the

Bar Hemo, Tomer Gafni and Kobi Cohen are with the School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. Email: {barh, gafnito, yakovsec}@bgu.ac.il

Qing Zhao is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA. Email: qz16@cornell.edu

Bar Hemo and Tomer Gafni contributed equally to this paper.

Part of this work was presented at the 16th IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Limassol, Cyprus, Dec. 2016 [1].

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

The work of B. Hemo, T. Gafni, and K. Cohen was supported by the Israel National Cyber Bureau under grant 076/16, and by the U.S.-Israel Binational Science Foundation (BSF) under grant 2017723. The work of Q. Zhao was supported by the U.S. Army Research Office under Grant W911NF-17-1-0464, and the National Science Foundation under Grant CCF-1815559.

finite sample regime. Extensive experiments on synthetic dataset and DARPA intrusion detection dataset are conducted, demonstrating strong performance of the proposed algorithm over existing methods.

## I. INTRODUCTION

We consider the problem of searching for an anomalous process (or few abnormal processes) among  $M$  processes. For convenience, we often refer to the processes as cells and the anomalous process as the target which can locate in any of the  $M$  cells. The decision maker is allowed to search for the target over  $K$  cells at a time ( $1 \leq K \leq M$ ). We consider the composite hypothesis case, where the observation distribution has an unknown parameter (vector). When taking observations from a certain cell, random continuous values are measured which are drawn from a common distribution  $f$ . The distribution  $f$  has an unknown parameter, belonging to parameter spaces  $\Theta^{(0)}$  or  $\Theta^{(1)}$ , depending on whether the target is absent or present, respectively. The objective is a sequential search strategy that minimizes the expected detection time subject to an error probability constraint. The anomaly detection problem finds applications in intrusion detection in cyber systems for quickly locating infected nodes by detecting statistical anomalies, spectrum scanning in cognitive radio networks for quickly detecting idle channels, and event detection in sensor networks.

### A. Main Results

Dynamic search algorithms can be broadly classified into two classes: (i) Algorithms that use open-loop selection rules, in which the decision of which cell to search is predetermined and independent of the sequence of observations. The stopping rule, that decides when to stop collecting observations from the current cell, and whether to switch to the next cell or stop the test, however, is dynamically updated based on past observations. In this class of algorithms, tractable optimal solutions have been obtained under various settings of observation distributions (see e.g., [2]–[5]). (ii) Algorithms that use closed-loop selection rules, in which the decision of which cell to search is based on past observations. The focus is on addressing the full-blown dynamic problem by jointly optimizing both selection and stopping rules in decision making (see e.g., [6]–[12]). In this setting, however, tractable optimal solutions have been obtained only for very special cases of observation distributions ([6], [7]). In this paper we focus on the latter setting.

Since observations are drawn in a one-at-a-time manner, we are facing a sequential detection problem over multiple composite hypotheses. Sequential detection problems involving multiple processes are partially-observed Markov decision processes (POMDP) [7] which have exponential complexity in general. As a result, computing optimal search policies is intractable (except for some special cases of observation distributions as in [6], [7]). When dealing with composite hypotheses, computing optimal policies is intractable even for the single process case. For tractability, a commonly adopted performance measure is asymptotic optimality in terms of minimizing the detection time as the error probability approaches zero (see, for example, classic and recent results in [8]–[10], [13]–[23]). The focus of this paper is thus on asymptotically optimal strategies with low computational complexity. Our main contributions are three fold, as detailed below.

*a) A general model for composite hypotheses:* Dynamic search problems have been investigated under various models of observation distributions in past and recent years. Closed-loop solutions have been obtained under known Wiener processes [6], known symmetric distributions [7], known general distributions [8], known Poisson point processes with unknown parameters [9], and unknown distributions in which the measurements can take values from a finite alphabet [10]. By contrast to these works, in this paper the dynamic search is conducted over a general known distribution model with unknown parameters that lie in disjoint normal and abnormal parameter sets, and the measurements can take continuous values. This distribution model finds applications in traffic analysis in computer networks [24] and spectrum scanning in cognitive radio networks [25] for instance. Handling this observation model in the dynamic search setting leads to fundamentally different algorithm design and analysis as compared to existing methods.

*b) Algorithm development:* In terms of algorithm development, the proposed algorithm is deterministic and has low-complexity implementations. Specifically, the proposed algorithm consists of exploration and exploitation phases. During exploration, the cells are probed in a round-robin manner for learning the unknown parameters. During exploitation, the most informative observations are collected based on the estimated parameters. We point out that our algorithm uses only bounded exploration time under the setting without side information (Section III-A) and when the null hypothesis is assumed known (Section III-B), which is of

particular significance. It is in sharp contrast with the logarithmic order of exploration time commonly seen in active search strategies (see, for example, [10], [26] or even linear order of exploration time in [9]).

*c) Performance analysis:* In terms of theoretical performance analysis, we prove that the proposed algorithm achieves asymptotic optimality when no additional side information on the process states is known, and a single location is probed at a time (as widely assumed in dynamic search studies for purposes of analysis, e.g., [6], [7], [9], [10], [27]–[29]). Furthermore, when the parameter value under the null hypothesis is known (i.e., as widely applied in anomaly detection cases, and also assumed in [10] for establishing asymptotic optimality), we establish asymptotic optimality as well, with better detection time determined by the true null state. We also consider the case where the parameter value under the null hypothesis is unknown, but is identical for all normal processes. In this case, the proposed algorithm is shown to be consistent in terms of achieving error probability that decays to zero with time. In addition to the asymptotic analysis, an explicit upper bound on the error probability is established under the finite sample regime. Extensive numerical experiments on synthetic dataset and DARPA intrusion detection dataset have been conducted to demonstrate the efficiency of the proposed algorithm.

## *B. Related work*

Optimal solutions for target search or target whereabouts problems have been obtained under some special cases when a single location is probed at a time. Modern application areas of search problems with limited sensing resources include narrowband spectrum scanning [30], [31], event detection by a fusion center that communicates with sensors using narrowband transmission [32], [33], and sensor visual search studied recently by neuroscientists [9]. Results under the sequential setting can be found in [5], [6], [32]–[36]. Specifically, optimal policies were derived in [6], [34], [35] for the problem of quickest search over Wiener processes. In [5], [36], optimal search strategies were established under the constraint that switching to a new process is allowed only when the state of the currently probed process is declared. Optimal policies under general distributions and unconstrained search model remain an open question. In this paper we address this question under the asymptotic regime as the error probability approaches zero. Optimal search strategies when a single location is probed at a time and a fixed sample size

have been established under binary-valued measurements [27]–[29], and under known symmetric distributions of continuous observations [7]. In this paper, however, we focus on the sequential setting and general composite hypothesis case.

Sequential tests for hypothesis testing problems have attracted much attention since Wald’s pioneering work on sequential analysis [37] due to their property of reaching a decision at a much earlier stage than would be possible with fixed-size tests. Wald established the Sequential Probability Ratio Test (SPRT) for a binary hypothesis testing of a single process. Under the simple hypothesis case, the SPRT is optimal in terms of minimizing the expected sample size under given type  $I$  and type  $II$  error probability constraints. Various extensions for  $M$ -ary hypothesis testing and testing composite hypotheses were studied in [14]–[17], [38] for a single process. In these cases, asymptotically optimal performance can be obtained as the error probability approaches zero. In this paper, we focus on asymptotically optimal strategies with low computational complexity for sequential search of a target over multiple processes. Different models considered the case of searching for targets without constraints on the probing capacity, whereas all processes are probed at each given time (i.e.,  $K = M$ , which is a special case of the setting considered in this paper) [17], [22], [23], [35].

Since the decision maker can choose which cells to probe, the anomaly detection problem has a connection with the classical sequential experimental design problem first studied by Chernoff [13]. Compared with the classical sequential hypothesis testing pioneered by Wald [37] where the observation model under each hypothesis is predetermined, the sequential design of experiments has a control aspect that allows the decision maker to choose the experiment to be conducted at each time. Chernoff has established a *randomized* strategy, referred to as the Chernoff test which is asymptotically optimal as the maximum error probability diminishes. Chernoff’s results were proved for a finite number of states of nature, and in [39] Albert extended Chernoff’s results to allow for an infinity of states of nature. More variations and extensions of the problem and the Chernoff test were studied in [8], [18]–[21], [40], [41]. In particular, when the distributions under both normal and abnormal states are completely known under the anomaly detection setting considered here, a modification of the randomized Chernoff test applies and achieves asymptotic optimality [18]. In our previous work [8], we have shown that a simpler deterministic algorithm applies and obtains the same asymptotic performance, with better performance in the finite sample regime. A modified algorithm has been developed recently in [30] for spectrum scanning

with time constraint. In this paper, however, we consider the composite hypothesis case, which is not addressed in [8], [18], [30].

In [9], searching over Poisson point processes with unknown rates has been investigated and asymptotic optimality has been established when a single location is probed at a time. The policy in [9] implements a randomized selection rule and also requires to dedicate a linear order of time for exploring the states of all processes. In our model, however, we consider general distributions (with disjoint parameter spaces) and show that deterministic selection rule, with bounded exploration time achieves asymptotic optimality. This result also extends a recent asymptotic result obtained in [10] for non-parametric detection when distributions are restricted to a finite observation space (in contrast to the general continuous valued observations considered here), where asymptotic optimality was shown when the distribution under the null hypothesis is known, a single location is probed at a time, and a logarithmic order of time is used for exploration. In [26], the problem of detecting abnormal processes over densities that have an unknown parameter was considered, where the process states are independent across cells (in contrast to the problem considered in this paper, in which there is a fixed number of abnormal processes). The objective was to minimize a cost function in the system occurred by abnormal processes, which does not capture the objective of minimizing the detection delay considered here.

Another set of related works is concerned with sequential detection over multiple independent processes [2]–[4], [26], [31], [42]–[45]. In particular, in [2], the problem of identifying the first abnormal sequence among an infinite number of i.i.d. sequences was considered. An optimal cumulative sum (CUSUM) test has been established under this setting. Further studies on this model can be found in [3], [4], [44]. While the objective of finding rare events or a single target considered in [2]–[4], [44] is similar to that of this paper, the main difference is that in [2]–[4], [44] the search is done over an infinite number of i.i.d processes, where the state of each process (normal or abnormal) is independent of other processes, resulting in open-loop search strategies, which is fundamentally different from the setting in this paper.

Other recent studies include searching for a moving Markovian target [46], and searching for correlation structures of Markov networks [47].

Finally, we point out that our setup is different from the change point detection setup. Our model is suitable to cases where a system has already raised an alarm for event (based on change

point detection, for instance), but the location of the event is unknown and needs to be located.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

We consider the problem of detecting a target located in one of  $M$  cells quickly and reliably. An extension to detecting multiple targets is discussed in Sec. III-C. If the target is in cell  $m$ , we say that hypothesis  $H_m$  is true. The *a priori* probability that  $H_m$  is true is denoted by  $\pi_m$ , where  $\sum_{m=1}^M \pi_m = 1$ . To avoid trivial solutions, it is assumed that  $0 < \pi_m < 1$  for all  $m$ .

We focus on the composite hypothesis case, where the observation distribution has an unknown parameter (or a vector of unknown parameters). Let  $\theta_m$  be the unknown parameter that specifies the observation distribution of cell  $m$ . The vector of unknown parameters is denoted by  $\boldsymbol{\theta} = (\theta_1 \dots \theta_M)$ . At each time, only  $K$  ( $1 \leq K \leq M$ ) cells can be observed. When cell  $m$  is observed at time  $n$ , an observation  $y_m(n)$  is drawn independently from a common density  $f(y|\theta_m)$ ,  $\theta_m \in \Theta$ , where  $\Theta \subset \mathbb{R}$  is the parameter space for all cells.

If the target is not located in cell  $m$ , then  $\theta_m \in \Theta^{(0)}$ ; otherwise,  $\theta_m \in (\Theta \setminus \Theta^{(0)})$ . The overall parameter space is the Cartesian product  $\Theta^M$ . Thus, under hypothesis  $H_m$ , the true vector of parameters  $\boldsymbol{\theta} \in \Theta_m \subset \Theta^M$ , where

$$\Theta_m = \{\boldsymbol{\theta} : \theta_i \in \Theta^{(0)}, \forall i \neq m, \theta_m \in \Theta \setminus \Theta^{(0)}\}.$$

Let  $\Theta^{(0)}$ ,  $\Theta^{(1)}$  be disjoint subsets of  $\Theta$ , where  $I = \Theta \setminus (\Theta^{(0)} \cup \Theta^{(1)}) \neq \emptyset$  is an indifference region<sup>1</sup>. When  $\theta^{(1)} \in I$ , the detector is indifferent regarding the location of the target. Hence, there are no constraints on the error probabilities for all  $\theta \in I$ . Shrinking  $I$  increases the sample size. We also assume that  $\Theta^{(0)}$ ,  $\Theta^{(1)}$  are open sets. Let  $\mathbf{P}_m$  be the probability measure under hypothesis  $H_m$  and  $\mathbf{E}_m$  be the operator of expectation with respect to the measure  $\mathbf{P}_m$ .

We define the stopping rule  $\tau$  as the time when the decision maker finalizes the search by declaring the location of the target.<sup>2</sup> Let  $\delta \in \{1, 2, \dots, M\}$  be a decision rule, where  $\delta = m$

<sup>1</sup>The assumption of an indifference region is widely used in the theory of sequential composite hypothesis testing to derive asymptotically optimal performance. Nevertheless, in some cases this assumption can be removed. For more details, the reader is referred to [15].

