# Finite-horizon Quickest Search in Correlated High-dimensional Data

Saeid Balaneshin
ECE Department
Wayne State University

Ali Tajer
ECE Department
Wayne State University

H. Vincent Poor
EE Department
Princeton University

***Abstract*—The problem of searching over a large number of data streams for identifying one that holds certain features of interest is considered. The data streams are assumed to be generated by one of two possible statistical distributions with cumulative distribution functions $F_0$ and $F_1$ and the objective is to identify one sequence generated by $F_1$ as quickly as possible, and prior to a pre-specified deadline. Furthermore, it is assumed that the generation of the data streams follows a known dependency kernel such that the likelihood of a sequence being generated by $F_1$ depends on the underlying distributions of the other data streams. The optimal sequential sampling strategy is characterized, and numerical evaluations are provided to illustrate the gains of incorporating the information about the dependency structure into the design of the sampling process.**

## I. Introduction

Recent advances in complex networks (e.g., energy grids) and information acquisition technologies (e.g., sensor networks) have led to the advent of high-dimensional data sets, which in turn propels the needs for computationally affordable and fast information processing mechanisms. Challenges associated with high-dimensional data analysis are multi-faceted and include modeling, communicating, storing, and searching, to name a few. Searching for features or anomalies in high-dimensional data is ubiquitous and has a central role in many domains, which often manifests itself as finding features (e.g., medical records [1]), detecting anomalies (e.g., fraud detection [2]), or identifying opportunities (e.g., spectrum sensing [3]–[5]).

In this paper we consider the problem of searching for features in high-dimensional data. Specifically, we consider a dataset consisting of a large number of data streams which are constantly generating information and aim to identify one of the data streams that holds a feature of interest as quickly as possible. In order to harness the complexity and quality of the decision, we focus on a sequential and data-adaptive information gathering process in which measurements are taken one at a time and it is dynamically decided whether to form a decision based on the information accumulated or to proceed with collecting more information. This problem when data streams are generated by one the two possible underlying mechanisms is studied in detail in [6], where it is assumed that all data streams, independently of each other, have equal chances of carrying the feature of interest. In this paper, we focus on a similar setting with the exception that we assume that there exists an inherent dependency among the generation of the data streams such that whether a data stream will bear the feature of interest depends on the rest of the data streams. Furthermore, we focus the attention on a finite-horizon quickest search strategy, in which the decision is delay-limited and has to be declared prior to a pre-specified deadline.

This research was supported in part by the U. S. National Science Foundation under Grants ECCS-1343326, DMS-1118605, and ECCS-1343210, and in part by the Marie Curie Outgoing Fellowship Program under Award No. FP7-PEOPLE-IOF-2011-298532.

The remainder of this paper is organized as follows. In Section II the quickest search problem and the underlying data generation model is formulated and Section III provides the optimal finite-horizon sampling strategy. Numerical evaluations are presented in Section IV and Section V concludes the paper.

## II. Problem Formulation

### A. Observation Model

Consider $n$ sequences of real-valued observations, denoted by $\{\mathcal{X}^i\}_{i=1}^n$, and define $X_j^i$ as the $j$th element of sequence $\mathcal{X}^i$, i.e.,

$$\mathcal{X}^i \triangleq \{X_1^i, X_2^i, \dots\} \ . \tag{1}$$

The elements within each sequence are independent and identically distributed (i.i.d.). The observations from the normal sequences are generated by a distribution with cumulative distribution function (CDF) $F_0$ and those from the outlier ones are generated by a different distribution with CDF $F_1$. By using this dichotomous model, the observations obey the following hypothesis model:

$$\begin{array}{ll} \mathsf{H}_0: & X_j^i \sim F_0, \\ \mathsf{H}_1: & X_j^i \sim F_1. \end{array} \tag{2}$$

The probability density functions corresponding to $F_0$ and $F_1$, are denoted by $f_0$ and $f_1$, respectively.

### B. Correlation Structure

Generation of sequences $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n\}$ is assumed to follow a known dependency kernel such that the prior probability that sequence $\mathcal{X}_i$ obeys $\mathsf{H}_1$ is controlled by the distribution of its proceeding sequence. Specifically, we have

$$\mathbb{P}(\mathsf{T}_1 = \mathsf{H}_1) \ = \ \epsilon \ , \tag{3}$$

and

$$\mathbb{P}(\mathsf{T}_i = \mathsf{H}_1 \mid \mathsf{T}_{i-1} = \mathsf{H}_j) \ = \ \epsilon_j \ , \quad \text{for } j \in \{0,1\} \ , \tag{4}$$

where $\epsilon_0 \neq \epsilon_1$. The prior probability that $\mathcal{X}_i$, for $i \in \{1,\dots,n\}$, is generated by $F_1$ can be obtained using (3) and (4), and it is

$$\mathbb{P}(\mathsf{T}_i = \mathsf{H}_1) = \epsilon \cdot (\epsilon_1 - \epsilon_0)^{i-1} + \epsilon_0 \cdot \frac{1-(\epsilon_1-\epsilon_0)^{i-1}}{1-(\epsilon_1-\epsilon_0)} \ . \tag{5}$$

### *C. Sampling Model*

In order to identify *one* sequence generated by $F_1$ as quickly as possible, the sampling procedure collects observations sequentially until a reliable decision can be formed. By denoting the sample taken at time $t$ by $Y_t$, and by defining $s_t$ as the index of the sequence observed at time $t \in