<sup>2</sup>We point out that it is assumed that the target exists with probability 1. Our model is suitable to cases where a security system has already raised an alarm for event (based on change point detection, for instance), but the location of the event is unknown and need to be located.

if the decision maker declares that  $H_m$  is true. Let  $\phi(n) \in \{1, 2, \dots, M\}^K$  be a selection rule indicating which  $K$  cells are chosen to be observed at time  $n$ . The time series vector of selection rules is denoted by  $\phi = (\phi(n), n = 1, 2, \dots)$ . Let  $\mathbf{y}_{\phi(n)}(n)$  be the vector of observations obtained from cells  $\phi(n)$  at time  $n$  and  $\mathbf{y}(n) = \{\phi(t), \mathbf{y}_{\phi(t)}(t)\}_{t=1}^n$  be the set of all cell selections and observations up to time  $n$ . A deterministic selection rule  $\phi(n)$  at time  $n$  is a mapping from  $\mathbf{y}(n-1)$  to  $\{1, 2, \dots, M\}^K$ . A randomized selection rule  $\phi(n)$  is a mapping from  $\mathbf{y}(n-1)$  to a probability mass function over  $\{1, 2, \dots, M\}^K$ . An admissible strategy  $\Gamma$  for the anomaly detection problem is given by the tuple  $\Gamma = (\tau, \delta, \phi)$ .

We adopt a Bayesian approach as in [13], [15], [37], [48] by assigning a cost of  $c$  for each observation and a loss of 1 for a wrong declaration. Let  $P_e(\Gamma) = \sum_{m=1}^M \pi_m \alpha_m(\Gamma)$  be the probability of error under strategy  $\Gamma$ , where  $\alpha_m(\Gamma) = \mathbf{P}_m(\delta \neq m | \Gamma)$  is the probability of declaring  $\delta \neq m$  when  $H_m$  is true. Let  $\mathbf{E}(\tau | \Gamma) = \sum_{m=1}^M \pi_m \mathbf{E}_m(\tau | \Gamma)$  be the average detection delay under  $\Gamma$ . The Bayes risk under strategy  $\Gamma$  when hypothesis  $H_m$  is true is given by:  $R_m(\Gamma) \triangleq \alpha_m(\Gamma) + c \mathbf{E}_m(\tau | \Gamma)$ . Note that  $c$  represents the ratio of the sampling cost to the cost of wrong detections. The average Bayes risk is given by:

$$R(\Gamma) = \sum_{m=1}^M \pi_m R_m(\Gamma) = P_e(\Gamma) + c \mathbf{E}(\tau | \Gamma).$$

The objective is to find a strategy  $\Gamma$  that minimizes the Bayes risk  $R(\Gamma)$ :

$$\inf_{\Gamma} R(\Gamma), \quad (1)$$

where the infimum is taken over all randomized and deterministic selection rules.

*Definition 1:* Let  $R^*$  be the solution of (1). We say that strategy  $\Gamma$  is asymptotically optimal if

$$\lim_{c \rightarrow 0} \frac{R(\Gamma)}{R^*} = 1. \quad (2)$$

We note that if the strategy that attains  $\inf$  does not exist, the definition of the first order asymptotic optimality would be:

$$\lim_{c \rightarrow 0} \frac{R(\Gamma)}{\inf_{\Gamma} R(\Gamma)} = 1. \quad (3)$$

A shorthand notation  $f \sim g$  will be used to denote  $\lim_{c \rightarrow 0} f/g = 1$ .

A dual formulation (i.e., a frequentist approach) of the problem is to minimize the sample complexity subject to an error constraint  $\alpha$ , i.e.,:

$$\inf_{\Gamma} \mathbb{E}_m(\tau|\Gamma), \quad s.t. \quad P_e(\Gamma) \leq \alpha \quad as \quad \alpha \rightarrow 0 \quad (4)$$

In Section III we develop an asymptotically optimal Deterministic Search (DS) algorithm for solving (1) and (4).

#### A. Notations

We provide next notations that will be used throughout the paper. Let

$$\hat{\theta}_m(n) \triangleq \arg \max_{\theta \in \Theta} f(\bar{\mathbf{y}}_m(n)|\theta) \quad (5)$$

be the maximum likelihood estimate (MLE) of the parameter over the parameter space  $\Theta$  (i.e., unconstrained MLE) at cell  $m$ , where  $\bar{\mathbf{y}}_m(n) = (y_m(r_1), \dots, y_m(r_{k(n)}))$  is the vector of  $k(n)$  observations (indicated by times  $r_1, \dots, r_{k(n)}$ ) collected from cell  $m$  up to time  $n$ . Regularity conditions for consistency of the MLE are given in App. VII-A.1.

Let:

$$\hat{\theta}_m^{(0)}(n) \triangleq \arg \max_{\theta \in \Theta^{(0)}} f(\bar{\mathbf{y}}_m(n)|\theta),$$

$$\hat{\theta}_m^{(1)}(n) \triangleq \arg \max_{\theta \in \Theta \setminus \Theta^{(0)}} f(\bar{\mathbf{y}}_m(n)|\theta)$$

be the MLE for cell  $m$  to be in normal or abnormal state, respectively.

Let  $\mathbf{1}_m(n)$  be the indicator function, where  $\mathbf{1}_m(n) = 1$  if cell  $m$  is observed at time  $n$ , and  $\mathbf{1}_m(n) = 0$  otherwise.

We now propose two optional statistics. Let

$$S_{m,LGLLR}^{(r)}(n) \triangleq \sum_{t=1}^n \mathbf{1}_m(t) \log \frac{f(y_m(t)|\hat{\theta}_m(n))}{f(y_m(t)|\hat{\theta}_m^{(r)}(n))} \quad (6)$$

be the sum of Local Generalized Log-Likelihood Ratio (LGLLR) of cell  $m$  at time  $n$  used to reject hypothesis  $r$  (for  $r = 0, 1$ ) regarding its state. We refer to the statistics as *local* since it uses the observations from cell  $m$  solely. In Section III-C we will define a statistics measure that uses observations from multiple cells, referred to as Multi-process Generalized Log-Likelihood Ratio (MGLLR). The LGLLR statistics is inspired by the Generalized Likelihood Ratio (GLR)

statistics used for sequential tests, first studied by Schwartz [14] for a one parameter exponential family, who assigned a cost of  $c$  for each observation and a loss function for wrong decisions. A refinement was studied by Lai [15], [49], who set a time-varying boundary value. Lai showed that for a multivariate exponential family this scheme asymptotically minimizes both the Bayes risk and the expected sample size subject to error constraints as  $c$  approaches zero [49].

The second statistics that we propose to use is obtained by replacing the parameter for the  $k$ th observation with the estimator  $\hat{\theta}_m(k-1)$  built upon samples  $\tilde{\mathbf{y}}_m(n) = (y_m(r_1), \dots, y_m(r_{k-1(n)}))$ . The statistics is given by:

$$S_{m,LALLR}^{(r)}(n) \triangleq \sum_{t=1}^n \mathbf{1}_m(t) \log \frac{f(y_m(t)|\hat{\theta}_m(t-1))}{f(y_m(t)|\hat{\theta}_m^{(r)}(n))}, \quad (7)$$

which we refer to as the sum of Local Adaptive Log Likelihood Ratio (LALLR). The LALLR statistics is inspired by the Adaptive Likelihood Ratio (ALR) statistics used for sequential tests, first introduced by Robbins and Siegmund [50] to design power-one sequential tests. Pavlov used it to design asymptotically (as the error probability approaches zero) optimal (in terms of minimizing the expected sample size subject to error constraints) tests for composite hypothesis testing of the multivariate exponential family [16]. Tartakovsky established asymptotically optimal performance for a more general multivariate family of distributions [17].

The advantage of using the LALLR statistics, is that it enables us to upper-bound the error probabilities of the sequential test by using simple threshold settings. Thus, implementing the LALLR is much simpler than implementing the LGLLR. The disadvantage of using the LALLR is that poor early estimates (for small number of observations) can never be revised even though one has a large number of observations. A numerical comparison for the performance of the two statistics is presented in Section IV-B.

Finally,

$$D(x||z) \triangleq \mathbf{E}_{f(y(n)|x)} \left( \log \frac{f(y(n)|x)}{f(y(n)|z)} \right)$$

denotes the KullbackLeibler (KL) divergence between two distributions,  $f(y(n)|x)$ ,  $f(y(n)|z)$ .

### III. A LOW-COMPLEXITY DETERMINISTIC SEARCH (DS) ALGORITHM

Sequential detection problems involving multiple processes are POMDP [7]. As a result, computing optimal search policies is intractable in general. In this section we present the

Deterministic Search (DS) algorithm, which has low complexity (linear with the number of processes) used for solving the anomaly detection problem asymptotically as the error approaches zero. Both proposed statistics (LGLLR and LALLR) can be used in the implementation of the algorithm.

We start by analyzing the case where no additional side information on the process states is known in Section III-A. Then, in Section III-B, we consider the case in which the parameter value under the null hypothesis is known and equal for all normal processes. In this case we show analytically the gain achieved in the detection time, by utilizing the side information on the normal state. Finally, in Section III-C, we consider the case where the parameter value under the null hypothesis is unknown, but is known to be equal for all normal processes.

#### A. Anomaly Detection Without Side Information

We assume that  $K = 1$  as widely assumed in dynamic search problems for purposes of analysis (e.g., [5], [6], [9], [10], [34]–[36]). In Section III-C we discuss the implementation under more general settings. We also assume that the parameter space is finite, and we assume a large-scale system where  $M \gg 1$  so that  $D(\theta^{(0)}||\theta^{(1)})/D(\theta^{(1)}||\theta^{(0)}) < M-1$  for all  $\theta^{(0)} \in \Theta^{(0)}$ ,  $\theta^{(1)} \in \Theta^{(1)}$ . Let  $\mathcal{H}_1(n) = \{m : \hat{\theta}_m(n) \notin \Theta^{(0)}\}$  be the set of cells whose MLEs lie outside  $\Theta^{(0)}$  at time  $n$  with cardinality  $|\mathcal{H}_1(n)| = N_{H_1}(n)$ . Let  $S_m^{(r)}(n)$  be the  $S_{m,LALLR}^{(r)}(n)$  or  $S_{m,LGLLR}^{(r)}(n)$  statistics defined in Section II-A. The DS algorithm has a structure of exploration and exploitation epochs. We start by addressing the Bayesian formulation, and we describe the DS algorithm with respect to time index  $n$ .

- 1) (Exploration phase:) If  $N_{H_1}(n) \neq 1$ , then probe the cells one by one in a round-robin manner, i.e.,  $\phi(n) = [(\phi(n-1) + 1) \bmod M]$  and go to Step 1 again. Otherwise, go to Step 2.
- 2) (Exploitation phase:) Update  $\hat{\theta}_m(n)$  for all  $m = 1, \dots, M$ , and let  $\hat{m}(n) = \{m : \hat{\theta}_m(n) \notin \Theta^{(0)}\}$  be the index of the cell whose MLE lies outside  $\Theta^{(0)}$  at time  $n$  (note that this cell is unique at the exploitation phase). Probe cell  $\phi(n) = \hat{m}(n)$  and go to Step 3.
- 3) (Sequential testing:) Update  $S_{\phi(n)}^{(0)}(n)$  based on the last observation. If  $S_{\hat{m}(n)}^{(0)}(n) \geq -\log c$  stop the test and declare  $\delta = \hat{m}(\tau)$  as the location of the target. Otherwise, go to Step 1.

Note that the selection rule constructed by Steps 1, 2 is deterministic and dynamically updated based on the current value of the MLEs. The proposed DS algorithm is intuitively satisfying. Consider first the simple hypothesis case (where asymptotic optimality was shown in [8]), in which  $\theta^{(0)}, \theta^{(1)}$  are assumed known. When  $K = 1$  and  $M \gg 1$ , the DS algorithm selects at each time the cell with the largest sum log likelihood ratio. The intuition behind this selection rule is that  $D(\theta^{(1)}||\theta^{(0)})$  and  $D(\theta^{(0)}||\theta^{(1)})/(M - 1)$  determine, respectively, the rates at which the state of the cell with the target and the states of the  $M - 1$  cells without the target can be accurately inferred. Since  $M \gg 1$  such that  $D(\theta^{(0)}||\theta^{(1)})/D(\theta^{(1)}||\theta^{(0)}) < M - 1$  for all  $\theta^{(0)} \in \Theta^{(0)}, \theta^{(1)} \in \Theta^{(1)}$ , the DS algorithm aims at identifying the cell with the target (which is equivalent to probe the most likely abnormal process as implemented during the exploitation phase). When handling the composite hypothesis case and  $\theta^{(1)}$  is unknown, the selection rule dedicates an exploration phase for estimating the parameter and adjusts the estimated KL divergences dynamically. Since the parameter spaces are disjoint, the exploration phase yields an estimate for the location of the abnormal process (i.e., the cell whose MLE lies outside  $\Theta^{(0)}$ ). The exploitation phase keeps taking samples until  $S_{\hat{m}(n)}^{(0)}(n) \geq -\log c$  first occurs to ensure a sufficiently accurate decision, i.e., error probability of order  $O(c)$  as shown in the analysis.

*Theorem 1:* Assume that the DS algorithm is implemented under the anomaly detection setting described in this section. Let  $R^*$  and  $R(\Gamma)$  be the Bayes risks under the DS algorithm and any other policy  $\Gamma$ , respectively. Then, the following statements hold:

1) (*Finite sample error bound:*) The error probability is upper bounded by  $(M - 1)c$  for all  $c$ .

2) (*Asymptotic optimality:*) The Bayes risk satisfies:

$$R^* \sim \frac{-c \log c}{D(\theta^{(1)})} \sim \inf_{\Gamma} R(\Gamma) \quad \text{as } c \rightarrow 0,$$

where  $D(\theta^{(1)}) \triangleq \min_{\varphi \in \Theta^{(0)}} D(\theta^{(1)}||\varphi)$ .

3) (*Bounded exploration time:*) The total expected time spent during the exploration phase (i.e., Step 1 in the DS algorithm) is  $O(1)$ .

The proof is given in Appendix VII-C.

We point out that bounded exploration time of the DS algorithm is of particular significance. It is in sharp contrast with the logarithmic order of exploration time commonly seen in active search strategies (see, for example, [10], [26]).

### B. Anomaly Detection under a Known Model of Normality

Here, we assume that the parameter under null hypothesis  $\theta = \theta^{(0)} \in \Theta^{(0)}$  is known, and equal for all empty cells, where  $\Theta^{(0)}$  is an open set that contains  $\theta^{(0)}$ . This setting models many anomaly detection situations, in which the distribution of the observations under a normal state is known, while there is uncertainty in the distribution under an abnormal state. To utilize this information, we adjust the LALLR statistics used to reject hypothesis  $H_0$  as follows:

$$\tilde{S}_{m,LALLR}^{(0)}(n) \triangleq \sum_{t=1}^n \mathbf{1}_m(t) \log \frac{f(y_m(t)|\hat{\theta}_m(t-1))}{f(y_m(t)|\theta^{(0)})}. \quad (8)$$

We define  $\tilde{S}_{m,LGLLR}^{(0)}(n)$  similarly.

In the following theorem we establish a finite-sample upper bound on the error probability and prove asymptotic optimality of the algorithm for the Bayesian formulation using the adjusted LALLR statistics, where only  $O(1)$  order of time is spent during the exploration phase. The proof is given in App. VII-A.

*Theorem 2:* Assume that the DS algorithm is implemented under the anomaly detection setting described in this section, using the adjusted LALLR statistics. Let  $R^*$  and  $R(\Gamma)$  be the Bayes risks under the DS algorithm and any other policy  $\Gamma$ , respectively. Then, the following statements hold:

- 1) (*Finite sample error bound:*) The error probability is upper bounded by  $(M-1)c$  for all  $c$ .
- 2) (*Asymptotic optimality:*) The Bayes risk satisfies:
 
$$R^* \sim \frac{-c \log c}{D(\theta^{(1)}||\theta^{(0)})} \sim \inf_{\Gamma} R(\Gamma) \quad \text{as } c \rightarrow 0.$$
- 3) (*Bounded exploration time:*) The total expected time spent during the exploration phase (i.e., Step 1 in the DS algorithm) is  $O(1)$ .

We point out that the side information on the true null hypothesis strengthens the algorithm performance. The improvement in the performance is clearly seen by the fact that  $D(\theta^{(1)}||\theta^{(0)}) \geq D(\theta^{(1)})$ . Hence, the risk in Theorem 2 is smaller than the risk in Theorem 1. Note also that in this setting we do not restrict  $\Theta^{(0)}$  to be a singleton set (the parameter still lies in an open set). The side information is utilized when constructing the statistics in (8).

For the frequentist formulation, in step 3 of the DS algorithm (i.e., sequential testing step) we define the threshold as  $a$ , i.e., if  $S_{\hat{m}(n)}^{(0)}(n) \geq a$  we stop the test and declare  $\delta = \hat{m}(\tau)$  as the

location of the target. We now present Theorem 3, which claims that the DS algorithm is first order asymptotically optimal in the sense of criterion (4). The proof is given in App.VII-B.

*Theorem 3:* Assume that the DS algorithm is implemented under the anomaly detection setting described in this section, using the adjusted LALLR statistics. We define the class of tests :

$$C(\alpha) = \{\Gamma : P_e(\Gamma) \leq \alpha\}.$$

Let  $\mathbb{E}_m(\tau|\Gamma^*)$  and  $\mathbb{E}_m(\tau|\Gamma)$  be, the detection time under the DS algorithm, and any other policy, respectively. Then, the following statement holds for each  $m = 1, \dots, M$ :

$$\mathbb{E}_m(\tau|\Gamma^*) \sim \inf_{\Gamma \in C(\alpha)} \mathbb{E}_m(\tau|\Gamma) = \frac{|\log \alpha|}{D(\theta^{(1)}||\theta^{(0)})}(1 + o(1)), \quad (9)$$

*as*  $\alpha \rightarrow 0$ ,

and  $\Gamma^* \in C(\alpha)$ .

### C. Anomaly Detection under Identical Parameter for All Normal Cells

Next, we consider the case where both parameter values under normal and abnormal states  $\theta^{(0)}$  and  $\theta^{(1)}$  are unknown. However, it is known that the unknown parameter is identical for all normal cells. Therefore, under hypothesis  $H_m$ , the true vector of parameters satisfy  $\boldsymbol{\theta} \in \Theta_m \subset \Theta^M$ , where

$$\Theta_m = \{\boldsymbol{\theta} : \theta_i = \theta^{(0)} \in \Theta^{(0)}, \forall i \neq m, \theta_m = \theta^{(1)} \in \Theta \setminus \Theta^{(0)}\}.$$

Note that in contrast to section III-A where observations from cell  $m$  does not contribute any information about the parameter value of cell  $r$ , for  $m \neq r$ , here the additional side information allows us to estimate the true value of  $\theta^{(0)}$  consistently using observations from each normal cell. Specifically, let  $\mathbf{y}_{\Theta^{(0)}}(n), \mathbf{y}_{\Theta \setminus \Theta^{(0)}}(n)$  be the set of all the observations collected from the cells whose MLEs lie inside  $\Theta^{(0)}$  (i.e.,  $\hat{\theta}_m(n) \in \Theta^{(0)}$ ), and inside  $\Theta \setminus \Theta^{(0)}$  (i.e.,  $\hat{\theta}_m(n) \in \Theta \setminus \Theta^{(0)}$ ) at time  $n$ , respectively. The global MLE of  $\theta^{(0)}$  is computed based on the observations from all the cells which are likely to be empty:

$$\hat{\theta}^{(0)}(n) \triangleq \arg \max_{\theta \in \Theta} f(\mathbf{y}_{\Theta^{(0)}}(n)|\theta),$$

where the global MLE of  $\theta^{(1)}$  is computed based on the observations from all the cells which are likely to contain the target:

$$\hat{\theta}^{(1)}(n) \triangleq \arg \max_{\theta \in \Theta} f(\mathbf{y}_{\Theta \setminus \Theta^{(0)}}(n) | \theta).$$

Intuitively, as more observations are collected from all cells, only the MLE at the cell that contains the target is likely to lie inside  $\Theta \setminus \Theta^{(0)}$ . Next, we define the statistics accordingly. Let:

$$S_{m,MGLLR}^{(r)}(n) \triangleq \sum_{t=1}^n \mathbf{1}_m(t) \log \frac{f(y_m(t) | \hat{\theta}_m(n))}{f(y_m(t) | \hat{\theta}^{(r)}(n))} \quad (10)$$

be the sum of Multi-process Generalized Log-Likelihood ratio (MGLLR) of cell  $m$  at time  $n$  used to reject hypothesis  $r$  (for  $r = 0, 1$ ) regarding its state. The modified adaptive statistics is defined by:

$$S_{m,MALLR}^{(r)}(n) \triangleq \sum_{t=1}^n \mathbf{1}_m(t) \log \frac{f(y_m(t) | \hat{\theta}_m(t-1))}{f(y_m(t) | \hat{\theta}^{(r)}(n))}, \quad (11)$$

which we refer to as the sum of Multi-process Adaptive Log-Likelihood Ratio (MALLR)<sup>3</sup>.

Let  $\mathcal{N}_s = \{n_1, n_2, \dots\}$  be a sequence of time instants, where  $O(|\mathcal{N}_s|)$  has a logarithmic order of time, in which the cells are selected in a round-robin manner during the algorithm. Intuitively speaking, the role of  $n_1, n_2, \dots$ , is to explore all the cells to infer the true value of  $\theta^{(0)}$  (which is not observed when testing the target cell) during the algorithm. This allows us to use the estimate values of both  $\theta^{(0)}$  and  $\theta^{(1)}$  when computing the statistics used in the algorithm. We also define  $m^{(i)}(n)$  for  $i = 1, 2, \dots, M-1$  as the index of the cell with the  $i^{th}$  smallest sum MALLR  $S_j^{(1)}(n)$  for  $j \neq \hat{m}(n)$  at time  $n$ . The DS algorithm has a structure of exploration and exploitation epochs. Let  $S_m^{(r)}(n)$  be the statistics used in the algorithm which can be the MALLR or MGLLR statistics. Next, we describe the DS algorithm with respect to time index  $n$ . We describe the algorithm for the general case where multiple processes can be probed at a time ( $K \geq 1$ ), and  $D(\theta^{(0)} || \theta^{(1)}) / D(\theta^{(1)} || \theta^{(0)}) < M-1$  does not necessarily hold.

- 1) (Exploration phase 1:) Exploration phase 1 is similar to the exploration phase described in Section III-B. If  $N_{H_1}(n) \neq 1$ , then cells are probed one by one in a round-robin manner. Otherwise, go to Step 2.

<sup>3</sup>Note that the adaptive LLR statistics and generalized LLR statistics used in sequential composite hypothesis testing of a single process contains a constrained MLE over the alternative parameter space in the denominator (see Section IV-A for more details). Here, we use unconstrained MLEs (which are computed over the entire parameter space  $\Theta$ ) in both numerator and denominator, depending on the cells from which the observations were taken. Thus, we refer to this statistics measure as a Multi-process Adaptive/Generalized LLR (MALLR/ MGLLR).

- 2) (Exploration phase 2:) If  $n \in \mathcal{N}_s$ , the cells are probed one by one in a round-robin manner. Otherwise, if  $N_{H_1}(n) \neq 1$ , go to Step 1. Otherwise, go to Step 3.
- 3) (Exploitation phase:) Update  $\hat{\theta}_m(n)$  for all  $m = 1, \dots, M$ , and let  $\hat{m}(n) = \left\{ m : \hat{\theta}_m(n) \notin \Theta^{(0)} \right\}$  be the index of the cell whose MLE lies outside  $\Theta^{(0)}$  at time  $n$  (note that this cell is unique at the exploitation phase). Then, probe cells  $\phi(n)$  which are given by:<sup>4</sup>

$$\phi(n) = \begin{cases} \left( \hat{m}(n), m^{(1)}(n), m^{(2)}(n), \dots, m^{(K-1)}(n) \right), & \text{if } D(\hat{\theta}^{(1)}(n) || \hat{\theta}^{(0)}(n)) \geq \frac{D(\hat{\theta}^{(0)}(n) || \hat{\theta}^{(1)}(n))}{(M-1)} \\ & \text{and } n \notin \{\text{exploration times}\} \\ \left( m^{(1)}(n), m^{(2)}(n), \dots, m^{(K)}(n) \right), & \text{if } D(\hat{\theta}^{(1)}(n) || \hat{\theta}^{(0)}(n)) < \frac{D(\hat{\theta}^{(0)}(n) || \hat{\theta}^{(1)}(n))}{(M-1)} \\ & \text{and } n \notin \{\text{exploration times}\} \end{cases} \quad (12)$$

and go to Step 4.

- 4) (Sequential testing:) Update the sum MALLRs based on the last observations. If  $S_{\hat{m}(n)}^{(0)}(n) + S_{m^{(1)}(n)}^{(1)}(n) \geq -\log c$  stop the test and declare  $\delta = \hat{m}(\tau)$  as the location of the target. Otherwise, go to Step 1.

The proposed DS algorithm under the general setting is intuitively satisfying. Since both  $\theta^{(0)}, \theta^{(1)}$  might be unknown, the selection rule dedicates exploration phases 1, 2 for estimating the parameters and adjusts the estimated KL divergences dynamically. Since the parameter spaces are disjoint, exploration phase 1 yields an estimate for the location of the abnormal process (i.e., the cell whose MLE lies outside  $\Theta^{(0)}$ ). The exploitation phase keeps taking samples until  $S_{\hat{m}(n)}^{(0)}(n) + S_{m^{(1)}(n)}^{(1)}(n) \geq -\log c$  first occurs, i.e., to ensure a sufficiently accurate decision. We show in the appendix that this stopping rule achieves error probability of order  $O(c)$  when the parameters are known under both normal and abnormal states, and polynomial decay with time is achieved under the general composite hypothesis testing setting (though only consistency can be shown, where asymptotic optimality still remains open in the general setting), which motivates the design of the stopping rule.

<sup>4</sup>Assume that  $K < M$ . Otherwise, all cells are probed.

In the theorem below, we prove the consistency of the DS algorithm using the MALLR statistics. The proof and regularity assumptions are given in Appendix VII-D.

*Theorem 4:* Assume that the DS algorithm is implemented under the anomaly detection setting described in this section. Assume also that the parameters  $\theta^{(0)}, \theta^{(1)}$  can take a finite number of values (where the observations are still continuous). Let  $H_m$  be true hypothesis. Then,  $\hat{m}(n) \rightarrow m$  as  $c \rightarrow 0$ , and the error probability decays polynomially with  $-\log c$ .

It should be noted that the expected detection time is of order  $O(-\log c)$ . Therefore, Theorem 4 implies that the error probability decays polynomially with the expected detection time. We point out that establishing asymptotic optimality for  $K > 1$  remains open. In this case, at each time slot the statistics is based on a mixed of samples from cells that contain the target and from cells that do not contain the target. As a result, bounding the error probability by  $O(c)$  while achieving the asymptotically optimal detection time is much more complex.

In Figure 1 we present simulation results, demonstrating strong performance of the DS algorithm under the setting considered in this section. The sum MALLRs use the exact values of  $\theta^{(0)}, \theta^{(1)}$  when they are known, and the MLEs of  $\theta^{(0)}, \theta^{(1)}$  when they are unknown. Although theoretical asymptotic optimality remains open when  $\theta^{(0)}, \theta^{(1)}$  are unknown (and  $\theta^{(0)}$  is identical for all normal cells), it can be seen by simulations that the DS algorithm nearly achieves asymptotically optimal performance in this case as well (since it approaches the performance of the DS algorithm when  $\theta^{(0)}, \theta^{(1)}$  are known).

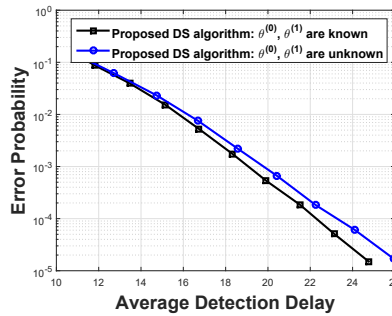


Fig. 1. The error probability as a function of the average detection delay under the proposed DS algorithm. A case of Laplace distributions with parameters  $\theta^{(0)} = 0, \theta^{(1)} = 1$  under normal and abnormal states, respectively, with  $K = 2, M = 5$ . We averaged over  $10^6$  Monte Carlo runs.

*Remark 1:* Note that in Sections III-A and III-B the exploitation phase collects observations from cell  $\hat{m}(n)$ . As a result, a sufficiently accurate MLE for  $\theta^{(1)}$  is computed based on observations collected during the exploitation phase, while exploration phase 2 is unnecessary. In the setting considered in this section, however, exploration phase 2 is required to guarantee a sufficiently accurate estimation of the unknown parameter  $\theta^{(0)}$ . Specifically, let  $N_O(t)$  denote the number of observations that have been collected in exploration phase 2, and let  $\tau_{ML}^{(0)}$  be the smallest integer such that  $\hat{\theta}^{(0)}(n) = \theta^{(0)}$  for all  $n > \tau_{ML}^{(0)}$ . Then, exploring cells such that  $N_O(t) > \frac{2}{I_0} \log(t)$  is met for all  $t$  is sufficient to ensure consistency of  $\hat{\theta}^{(0)}(n)$ , where

$$I_0 \triangleq \inf_{\theta^{(0)}, \theta \in \Theta: \theta^{(0)} \neq \theta} \sup_{s > 0} \left\{ -\log \mathbf{E}_{\sim f(y|\theta^{(0)})} \left[ e^{s(-\ell^{(\theta^{(0)}, \theta)})} \right] \right\} \quad (13)$$

is the Legendre-Fenchel transformation of

$$\ell^{(\theta^{(0)}, \theta)} = \log \frac{f(y|\theta^{(0)})}{f(y|\theta)}.$$

Below, we prove the statement (under hypothesis  $H_m$  w.l.o.g.):

$$\mathbf{P}_m \left( \tau_{ML}^{(0)} > n \right) \leq \sum_{t=n}^{\infty} \mathbf{P}_m \left( \hat{\theta}^{(0)}(t) \neq \theta^{(0)} \right).$$

By the definition of  $\hat{\theta}^{(0)}(n)$ , the event  $\hat{\theta}^{(0)}(t) \neq \theta^{(0)}$  implies:

$$\sum_{i=1}^t \ell^{(\theta^{(0)}, \tilde{\theta}^{(t)})}(i) < 0, \quad (14)$$

for some  $\tilde{\theta}^{(t)} \neq \theta^{(0)}$ , where the index  $i$  refers to measurement  $i$  taken from cells which are likely to be empty. Since the expected last exit time (say  $t'$ ) from exploration phase 1 is bounded (see Appendix VII-A), applying the Chernoff bound for all  $t > t'$  and using the i.i.d property yields:

$$\begin{aligned} & \mathbf{P}_m \left( \sum_{i=1}^t \ell^{(\theta^{(0)}, \tilde{\theta}^{(t)})}(i) < 0 \right) \\ & \leq \min_{s > 0} \left\{ \mathbf{E}_{\sim f(y|\theta^{(0)})} \left[ e^{s(-\ell^{(\theta^{(0)}, \tilde{\theta}^{(t)})}(i))} \right] \right\}^{N_O(t)} \\ & = \min_{s > 0} \left\{ e^{-N_O(t) \left( -\log \mathbf{E}_{\sim f(y|\theta^{(0)})} \left[ e^{s(-\ell^{(\theta^{(0)}, \tilde{\theta}^{(t)})}(i))} \right] \right)} \right\} \\ & = e^{-N_O(t) \left( \sup_{s > 0} \left\{ -\log \mathbf{E}_{\sim f(y|\theta^{(0)})} \left[ e^{s(-\ell^{(\theta^{(0)}, \tilde{\theta}^{(t)})}(i))} \right] \right\} \right)}. \end{aligned} \quad (15)$$

Since  $N_O(t) > \frac{2}{I_0} \log(t)$ ,  $\mathbf{E}_{\sim f(y|\theta^{(0)})}[\tau_{ML}^{(0)}] = O(1)$  is satisfied.

*Remark 2:* It should be noted that the proposed DS algorithm can be extended to handle multiple (say  $L$ ) abnormal processes as well. The exploration phase can be implemented in a similar manner until exactly  $L$  MLEs lie outside  $\Theta^{(0)}$ . The exploitation phase will prioritize processes which are likely to be abnormal if the conditions on the first line of (12) hold. Otherwise, it will prioritize processes which are likely to be normal if the conditions on the second line of (12) hold. The test terminates once all the abnormal processes are distinguished from the rest  $M - L$  normal processes, i.e., when the  $L^{th}$  highest sum MALLR among the processes which are likely to be abnormal plus the smallest sum MALLR among the processes which are likely to be normal is greater than  $-\log c$ .

#### D. Comparison with Chernoff's test

In this section, we discuss the differences between our problem and the classical sequential experimental design problem studied by Chernoff, first presented in [13]. While we presented a deterministic algorithm search, Chernoff proposed a test with a randomized selection rule. Specifically, let  $q = (q_1, \dots, q_N)$  be a probability mass function over a set of  $N$  available experiments  $u = \{u_i\}_{i=1}^N$  that the decision maker can choose from, where  $q_i$  is the probability of choosing experiment  $u_i$ . For a general M-ary sequential design of experiments problem, the action at time  $n$  under the Chernoff test is drawn from a distribution  $q^*(n) = (q_1^*(n), \dots, q_N^*(n))$  that depends on the past actions and observations:

$$q^*(n) = \arg \max_q \min_{j \in \mathcal{M} \setminus \{\hat{i}(n)\}} \sum_{u_i} q_i D(p_{\hat{i}(n)}^{u_i} || p_j^{u_i}), \quad (16)$$

where  $\mathcal{M}$  is the set of the  $M$  hypotheses,  $\hat{i}(n)$  is the MLE of the true hypothesis at time  $n$  based on past actions and observations, and  $p_j^{u_i}$  is the observation distribution under hypothesis  $j$  when action  $u_i$  is taken.

Chernoff's results were proved only for a finite number of states of nature (set of possible parameters). Albert [39] extended Chernoff's results to allow for an infinity of states of nature. Beyond the differences in the deterministic versus randomized selection rules, we will now discuss in details the connection with the model considered by Chernoff and Albert. (i) *Violating the positivity assumption on the KL divergence:* The asymptotic optimality of the Chernoff test as shown in [13], [39] requires that under any experiment, any pair of hypotheses are distinguishable (i.e., has positive KL divergence). This assumption does not hold in the anomaly

detection settings considered in this paper. For instance, under the experiment of searching the  $i^{th}$  cell, the hypotheses of the target being in the  $j^{th}$  ( $j \neq i$ ) and the  $k^{th}$  ( $k \neq i$ ) cells yield the same observation distribution. In [18], the authors relaxed this assumption, and developed a modified Chernoff test in order to handle indistinguishable hypotheses under some (but not all) actions. The basic idea of the modified test is to implement an exploration phase with a uniform distribution for a subsequence of time instants that grows logarithmically with time. Although asymptotic optimality was proved under the modified Chernoff test, its exploration time is unbounded, and affects the finite-time performance. Nevertheless, in this paper we have shown that the DS algorithm achieves asymptotic optimality under both settings in Sections III-A, III-B, using a bounded exploration time. (ii) *Utilizing the side information in the anomaly detection setting*: The model in [13], [39] can be embedded to the model in Section III-A (with the extension in [18] as discussed earlier). This embedding does not contain side information on the parameter values under different hypotheses. The analysis in [13], [39] relies on rejecting the alternative hypothesis with respect to the closest alternative. Indeed, the DS algorithm achieves the same asymptotic optimality as in [13], but with deterministic selection rule, with better finite-time performance as demonstrated in the simulation results. The asymptotically optimal Bayes risk is given in this case by  $\sim -c \log c / \inf_{\varphi \in \Theta^{(0)}} D(\theta^{(1)} || \varphi)$  which matches with the asymptotically optimal performance in [13], [39]. Asymptotic optimality of the Chernoff test is achieved under the model setting in Section III-B by embedding the parameter set under  $\theta^{(0)}$  to a singleton, and thus the same asymptotic performance can be achieved, where the asymptotically optimal Bayes risk is given in this case by  $\sim -c \log c / D(\theta^{(1)} || \theta^{(0)})$ . Indeed, we have shown that the DS algorithm achieves the same asymptotic optimality as in [13], [39] in this case. However, asymptotic optimality under the Chernoff test remains open in the setting considered in Section III-C, since it cannot be embedded as in Section III-B. The asymptotic analysis in [13], [39] is established with respect to the entire parameter space (as in Section III-A), while the lower bound on the risk must be developed with respect to the true parameter values that satisfy the side information. Nevertheless, intuitively, one can expect to improve performance by estimating the parameter  $\theta^{(0)}$  consistently and improve the detection performance by approaching the performance in Section III-B. We indeed showed that the DS algorithm achieves consistency in this setting.

Despite the differences between the two models, we extended the randomized Chernoff test

for the anomaly detection problem over composite hypotheses as follows. We select cells from a uniform distribution at exploration phase until only a single MLE lies outside  $\Theta^{(0)}$ . Then, the solution of (16) is executed in the exploitation phase. The randomized test in (16) chooses, at each time, a probability distribution that governs the selection of the experiment to be carried out at this time. This distribution is obtained by solving a maximin problem so that the next observation will best differentiate the current MLE of the true hypothesis from its closest alternative, where the distance is measured by the KL divergence. It can be shown that when applied to the anomaly detection problem, the solution of (16) works as follows. Consider for example the setting in Section III-B (i.e., when the parameter under the null hypothesis is known). When  $D(\hat{\theta}^{(1)}(n)||\theta^{(0)}) \geq \frac{D(\theta^{(0)}||\hat{\theta}^{(1)}(n))}{(M-1)}$ , the Chernoff test selects cell  $\hat{m}(n)$  and draws the rest  $K - 1$  cells randomly with equal probability from the remaining  $M - 1$  cells. When  $D(\hat{\theta}^{(1)}(n)||\theta^{(0)}) < \frac{D(\theta^{(0)}||\hat{\theta}^{(1)}(n))}{(M-1)}$ , all  $K$  cells are drawn randomly with equal probability from cells  $\{m^{(1)}(n), m^{(2)}(n), \dots, m^{(M-1)}(n)\}$ . The same selection rule is obtained when setting the alternative hypothesis according to the settings in Sections III-A, III-C. We refer to this policy as the modified Chernoff test. We present numerical examples to illustrate the performance of the proposed deterministic policy as compared to the randomized Chernoff test, under the setting considered in Section III-C. It can be seen in Figures 2, 3, that the proposed deterministic DS algorithm significantly outperforms the randomized Chernoff test.

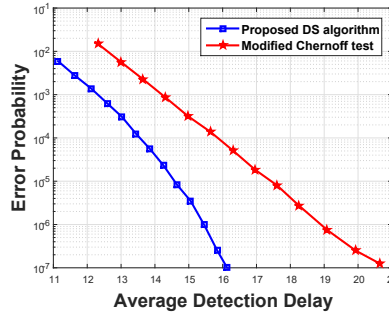


Fig. 2. The error probability as a function of the average detection delay under various algorithms: (i) The proposed DS algorithm that uses the MALLR statistics (referred to as the proposed DS algorithm); and (ii) The modified randomized Chernoff test as described in Section III-D. A case of exponential distributions with parameters  $\theta^{(0)} = 1, \theta^{(1)} = 10$  under normal and abnormal states, respectively, where  $M = 15$  and  $K = 5$ . We averaged over  $4 \cdot 10^7$  Monte Carlo runs.

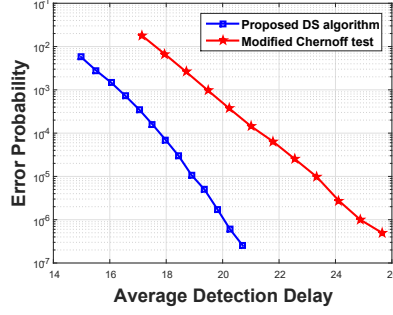


Fig. 3. The error probability as a function of the average detection delay under various algorithms: (i) The proposed DS algorithm that uses the MALLR statistics (referred to as the proposed DS algorithm); and (ii) The modified randomized Chernoff test as described in Section III-D. A case of exponential distributions with parameters  $\theta^{(0)} = 1, \theta^{(1)} = 10$  under normal and abnormal states, respectively, where  $M = 20$  and  $K = 5$ . We averaged over  $4 \cdot 10^7$  Monte Carlo runs.

#### IV. EMPIRICAL STUDIES

In this section, we present additional numerical experiments<sup>5</sup> for demonstrating the performance of the proposed DS algorithm as compared to existing methods.

##### A. Comparison between MALLR and LALLR statistics

We first compare the proposed DS algorithm under the settings of Section III-C, using the MALLR statistics defined in (11) and the LALLR statistics defined in (7), which is a popular method for performing sequential composite hypothesis testing, first introduced by Robbins and Siegmund in [51] (variations can be found in [16], [17], [52]). It can be seen that the proposed DS algorithm using the MALLR statistics adopts a variation of the LALLR statistics in the design of the stopping rule for anomaly detection over multiple composite hypotheses. However, since both empty cells and the cell that contains the target can be observed by the decision maker, the unconstrained MLEs of the unknown parameters  $\theta^{(0)}$  and  $\theta^{(1)}$  can be applied in both numerator and denominator (which we referred to as MALLR). We next simulate the case of searching for a target over processes that follow Laplace distributions with unknown means, where the observations  $y_j$  are drawn from distribution  $f(y_j|\theta) = \frac{1}{2} \exp\{|y - \theta|\}$ . We note that by using the global MLE we expect for better performance. The simulation results demonstrate

<sup>5</sup>The indifference region in the simulations was set to  $I = [\frac{\theta^{(0)} + \theta^{(1)}}{2} - 10^{-3}, \frac{\theta^{(0)} + \theta^{(1)}}{2} + 10^{-3}]$ . We ran Monte-Carlo experiments for generating the simulation results.

the performance gain that we can get in this setting. It can be seen in Figure 4 that implementing the DS algorithm with MALLR statistics as proposed in Section III-C significantly outperforms an algorithm that uses the selection rule of DS algorithm with the LALLR statistics as proposed in Section III-A. It can be seen that the error exponent is significantly better when using the MALLR statistics in the algorithm design. Thus, the performance gain by using the proposed DS algorithm is expected to further increase as the error decreases.

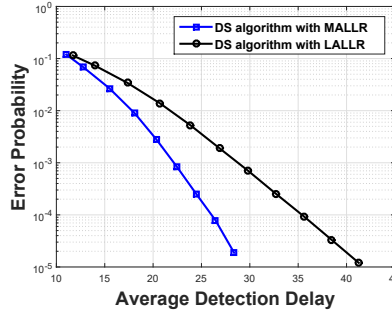


Fig. 4. The error probability as a function of the average detection delay. Performance comparison between the following algorithms: (i) The proposed DS algorithm that uses the MALLR statistics as described in Section III-C (referred to as the DS selection rule with MALLR); and (ii) The proposed DS algorithm that uses the LALLR statistics as described in Section III-A. A case of Laplace distributions with parameters  $\theta^{(0)} = 0$ ,  $\theta^{(1)} = 1$ , under normal and abnormal states, respectively, with  $K = 2$ ,  $M = 5$ . We averaged over  $10^6$  Monte Carlo runs.

### B. Comparison between MALLR and MGLLR

In Figure 5, we compare the performance of the two proposed statistics suggested in section III-C. As discussed earlier, using the MALLR statistics allows us to establish asymptotic optimality theoretically, whereas asymptotic optimality remains open when using the MGLLR. However, in practice, we expect that using the MGLLR will perform better since it uses all samples when updating the MLE. It can be seen in Figure 5 that the DS algorithm using the MGLLR statistics slightly outperforms the DS algorithm using the MALLR.

### C. Network Traffic Analysis

Finally, we demonstrate the performance of the DS algorithm using the MALLR statistics in intrusion detection applications, by detecting statistical deviations in network traffic. We examine

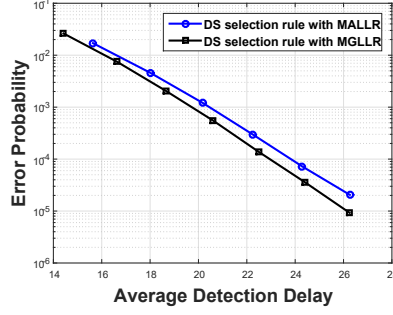


Fig. 5. The error probability as a function of the average detection delay. Performance comparison between the following algorithms: (i) The proposed DS algorithm that uses the MALLR statistics as described in Section III-C (referred to as the DS selection rule with MALLR); and (ii) The proposed DS algorithm that uses the MGLLR statistics. A case of Laplace distributions with parameters  $\theta^{(0)} = 0$ ,  $\theta^{(1)} = 1$ , under normal and abnormal states, respectively, with  $K = 2$ ,  $M = 5$ . We averaged over  $10^6$  Monte Carlo runs.

anomaly detection in packet size statistics, which has been mostly investigated using open loop strategies for detecting malicious activity. We use the model in [24] that proposed a sample entropy for packet-size modeling and demonstrated strong performance in detecting anomalous data using the GLR statistics in the sequential detection test. Specifically, for a given interval, let  $\mathcal{S}$  be the set of packet size values that have arrived in this interval, and let  $q^{(i)}$  be the proportion of number of packets of size  $i$  to the total number of packets that have arrived in that interval. The sample entropy  $y$  is thus computed as  $y = - \sum_{i \in \mathcal{S}} q^{(i)} \log q^{(i)}$ . The sample entropy is modeled by Gaussian distribution and given by:

$$p(y|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2\sigma^2} (y - \mu)^2 \right],$$

where  $\theta^{(0)} = (\mu_0, \sigma_0)$ , and  $\theta^{(1)} = (\mu_1, \sigma_1)$ , under normal state, or abnormal state, respectively. We simulated a network with  $M$  flows of data, in which a single flow is abnormal. We used the DARPA intrusion detection data set [53], which contains 5-million labeled network connections, for generating the normal and abnormal flows. When testing the algorithms, the sample entropy has been learned online from the data. We implemented both the proposed DS algorithm, and the entropy-based algorithm with the GLR statistics that has been proposed in [24]. We set the thresholds so that both algorithms satisfy error probability  $10^{-4}$ . It can be seen in Figure 6 that the DS algorithm achieves strong performance and significantly outperforms the entropy-based

algorithm with the GLR statistics.

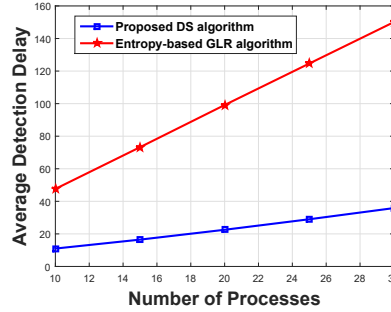


Fig. 6. The average detection delay as a function of the number of processes  $M$  using the DARPA intrusion detection dataset. Performance comparison between the following algorithms: (i) The proposed DS algorithm that uses the MALLR statistics as described in Section III-C (referred to as the proposed DS algorithm); and (ii) a policy that applies an open loop selection rule when probing cells and uses the GLR statistics for the packet size modeling in the stopping rule as proposed in [24] (referred to as entropy-based GLR algorithm). We averaged over  $10^5$  Monte Carlo runs.

## V. CONCLUSION

We considered the problem of searching for anomalies among  $M$  processes (i.e., cells). The observations follow a common distribution with an unknown parameter, belonging to disjoint parameter spaces depending on whether the target is absent or present. The decision maker is allowed to probe a subset of the cells at a time and the objective is a sequential search strategy that minimizes the expected detection time subject to an error probability constraint. We have developed a deterministic search algorithm to solve the problem that enjoys the following properties. First, when no additional side information on the process states is known, the proposed algorithm was shown to be asymptotically optimal. Second, when the parameter value under the null hypothesis is known and equal for all normal processes, asymptotic optimality was shown as well, with better detection time determined by the true null state. Third, when the parameter value under the null hypothesis is unknown, but is known to be equal for all normal processes, consistency was shown in terms of achieving error probability that decays to zero with the detection delay. Finally, an explicit upper bound on the error probability under the proposed algorithm was established under the finite sample regime. Extensive experiments have demonstrated the efficiency of the algorithm over existing methods.

## VI. ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for comments that significantly improved the technical results and presentation of this paper.

## VII. APPENDIX

For purposes of presentation, we start by proving Theorem 2. Then, we focus on the key steps for extending the results to the other models presented in Section III.

### A. Proof of Theorem 2

Without loss of generality we prove the theorem when hypothesis  $m$  is true. For simplifying the presentation, we start with proving the theorem when the parameter space is finite, so that  $\theta^{(0)}, \theta^{(1)}$  can take a finite number of values (but the measurements can still be continuous). We will then extend the proof for continuous parameter space under mild regularity conditions. The proof is derived using the adjusted LALLR statistics defined in (8), i.e.,  $S_m^{(0)}(n) = \sum_{t=1}^n \mathbf{1}_m(t) \log \frac{f(y_m(t)|\hat{\theta}_m(t-1))}{f(y_m(t)|\theta^{(0)})}$ .

#### *Step 1: Bounding the error probability:*

We first prove the upper bound on the error probability for all  $c$ . Specifically, we show below that the error probability is upper bounded by:

$$P_e = \sum_{m=1}^M \pi_m \alpha_m \leq (M-1)c. \quad (17)$$

Let

$$\alpha_{m,j} = \mathbf{P}_m(\delta = j)$$

for all  $j \neq m$ . Thus,

$$\alpha_m = \sum_{j \neq m} \alpha_{m,j}.$$

Therefore, we need to show that  $\alpha_{m,j} \leq c$  for proving (17). Note that  $\alpha_{m,j}$  can be rewritten as follows:

$$\begin{aligned}\alpha_{m,j} &= \mathbf{P}_m(\delta = j) \\ &= \mathbf{P}_m\left(S_j^{(0)}(\tau) \geq -\log c \text{ for some } \tau \geq 1\right) \\ &\leq \mathbf{P}_m\left(Z(N_j(\tau)) \geq \frac{1}{c} \text{ for some } N_j(\tau) \geq 1\right),\end{aligned}\tag{18}$$

where

$$Z(N_j(\tau)) \triangleq e^{S_j^{(0)}(\tau)} = \prod_{i=1}^{N_j(\tau)} \frac{f(y_j(r_i)|\hat{\theta}_j(r_i))}{f(y_j(r_i)|\theta^{(0)})},\tag{19}$$

and  $r_1, \dots, r_{N_j(\tau)}$  are the time indices in which observations are taken from cell  $j$ . Next, note that  $Z(N_j(\tau))$  is a nonnegative martingale,

$$\begin{aligned}\mathbf{E}_{\theta^{(0)}}\left[Z(N_j(\tau)) \mid \{y_j(r_i)\}_{i=1}^{N_j(\tau)-1}\right] \\ = Z(N_j(\tau) - 1) \mathbf{E}_{\theta^{(0)}}\left[\frac{f(y_j(r_{N_j(\tau)})|\hat{\theta}_j(r_{N_j(\tau)}))}{f(y_j(r_{N_j(\tau)})|\theta^{(0)})}\right] \\ = Z(N_j(\tau) - 1).\end{aligned}\tag{20}$$

Therefore, applying Lemma 1 in [51] for nonnegative martingales yields:

$$\begin{aligned}\mathbf{P}_m\left(Z(N_j(\tau)) \geq \frac{1}{c} \text{ for some } N_j(\tau) \geq 1\right) \\ \leq c \mathbf{E}_{\theta^{(0)}}[Z(1)].\end{aligned}\tag{21}$$

Finally, since  $\mathbf{E}_{\theta^{(0)}}[Z(1)] = 1$ , we have  $\alpha_{m,j} \leq c$ , which completes Statement 1 of the theorem.

Next, we define the following major event:

*Definition 2:*  $\tau_{ML}$  is the smallest integer such that  $\hat{\theta}_m(n) = \theta^{(1)}$ , and  $\hat{\theta}_j(n) = \theta^{(0)}$  for all  $j \neq m$  for all  $n > \tau_{ML}$ , when  $H_m$  is the true hypothesis.

*Remark 3:* Note that for all  $n > \tau_{ML}$  only the exploitation phase is implemented. As a result, the time spent during the round-robin exploration phase is upper bounded by  $\tau_{ML}$ . In the next step of the proof we show that  $\tau_{ML}$  is bounded, which also yields Statement 3 of the theorem. It should be noted that  $\tau_{ML}$  is not a stopping time. The decision maker does not know whether it has arrived. However, it is used to upper bound the actual stopping time under the DS algorithm.

*Remark 4:* For evaluating the detection time under the DS algorithm, we analyze the case where the DS algorithm is implemented indefinitely. When we say that the DS algorithm is implemented indefinitely we mean that we probe the cells as described by the DS algorithm, while disregarding the stopping rule. This analysis enables us to upper bound the actual detection time when the stopping rule is applied.

*Step 2: Bounding  $\tau_{ML}$ :*

*Lemma 1:* Assume that the DS algorithm is implemented indefinitely. Then, there exist  $C > 0$  and  $\gamma > 0$  such that

$$\mathbf{P}_m(\tau_{ML} > n) \leq Ce^{-\gamma n}. \quad (22)$$

*Proof:* Note that event  $\tau_{ML} > n$  implies one of the following events: (i) There exists a time instant  $t > n$  at the round-robin exploration phase, in which  $\hat{\theta}_m(t) \neq \theta^{(1)}$ , or  $\hat{\theta}_j(t) \neq \theta^{(0)}$  for some  $j \neq m$ . When such time  $t$  occurs we say that  $E_1(t)$  occurs. (ii) At the beginning of an exploitation phase (say at time  $n'$ )  $\hat{\theta}_m(n') = \theta^{(1)}$ , and  $\hat{\theta}_j(n') = \theta^{(0)}$  for all  $j \neq m$ . However, there exists a time instant  $t > n$  during the exploitation phase, in which  $\hat{\theta}_m(t) \neq \theta^{(1)}$ . When such time  $t$  occurs we say that  $E_2(t)$  occurs.

We can rewrite (22) as follows:

$$\begin{aligned} \mathbf{P}_m(\tau_{ML} > n) &\leq \mathbf{P}_m(E_1(t) \text{ occurs for some } t \geq n) \\ &\quad + \mathbf{P}_m(E_2(t) \text{ occurs for some } t \geq n) \\ &\leq \sum_{t=n}^{\infty} \mathbf{P}_m(E_1(t) \text{ occurs}) + \sum_{t=n}^{\infty} \mathbf{P}_m(E_2(t) \text{ occurs}). \end{aligned} \quad (23)$$

Next, we upper bound the first term on the RHS of (23). It suffices to show that there exist  $C > 0$  and  $\gamma > 0$  such that  $\mathbf{P}_m(E_1(n) \text{ occurs}) < Ce^{-\gamma n}$ . Let  $N_{RR}(n)$  be the total number of time instants spent during the round-robin exploration phase up to time  $n$ , and fix  $0 < r < 1$ . Then,  $\mathbf{P}_m(E_1(n) \text{ occurs})$  can be rewritten as follows:

$$\begin{aligned} \mathbf{P}_m(E_1(n) \text{ occurs}) &= \mathbf{P}_m(E_1(n) \text{ occurs}, N_{RR}(n) \geq rn) \\ &\quad + \mathbf{P}_m(E_1(n) \text{ occurs}, N_{RR}(n) < rn). \end{aligned} \quad (24)$$

We first upper bound the first term on the RHS of (24). Since that more than  $rn$  observations were taken in a round-robin manner, then at least  $rn/M$  observations were taken from each cell.

Thus,

$$\begin{aligned} & \mathbf{P}_m(E_1(n) \text{ occurs}, N_{RR}(n) \geq rn) \\ & \leq \mathbf{P}_m\left(\hat{\theta}_m(n) \neq \theta^{(1)}, N_m(n) \geq rn/M\right) \\ & \quad + \sum_{j \neq m} \mathbf{P}_m\left(\hat{\theta}_j(n) \neq \theta^{(0)}, N_j(n) \geq rn/M\right). \end{aligned} \quad (25)$$

Next, we show that the first term on the RHS of (25) decreases exponentially with  $n$ . Let  $(y_m(r_1), \dots, y_m(r_{N_m(n)}))$  be the vector of all  $N_m(n)$  observations (indicated by times  $r_1, \dots, r_{N_m(n)}$ ) collected from cell  $m$  up to time  $n$ , and let  $\tilde{\theta}_m(n') = \hat{\theta}_m(n)$  denotes the MLE based on  $N_m(n) = n'$  observations collected from cell  $m$  up to time  $n$ . We can upper bound  $\mathbf{P}_m\left(\hat{\theta}_m(n) \neq \theta^{(1)}, N_m(n) \geq rn/M\right)$  by:

$$\begin{aligned} & \mathbf{P}_m\left(\hat{\theta}_m(n) \neq \theta^{(1)}, N_m(n) \geq rn/M\right) \\ & \leq \sum_{q=\lceil rn/M \rceil}^{\infty} \mathbf{P}_m\left(\tilde{\theta}_m(q) \neq \theta^{(1)}\right). \end{aligned} \quad (26)$$

Then, by the definition of the MLE (5), the event  $\tilde{\theta}_m(n) \neq \theta^{(1)}$  implies:

$$\sum_{i=1}^n \ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i) < 0, \quad (27)$$

for some  $\tilde{\theta}_m(n) \neq \theta^{(1)}$ , where

$$\ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i) \triangleq \log \frac{f(y_m(i)|\theta^{(1)})}{f(y_m(i)|\tilde{\theta}_m(n))}.$$

Note that we only refer to the number of observations irrespective of the probing times due to i.i.d. property. Hence, it remains to show that  $\mathbf{P}_m\left(\sum_{i=1}^n \ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i) < 0\right)$  decreases exponentially with  $n$  for each  $\tilde{\theta}_m(n) \neq \theta^{(1)}$ . Applying the Chernoff bound and using the i.i.d. property yields:

$$\begin{aligned} & \mathbf{P}_m\left(\sum_{i=1}^n \ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i) < 0\right) \\ & \leq \left[\mathbf{E}_m\left(e^{s(-\ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i))}\right)\right]^n. \end{aligned} \quad (28)$$

Note that a moment generating function (MGF) is equal to one at  $s = 0$ . Furthermore, since  $\mathbf{E}_m(-\ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i)) = -D\left(\theta^{(1)} || \tilde{\theta}_m(n)\right) < 0$  is strictly negative, differentiating the MGFs of  $-\ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i)$  with respect to  $s$  yields a strictly negative derivative at  $s = 0$ . Hence, there exist  $s > 0$  and  $\gamma' > 0$  such that  $\mathbf{E}_m\left(e^{s(-\ell_{\theta^{(1)}, \tilde{\theta}_m(n)}(i))}\right)$  is strictly less than  $e^{-\gamma'} < 1$ , which yields the desired exponential decay. A similar argument applies for showing that the second term on the RHS of (25) decreases exponentially with  $n$ .

Next, we upper bound the second term on the RHS of (24). Let  $N_{XT}$ ,  $N_{XT,j}$  be the total number of time instants spent during the exploitation phase up to time  $n$  at all cells and cell  $j$ , respectively. Since  $N_{RR}(n) < rn$ , then  $N_{XT} \geq (1-r)n$ . Fix  $0 < r_2 < 1$ . We can rewrite the second term on the RHS of (24) as follows:

$$\begin{aligned}
& \mathbf{P}_m(E_1(n) \text{ occurs}, N_{RR}(n) < rn) \\
& \leq \mathbf{P}_m(E_1(n) \text{ occurs}, N_{XT}(n) \geq (1-r)n, \\
& \quad N_{XT,m}(n) \geq r_2(1-r)n) \\
& + \mathbf{P}_m(E_1(n) \text{ occurs}, N_{XT}(n) \geq (1-r)n, \\
& \quad N_{XT,m}(n) < r_2(1-r)n).
\end{aligned} \tag{29}$$

We first upper bound the first term on the RHS of (29). Note that  $E_1(n)$  occurs implies that there exists an exploitation time  $t$  before time  $n$ , in which cell  $m$  has been probed, its MLE was computed based on more than  $r_2(1-r)n$  observations, and error event was occurred,  $\hat{\theta}_m(t) \neq \theta^{(1)}$  (so that the algorithm moved back to exploration phase and  $E_1(n)$  occurred). Therefore, we can write:

$$\begin{aligned}
& \mathbf{P}_m(E_1(n) \text{ occurs}, N_{XT}(n) \geq (1-r)n, \\
& \quad N_{XT,m}(n) \geq r_2(1-r)n) \\
& \leq \sum_{t=\lceil r_2(1-r)n \rceil}^{\infty} \mathbf{P}_m(\hat{\theta}_m(t) \neq \theta^{(1)}, N_m(t) \geq r_2(1-r)n).
\end{aligned} \tag{30}$$

By a similar argument as we developed when proving (25), each of the terms in the summation decreases exponentially with  $n$ , implying exponentially decreasing of the first term on the RHS of (29).

Next, we upper bound the second term on the RHS of (29). Since less than  $r_2(1-r)n$  observations were taken from cell  $m$  during exploitation and the total number of observations during exploitation is more than  $(1-r)n$ , then there exists cell  $j \neq m$  that have been observed more than  $(1-r_2)(1-r)n/(M-1)$  times during exploitation phase. This implies that there exists an exploitation time  $t$  before time  $n$ , in which cell  $j$  has been probed, its MLE was computed based on more than  $(1-r_2)(1-r)n/(M-1) = qn$  observations, where  $0 < q \triangleq$

$(1 - r_2)(1 - r)/(M - 1) < 1$ , and  $\hat{\theta}_j(t) \neq \theta^{(0)}$  Therefore, we can write:

$$\begin{aligned} & \mathbf{P}_m(E_1(n) \text{ occurs}, N_{XT}(n) \geq (1 - r)n, \\ & \quad N_{XT,m}(n) < r_2(1 - r)n) \\ & \leq \sum_{t=\lceil qn \rceil}^{\infty} \mathbf{P}_m\left(\hat{\theta}_j(t) \neq \theta^{(0)}, N_j(t) \geq qn\right). \end{aligned} \quad (31)$$

By a similar argument as we developed when proving (25), each of the terms in the summation decreases exponentially with  $n$ , implying exponentially decreasing of the second term on the RHS of (29). Therefore, we have shown exponentially decreasing of the first term on the RHS of (23).

It remains to show an exponentially decreasing of the second term on the RHS of (23). It suffices to show that there exist  $C > 0$  and  $\gamma > 0$  such that  $\mathbf{P}_m(E_2(n) \text{ occurs}) < Ce^{-\gamma n}$ . Fix  $0 < r < 1$ . We can rewrite  $\mathbf{P}_m(E_2(n) \text{ occurs})$  as follows:

$$\begin{aligned} \mathbf{P}_m(E_2(n) \text{ occurs}) & \leq \mathbf{P}_m(E_2(n) \text{ occurs}, N_m(n) \geq rn) \\ & \quad + \mathbf{P}_m(E_2(n) \text{ occurs}, N_m(n) < rn). \end{aligned} \quad (32)$$

We first upper bound the first term on the RHS of (32). Since  $E_2(n)$  occurs and more than  $rn$  observations were taken from cell  $m$  we have:

$$\begin{aligned} & \mathbf{P}_m(E_2(n) \text{ occurs}, N_m(n) \geq rn) \\ & \leq \mathbf{P}_m\left(\hat{\theta}_m(n) \neq \theta^{(1)}, N_m(n) \geq rn\right). \end{aligned} \quad (33)$$

By a similar argument as we developed when proving (25), the RHS of (33) decreases exponentially with  $n$ .

Next, we upper bound the second term on the RHS of (32). Since  $N_m(n) < rn$  then at least  $(1 - r)n$  observations were taken from other cells. Let  $\tilde{N}_{RR}(n)$  be the total number of observations collected from all cells excepts cell  $m$  during the round-robin exploration phase up to time  $n$ , and fix  $0 < r_2 < 1$ . Then, the second term on the RHS of (32) can be rewritten as follows:

$$\begin{aligned} & \mathbf{P}_m(E_2(n) \text{ occurs}, N_m(n) < rn) \\ & = \mathbf{P}_m\left(E_2(n) \text{ occurs}, N_m(n) < rn, \tilde{N}_{RR}(n) \geq r_2(1 - r)n\right) \\ & \quad + \mathbf{P}_m\left(E_2(n) \text{ occurs}, N_m(n) < rn, \tilde{N}_{RR}(n) < r_2(1 - r)n\right). \end{aligned} \quad (34)$$

Next, we upper bound the first term on the RHS of (34). Since more than  $r_2(1 - r)n$  observations were taken from all cells excepts cell  $m$  during round-robin exploration, then at least  $r_2(1 -$

$r)n/(M-1)$  observations were taken from each cell  $j \neq m$  (and the same number of observations must have been taken from cell  $m$  as well during round-robin exploration). Then, at time  $n$  during exploration phase, its MLE was computed based on more than  $r_2(1-r)n/(M-1)$  observations, and error event was occurred,  $\hat{\theta}_m(t) \neq \theta^{(1)}$  (so that  $E_2(n)$  occurred). Then,

$$\begin{aligned} & \mathbf{P}_m \left( E_2(n) \text{ occurs}, N_m(n) < rn, \tilde{N}_{RR}(n) \geq r_2(1-r)n \right) \\ & \leq \mathbf{P}_m \left( \hat{\theta}_m(n) \neq \theta^{(1)}, N_m(n) \geq r_2(1-r)n/(M-1) \right). \end{aligned} \quad (35)$$

By a similar argument as we developed when proving (25), the RHS of (35) decreases exponentially with  $n$ .

Next, we upper bound the second term on the RHS of (34). Since less than  $r_2(1-r)n$  observations were taken from all cells excepts cell  $m$  during round-robin exploration, then there exists cell  $j \neq m$  in which more than  $(1-r_2)(1-r)n/(M-1)$  observations were taken from it during exploitation phase. By subtracting all time instants in which the test might switch between exploration to exploitation phases, at least  $(1-r_2)(1-r)n/(M-1) - r_2(1-r)n$  observations were taken during exploitation, where  $\hat{\theta}_j(t) \neq \theta^{(0)}$ . We can choose small  $r_2$  (e.g.,  $r_2 = 1/(3(M-1))$ ) so that  $(1-r_2)(1-r)n/(M-1) - r_2(1-r)n = qn$  for  $0 < q < 1$ . Thus,

$$\begin{aligned} & \mathbf{P}_m \left( E_2(n) \text{ occurs}, N_m(n) < rn, \tilde{N}_{RR}(n) < r_2(1-r)n \right) \\ & \leq \mathbf{P}_m \left( \hat{\theta}_j(n) \neq \theta^{(0)}, N_j(n) \geq qn \right). \end{aligned} \quad (36)$$

By a similar argument as we developed when proving (25), the RHS of (35) decreases exponentially with  $n$ . Hence, (22) follows.  $\blacksquare$

Note that the total time spent during the round-robin exploration phase is upper bounded by  $\tau_{ML}$ . Hence, Statement 3 in Theorem 2 follows.

*Step 3: Bounding the detection time:*

*Definition 3:* Assume that the DS algorithm is implemented indefinitely. Then,  $\tau_U$  denotes the first time that  $S_m^{(0)}(n) \geq -\log(c)$  for  $n > \tau_{ML}$ :

$$\tau_U \triangleq \inf \{ n > \tau_{ML} : S_m^{(0)}(n) \geq -\log c \}, \quad (37)$$

and  $n_U \triangleq \tau_U - \tau_{ML}$  denotes the total amount of time between  $\tau_{ML}$  and  $\tau_U$ .

It should be noted that the actual detection time  $\tau$  under DS algorithm (when the stopping rule is applied) is upper bounded by  $\tau_U$ . In the next lemma we show that  $n_U$  cannot be significantly larger than  $-(1+\epsilon)\log c/D(\theta^{(1)}||\theta^{(0)})$  with high probability.

*Lemma 2:* Assume that the DS algorithm is implemented indefinitely and  $H_m$  is true. Then, for every fixed  $\epsilon > 0$  there exist  $C > 0$  and  $\gamma > 0$  such that

$$\begin{aligned} \mathbf{P}_m(n_U > n) &\leq C e^{-\gamma n} \\ \forall n &> -(1 + \epsilon) \log c / D(\theta^{(1)} || \theta^{(0)}) . \end{aligned} \quad (38)$$

*Proof:* We define

$$\tilde{\ell}_m(t) \triangleq \ell_m(t) - D(\theta^{(1)} || \theta^{(0)}), \quad (39)$$

where the MALLR  $\ell_m(t)$  at time  $n \geq t$  is given by:

$$\ell_m(t) \triangleq \log \frac{f(y_m(t) | \hat{\theta}_m(t))}{f(y_m(t) | \hat{\theta}^{(r)}(n))}. \quad (40)$$

Recall that the test statistics is given by  $S_m^{(0)}(n) = \sum_{t=1}^n \mathbf{1}_m(t) \ell_m(t)$ . Since that for all  $t \geq \tau_{ML}$  the DS algorithm collects observations from cell  $m$ , then  $\mathbf{1}_m(t) = 1$  for all  $t \geq \tau_{ML}$ . Let  $\epsilon_1 = D(\theta^{(1)} || \theta^{(0)}) \epsilon / (1 + \epsilon) > 0$ . Then, we can write

$$\begin{aligned} &\sum_{i=1}^{\tau_{ML}+n} \mathbf{1}_m(i) \ell_m(i) + \log c \\ &= \sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) + \sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \ell_m(i) + \log c \\ &= \sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) + \sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \tilde{\ell}_m(i) + n D(\theta^{(1)} || \theta^{(0)}) + \log c \\ &\geq \sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) + \sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \tilde{\ell}_m(i) + n \epsilon_1 , \end{aligned} \quad (41)$$

for all  $n > -(1 + \epsilon) \log c / D(\theta^{(1)} || \theta^{(0)})$ .

As a result,

$$\sum_{i=1}^{\tau_{ML}+n} \mathbf{1}_m(i) \ell_m(i) \leq -\log c . \quad (42)$$

implies

$$\sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) + \sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \tilde{\ell}_m(i) \leq -n \epsilon_1 . \quad (43)$$

Hence, for any  $\epsilon > 0$  there exists  $\epsilon_1 > 0$  such that

$$\begin{aligned}
& \mathbf{P}_m(n_U > n) \\
& \leq \mathbf{P}_m\left(\sum_{i=1}^{\tau_{ML}+n} \mathbf{1}_m(i) \ell_m(i) \leq -\log c\right) \\
& \leq \mathbf{P}_m\left(\sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) + \sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \tilde{\ell}_m(i) \leq -n\epsilon_1\right) \\
& \leq \mathbf{P}_m\left(\sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) \leq -n\epsilon_1/2\right) \\
& \quad + \mathbf{P}_m\left(\sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \tilde{\ell}_m(i) \leq -n\epsilon_1/2\right) \\
& \leq \mathbf{P}_m\left(\sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) \leq -n\epsilon_1/2, \tau_{ML} > \epsilon_2 n\right) \\
& \quad + \mathbf{P}_m\left(\sum_{i=1}^{\tau_{ML}} \mathbf{1}_m(i) \ell_m(i) \leq -n\epsilon_1/2, \tau_{ML} \leq \epsilon_2 n\right) \\
& \quad + \mathbf{P}_m\left(\sum_{i=\tau_{ML}+1}^{\tau_{ML}+n} \tilde{\ell}_m(i) \leq -n\epsilon_1/2\right)
\end{aligned} \tag{44}$$

for all  $n > -(1 + \epsilon) \log c / D(\theta^{(1)} || \theta^{(0)})$ , and  $0 < \epsilon_2 < 1$ . The first term on the RHS decreases exponentially by Lemma 1. Since  $\epsilon_2 > 0$  can be arbitrarily small, and  $\ell_m(i)$  has finite expectation, then the second term decreases exponentially by applying the Chernoff bound. Since  $\tilde{\ell}_m(i)$  has zero mean for all  $i > \tau_{ML}$ , then the third term decrease exponentially by applying the Chernoff bound. Hence, (38) follows.  $\blacksquare$

Next, we can upper bound the actual detection time under DS algorithm by combining Lemmas 1, 2:

$$\mathbf{E}_m(\tau) \leq \mathbf{E}_m(\tau_{ML}) + \mathbf{E}_m(n_U) \leq -(1 + o(1)) \frac{\log(c)}{D(\theta^{(1)} || \theta^{(0)})}. \tag{45}$$

Next, we can upper bound the Bayes risk under DS algorithm By combining (45) and (17):

$$R_m(\Gamma) \leq -(1 + o(1)) \frac{c \log(c)}{D(\theta^{(1)} || \theta^{(0)})}. \tag{46}$$

Finally, Combining the upper bound on the Bayes risk with the lower bound on the Bayes risk  $R_m(\Gamma) \geq -(1 + o(1)) \frac{c \log(c)}{D(\theta^{(1)} || \theta^{(0)})}$  that was obtained in [8] under simple hypotheses completes the proof.  $\blacksquare$

1) *Extending the Proof of Theorem 2 for Continuous Parameter Space:* Next, we focus on the key steps used for extending the proof of Theorem 2 for continuous parameter space. We need the following requirement for the consistency of the MLE: For all  $\epsilon > 0$ , we require that  $\mathbf{P}_m \left( |\tilde{\theta}_m(n) - \theta^{(1)}| > \epsilon \right)$  decays only polynomially with  $n$ . To achieve this, we require that the parameters space  $\Theta^{(0)}$ ,  $\Theta^{(1)}$  are open sets. Then, the condition holds for a wide class of distributions, including exponential family distributions (see e.g., [54]).

*Step 1: Bounding the error probability:*

Bounding the error probability in Step 1 in Appendix VII-A directly applies to continuous parameter space.

*Step 2: Bounding  $\tau_{ML}(\epsilon_3)$ :*

Since the MLEs take continuous values, instead of defining  $\tau_{ML}$  as in Appendix VII-A, we define  $\tau_{ML}(\epsilon_3)$  for some  $\epsilon_3 > 0$  as the smallest integer such that  $|\hat{\theta}_m(n) - \theta^{(1)}| \leq \epsilon_3$ , and  $|\hat{\theta}_j(n) - \theta^{(0)}| \leq \epsilon_3$  for all  $j \neq m$  for all  $n > \tau_{ML}(\epsilon_3)$ , when  $H_m$  is the true hypothesis. We require that the parameters take values in the interior of the parameter spaces. Then, we can choose a sufficiently small  $\epsilon_3$  so that for all  $n > \tau_{ML}(\epsilon_3)$  only the exploitation phase is implemented. As a result, the time spent in a round-robin exploration phase is upper bounded by  $\tau_{ML}(\epsilon_3)$  similarly to upper bounding the round-robin exploration time by  $\tau_{ML}$  as in Appendix VII-A.

We then modify Lemma 1 so that to show at least polynomial decay of  $\mathbf{P}_m(\tau_{ML}(\epsilon_3) > n)$  for any  $\epsilon_3 > 0$  (since polynomial decay is sufficient to guarantee a finite expected value of  $\tau_{ML}(\epsilon_3)$ ). Proving the modified lemma requires similar steps as in Appendix VII-A with the following modification. Since the MLEs take continuous values, instead of referring to the events in which the MLEs are not equal to the true parameter values  $\hat{\theta}_m(n) \neq \theta^{(1)}$ , and  $\hat{\theta}_j(n) \neq \theta^{(0)}$  for all  $j \neq m$  as in Appendix VII-A, we refer to the events in which the MLEs deviate from the true parameter values by  $\epsilon_3$ . As a result, for bounding  $\mathbf{P}_m(\tau_{ML}(\epsilon_3) > n)$  (as in equation (26)) we need to require only weak consistency of the MLEs so that  $\mathbf{P}_m \left( |\tilde{\theta}_m(n) - \theta^{(1)}| > \epsilon_3 \right)$  decays only polynomially with  $n$  as mentioned above.

*Step 3: Bounding the detection time:*

Step 3 follows similar steps as in Appendix VII-A for any  $\epsilon_3 > 0$ . Since  $\epsilon_3 > 0$  is arbitrarily small, the theorem follows.

### B. Proof of Theorem 3

We now prove the asymptotic optimality criterion given in (4). Here we are interested in detection procedures that satisfy the constraint  $P_e(\Gamma) \leq \alpha$ . The class of such detection algorithms will be denoted by  $C(\alpha)$ .

By applying the same steps as in the proof of Theorem 1, we can show that:  $P_e(\Gamma) \leq (M-1)e^{-a}$ . It follows that  $a = \log(\frac{M-1}{\alpha})$  implies that  $\Gamma \in C(\alpha)$ .

Next, the upper bound for  $\tau_{ML}$  holds with the same steps as in Lemma 1. We define  $\tau_U$  and  $n_U$  as in VII-A. To prove the asymptotic optimality, first note that the lower bound on the detection time is given by:

$$\inf_{\Gamma \in C(\alpha)} \mathbb{E}_m(\tau|\Gamma) \geq \frac{|\log \alpha|}{D(\theta^{(1)}||\theta^{(0)})}(1 + o(1)), \quad \alpha \rightarrow 0, \quad (47)$$

which can be derived following the same steps as in [55]. We next provide the proof for the upper bound on the detection time.

*Lemma 3:* Assume that the DS algorithm is implemented indefinitely. Then,

$$\mathbb{E}_m(\tau|\Gamma^*) \leq \frac{|\log \alpha|}{D(\theta^{(1)}||\theta^{(0)})}(1 + o(1)), \quad \alpha \rightarrow 0. \quad (48)$$

*Proof:* We define the last exit times  $L(\epsilon, \theta)$ . For all  $\epsilon > 0$ :

$$L(\epsilon, \theta) = \sup\{n \geq \tau_{ML} \mid \frac{S_m^{(0)}(n)}{n} - D(\theta^{(1)}||\theta^{(0)}) > \epsilon\}. \quad (49)$$

Under  $H_m$ ,

$$S_m^{(0)}(n_U - 1) \geq (n_U - 1)(D(\theta^{(1)}||\theta^{(0)}) - \epsilon)$$

on  $\{n_U > L(\epsilon, \theta) + 1\}$ , and

$$S_m^{(0)}(n_U - 1) < a, \text{ on } \{n_U < \infty\}.$$

Therefore, for every  $0 < \epsilon < D(\theta^{(1)}||\theta^{(0)})$ ,

$$\begin{aligned} n_U &< \left(1 + \frac{a}{D(\theta^{(1)}||\theta^{(0)}) - \epsilon}\right) \mathbf{1}_{\{n_U > 1 + L(\epsilon, \theta)\}} \\ &+ [1 + L(\epsilon, \theta)] \mathbf{1}_{\{n_U < 1 + L(\epsilon, \theta)\}} \\ &\leq 1 + L(\epsilon, \theta) + \frac{a}{D(\theta^{(1)}||\theta^{(0)}) - \epsilon}. \end{aligned}$$

By using Chernoff bound we can show that  $E[L(\epsilon, \theta)] < \infty$ , and by letting  $\epsilon \rightarrow 0$  and choosing

$a = \log(\frac{M-1}{\alpha})$  we get:

$$E_m[n_U] \leq \frac{|\log \alpha|}{D(\theta^{(1)} || \theta^{(0)})} (1 + o(1)) \text{ as } \alpha \rightarrow 0,$$

and combining with the upper bound of  $\tau_{ML}$  derived in Lemma 1, we prove the Lemma.  $\blacksquare$

Finally, by combining Lemma 3, with (47) we complete the proof.

### C. Proof of Theorem 1

We focus on the key steps used for extending the proof of Theorem 2 to the settings in which both parameters under normal and abnormal states  $\theta^{(0)}, \theta^{(1)}$  are unknown, and no additional side information of the parameter values are given. Without loss of generality we prove the theorem when hypothesis  $m$  is true, and the proof is derived using the LALLR statistics defined in (7), i.e.,  $S_m^{(0)}(n) = S_{m,LALLR}^{(0)}(n)$ .

*Step 1: Bounding the error probability:*

We begin by upper bounding the error probability for all  $c$ . With the same notation as in Step 1 in Appendix VII-A, we need to show that  $\alpha_{m,j} = \mathbf{P}_m(\delta = j) \leq c$ . We first notice:

$$\begin{aligned} S_j^{(0)}(\tau) &= \sum_{t=1}^{\tau} \mathbf{1}_j(t) \log \frac{f(y_j(t) | \hat{\theta}_j(t-1))}{f(y_j(t) | \hat{\theta}_j^{(0)}(\tau))} \\ &= \min_{\varphi \in \Theta^{(0)}} \sum_{t=1}^{\tau} \mathbf{1}_j(t) \log \frac{f(y_j(t) | \hat{\theta}_j(t-1))}{f(y_j(t) | \varphi)} \\ &\leq \sum_{t=1}^{\tau} \mathbf{1}_j(t) \log \frac{f(y_j(t) | \hat{\theta}_j(t-1))}{f(y_j(t) | \theta^{(0)})}. \end{aligned}$$

Hence, we have:

$$\begin{aligned} \alpha_{m,j} &= \mathbf{P}_m(\delta = j) \\ &= \mathbf{P}_m \left( S_j^{(0)}(\tau) \geq -\log c \text{ for some } \tau \geq 1 \right) \\ &\leq \mathbf{P}_m \left( \sum_{t=1}^{\tau} \mathbf{1}_j(t) \log \frac{f(y_j(t) | \hat{\theta}_j(t-1))}{f(y_j(t) | \theta^{(0)})} \geq -\log c \right. \\ &\quad \left. \text{for some } \tau \geq 1 \right). \end{aligned}$$

Next, we can use similar steps as in Appendix VII-A, starting at (18) onwards, to prove that  $\alpha_{m,j} \leq c$ , which implies that the error probability is upper bounded by  $(M-1)c$  for all  $c$ . Thus, Statement 1 in Theorem 1 follows.

*Step 2: Bounding  $\tau_{ML}$ :*

Upper bounding  $\tau_{ML}$  (see (22)) follows the same steps as in Lemma 1. Hence, Statement 3 in

Theorem 1 follows.

*Step 3: Bounding the detection time:*

We define  $\tau_U$  and  $n_U$  similarly as in step 3 in Appendix VII-A:

*Definition 4:* Assume that the DS algorithm is implemented indefinitely. Then,  $\tau_U$  denotes the first time that  $S_m^{(0)}(n) \geq -\log(c)$  for  $n > \tau_{ML}$ :

$$\tau_U \triangleq \inf \{n > \tau_{ML} : S_m^{(0)}(n) \geq -\log c\}, \quad (50)$$

and  $n_U \triangleq \tau_U - \tau_{ML}$  denotes the total amount of time between  $\tau_{ML}$  and  $\tau_U$ .

We also define  $\ell_m^{(\theta, \varphi)}(t) = \log \frac{f(y_m(t)|\theta)}{f(y_m(t)|\varphi)}$ .

Note that  $S_m^{(0)}(n) = \sum_{t=1}^n \mathbf{1}_j(t) \ell_m^{(\theta^{(1)}, \hat{\theta}_m^{(0)}(n))}(t)$  for all  $n > \tau_{ML}$ . Define  $\tau_U(\varphi)$  to be the first time that  $\sum_{t=1}^n \mathbf{1}_m(t) \ell_m^{(\theta^{(1)}, \varphi)}(t) \geq -\log c$  for  $n > \tau_{ML}$ , and define  $n_U(\varphi) = \tau_U(\varphi) - \tau_{ML}$ . Clearly,  $n_U \leq n_U(\varphi)$  for each  $\varphi \in \Theta^{(0)}$ . We now bound  $n_U(\varphi)$  for each  $\varphi \in \Theta^{(0)}$ .

*Lemma 4:* Assume that the DS algorithm is implemented indefinitely and  $H_m$  is true. Then, for each  $\varphi \in \Theta^{(0)}$  and for every fixed  $\epsilon > 0$  there exist  $C > 0$  and  $\gamma > 0$  such that

$$\begin{aligned} \mathbf{P}_m(n_U(\varphi) > n) &\leq C e^{-\gamma n} \\ \forall n &> -(1 + \epsilon) \log c / D(\theta^{(1)}) . \end{aligned} \quad (51)$$

*Proof:* Define  $\tilde{\ell}_m^{(\theta^{(1)}, \varphi)}(t) = \ell_m^{(\theta^{(1)}, \varphi)}(t) - D(\theta^{(1)}||\varphi)$ . Using the same steps as in the proof of Lemma 2 (choosing this time  $\epsilon_1 = D(\theta^{(1)}||\varphi)\epsilon/(1 + \epsilon) > 0$ ), equation (44) holds with  $\ell_m^{(\theta^{(1)}, \varphi)}(t)$  and  $\tilde{\ell}_m^{(\theta^{(1)}, \varphi)}(t)$  instead of  $\ell_m(t)$  and  $\tilde{\ell}_m(t)$ , respectively. Again, since  $\tilde{\ell}_m^{(\theta^{(1)}, \varphi)}(t)$  has zero mean for all  $t > \tau_{ML}$ , all the three terms can be bounded as done in (44).

Since  $D(\theta^{(1)}||\varphi) \geq D(\theta^{(1)})$ ,  $\forall \varphi \in \Theta^{(0)}$ , (51) follows. ■

Using Lemma 4 we have:

$$\mathbf{P}_m(n_U > n) \leq \mathbf{P}_m(n_U(\varphi) > n) \leq C e^{-\gamma n}$$

$\forall n > -(1 + \epsilon) \log c / D(\theta^{(1)})$ , thus, the actual detection time is upper

bounded by:

$$\mathbf{E}_m(\tau) \leq \mathbf{E}_m(\tau_{ML}) + \mathbf{E}_m(n_U) \leq -(1 + o(1)) \frac{\log(c)}{D(\theta^{(1)})},$$

and using the bound on the error probability obtained in step 1, the Bayes risk is upper bounded by:

$$R_m(\Gamma) \leq -(1 + o(1)) \frac{c \log(c)}{D(\theta^{(1)})}.$$

Combining the upper bound with the lower bound from [13] completes the proof.  $\blacksquare$

#### D. Proof of Theorem 4

For purposes of analysis we require that the stopping rule does not stop the test before  $-\epsilon \log c$  samples have been taken, where  $\epsilon > 0$  is arbitrarily small. Also, when updating the sum MALLRs at time  $t$ , we use the current estimates for all  $n = 1, \dots, t$ .

Without loss of generality, let  $H_m$  be the true hypothesis. Let  $P_e = \sum_{m=1}^M \pi_m \alpha_m$  be the error probability, where

$$\alpha_{m,j} = \mathbf{P}_m(\delta = j)$$

for all  $j \neq m$ . Thus,

$$\alpha_m = \sum_{j \neq m} \alpha_{m,j}.$$

Therefore, we need to show that  $\alpha_{m,j}$  decays polynomially with  $-\log c$ . Note that  $\alpha_{m,j}$  can be rewritten as follows:

$$\begin{aligned} \alpha_{m,j} &= \mathbf{P}_m(\delta = j) = \mathbf{P}_m(\delta = j, \tau_{ML} > \tau) \\ &\quad + \mathbf{P}_m(\delta = j, \tau_{ML} \leq \tau). \end{aligned} \tag{52}$$

Since the stopping rule does not stop the test before  $-\epsilon \log c$  samples have been taken, the first term on the RHS is upper bounded by  $C\tau^{-\gamma} \leq C(-\epsilon \log c)^{-\gamma}$ , for some constants  $C, \gamma, \epsilon > 0$ , resulting in a polynomial decay with  $-\log c$ . Thus, it remains to show that the second term on the RHS decreases polynomially with  $-\log c$ .

Accepting  $H_j$  at time  $n$  implies  $S_j^{(0)}(n) + S_{m^{(1)}(n)}^{(1)}(n) \geq -\log c$ , which implies  $S_j^{(0)}(n) + S_m^{(1)}(n) \geq -\log c$ . Hence, for all  $j \neq m$  we obtain:

$$\begin{aligned} &\mathbf{P}_m(\delta = j, \tau_{ML} \leq \tau) \\ &\leq \mathbf{P}_m\left(S_j^{(0)}(n) + S_m^{(1)}(n) \geq -\log c, \tau_{ML} \leq \tau\right) \\ &\leq c\mathbf{P}_j\left(S_j^{(0)}(n) + S_m^{(1)}(n) \geq -\log c, \tau_{ML} \leq \tau\right) \leq c, \end{aligned} \tag{53}$$

where changing the measure in the second inequality follows by the fact that  $S_j^{(0)}(n) + S_m^{(1)}(n) \geq -\log c$ , and that the estimates are given by the true parameters for all  $\tau \geq \max\{\tau_{ML}, \tilde{\tau}_{ML}\}$  (where the current estimates are updated for all  $n = 1, \dots, \tau$ ). As a result, the theorem follows. ■

## REFERENCES

- [1] B. Hemo, K. Cohen, and Q. Zhao, "Asymptotically optimal search of unknown anomalies," in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 75–80, 2016.
- [2] L. Lai, H. V. Poor, Y. Xin, and G. Georgiadis, "Quickest search over multiple sequences," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5375–5386, 2011.
- [3] M. L. Malloy, G. Tang, and R. D. Nowak, "Quickest search for a rare distribution," *IEEE Annual Conference on Information Sciences and Systems*, pp. 1–6, 2012.
- [4] A. Tajer and H. V. Poor, "Quick search for rare events," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4462–4481, 2013.
- [5] K. Cohen, Q. Zhao, and A. Swami, "Optimal index policies for anomaly localization in resource-constrained cyber systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4224–4236, 2014.
- [6] K. S. Zigangirov, "On a problem in optimal scanning," *Theory of Probability and Its Applications*, vol. 11, no. 2, pp. 294–298, 1966.
- [7] D. A. Castanon, "Optimal search strategies in dynamic hypothesis testing," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 7, pp. 1130–1138, 1995.
- [8] K. Cohen and Q. Zhao, "Active hypothesis testing for anomaly detection," *IEEE Transactions on Information Theory*, vol. 61, no. 3, pp. 1432–1450, 2015.
- [9] N. K. Vaidhiyan and R. Sundaresan, "Learning to detect an oddball target," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 831–852, 2018.
- [10] S. Nitinawarat and V. V. Veeravalli, "Universal scheme for optimal search and stop," in *Bernoulli*, vol. 23, pp. 1759–1783, March 2017.
- [11] B. Huang, K. Cohen, and Q. Zhao, "Active anomaly detection in heterogeneous processes," *IEEE Transactions on Information Theory*, vol. 65, no. 4, pp. 2284–2301, 2018.
- [12] A. Gurevich, K. Cohen, and Q. Zhao, "Sequential anomaly detection under a nonlinear system cost," *IEEE Transactions on Signal Processing*, 2019.
- [13] H. Chernoff, "Sequential design of experiments," *The Annals of Mathematical Statistics*, vol. 30, no. 3, pp. 755–770, 1959.
- [14] G. Schwarz, "Asymptotic shapes of Bayes sequential testing regions," *The Annals of mathematical statistics*, pp. 224–236, 1962.
- [15] T. L. Lai, "Nearly optimal sequential tests of composite hypotheses," *The Annals of Statistics*, pp. 856–886, 1988.
- [16] I. V. Pavlov, "Sequential procedure of testing composite hypotheses with applications to the Kiefer-Weiss problem," *Theory of Probability and Its Applications*, vol. 35, no. 2, pp. 280–292, 1990.
- [17] A. G. Tartakovsky, "An efficient adaptive sequential procedure for detecting targets," in *IEEE Aerospace Conference Proceedings, 2002*, vol. 4, pp. 1581–1596, 2002.

- [18] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2451–2464, 2013.
- [19] M. Naghshvar and T. Javidi, "Active sequential hypothesis testing," *The Annals of Statistics*, vol. 41, no. 6, pp. 2703–2738, 2013.
- [20] M. Naghshvar and T. Javidi, "Sequentiality and adaptivity gains in active hypothesis testing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 768–782, 2013.
- [21] S. Nitinawarat and V. V. Veeravalli, "Controlled sensing for sequential multihypothesis testing with controlled markovian observations and non-uniform control cost," *Sequential Analysis*, vol. 34, no. 1, pp. 1–24, 2015.
- [22] Y. Song and G. Fellouris, "Sequential multiple testing with generalized error control: an asymptotic optimality theory," *arXiv preprint arXiv:1608.07014*, 2016.
- [23] Y. Song and G. Fellouris, "Asymptotically optimal, sequential, multiple testing procedures with prior information on the number of signals," *Electronic Journal of Statistics*, vol. 11, no. 1, pp. 338–363, 2017.
- [24] G. Thatte, U. Mitra, and J. Heidemann, "Parametric methods for anomaly detection in aggregate traffic," *IEEE/ACM Transactions on Networking*, vol. 19, no. 2, pp. 512–525, 2011.
- [25] J. Font-Segura and X. Wang, "Glrt-based spectrum sensing for cognitive radio with prior information," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2137–2146, 2010.
- [26] K. Cohen and Q. Zhao, "Asymptotically optimal anomaly detection via sequential testing," *IEEE Transactions on Signal Processing*, vol. 63, no. 11, pp. 2929–2941, 2015.
- [27] K. P. Tognetti, "An optimal strategy for a whereabouts search," *Operations Research*, vol. 16, no. 1, pp. 209–211, 1968.
- [28] J. B. Kadane, "Optimal whereabouts search," *Operations Research*, vol. 19, no. 4, pp. 894–904, 1971.
- [29] Y. Zhai and Q. Zhao, "Dynamic search under false alarms," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2013.
- [30] M. Egan, J.-M. Gorce, and L. Cardoso, "Fast initialization of cognitive radio systems," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, 2017.
- [31] Q. Zhao and J. Ye, "Quickest detection in multiple on-off processes," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 5994–6006, 2010.
- [32] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3229–3235, 2008.
- [33] K. Cohen and A. Leshem, "Energy-efficient detection in wireless sensor networks using likelihood ratio and channel state information," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1671–1683, 2011.
- [34] E. Klimko and J. Yackel, "Optimal search strategies for Wiener processes," *Stochastic Processes and their Applications*, vol. 3, no. 1, pp. 19–33, 1975.
- [35] V. Dragalin, "A simple and effective scanning rule for a multi-channel system," *Metrika*, vol. 43, no. 1, pp. 165–182, 1996.
- [36] L. D. Stone and J. A. Stanshine, "Optimal search using uninterrupted contact investigation," *SIAM Journal on Applied Mathematics*, vol. 20, no. 2, pp. 241–263, 1971.
- [37] A. Wald, "Sequential analysis," *New York: Wiley*, 1947.
- [38] V. Draglin, A. G. Tartakovsky, and V. V. Veeravalli, "Multihypothesis sequential probability ratio tests - part i: Asymptotic optimality," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2448–2461, 1999.
- [39] A. E. Albert, "The sequential design of experiments for infinitely many states of nature," *The Annals of Mathematical Statistics*, pp. 774–799, 1961.

- [40] S. Bessler, "Theory and applications of the sequential design of experiments, k-actions and infinitely many experiments: Part I-Theory," *Tech. Rep. Applied Mathematics and Statistics Laboratories, Stanford University*, no. 55, 1960.
- [41] A. Deshmukh, S. Bhashyam, and V. V. Veeravalli, "Sequential controlled sensing for composite multihypothesis testing," *arXiv preprint arXiv:1910.12697*, 2019.
- [42] H. Li, "Restless watchdog: Selective quickest spectrum sensing in multichannel cognitive radio systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, 2009.
- [43] R. Caromi, Y. Xin, and L. Lai, "Fast multiband spectrum scanning for cognitive radio systems," *IEEE Transaction on Communications*, vol. 61, no. 1, pp. 63–75, 2013.
- [44] M. L. Malloy and R. D. Nowak, "Sequential testing for sparse recovery," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7862–7873, 2014.
- [45] J. Heydari, A. Tajer, and H. V. Poor, "Quickest linear search over correlated sequences," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5786–5808, 2016.
- [46] K. Leahy and M. Schwager, "Always choose second best: Tracking a moving target on a graph with a noisy binary sensor," in *European Control Conference (ECC)*, 2016, pp. 1715–1721, 2016.
- [47] J. Heydari, A. Tajer, and H. V. Poor, "Quickest detection of markov networks," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 1341–1345, 2016, *arXiv version: arXiv:1711.04268v2*, 2017.
- [48] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for hypothesis testing," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5277–5280, IEEE, 2012.
- [49] T. L. Lai and L. M. Zhang, "Nearly optimal generalized sequential likelihood ratio tests in multivariate exponential families," *Lecture Notes-Monograph Series*, pp. 331–346, 1994.
- [50] H. Robbins and D. Siegmund, "The expected sample size of some tests of power one," *The Annals of Statistics*, pp. 415–436, 1974.
- [51] H. Robbins and D. Siegmund, "A class of stopping rules for testing parametric hypotheses," in *Proc. Sixth Berkeley Symp. on Mathematical Statistics and Probability (Univ. Calif. Press, Berkeley, CA, 1972)*, pp. 37–41, 1972.
- [52] H. Robbins and D. Siegmund, "The expected sample size of some tests of power one," *The Annals of Statistics*, pp. 415–436, 1974.
- [53] "DARPA Intrusion Detection Data Sets," *available at <https://www.ll.mit.edu/ideval/data/>*.
- [54] J. Fu, G. Li, and D. Zhao, "On large deviation expansion of distribution of maximum likelihood estimator and its application in large sample estimation," *Annals of the Institute of Statistical Mathematics*, vol. 45, no. 3, pp. 477–498, 1993.
- [55] A. G. Tartakovsky, "An efficient adaptive sequential procedure for detecting targets," in *Aerospace Conference Proceedings, 2002. IEEE*, vol. 4, pp. 4–4, IEEE, 2002.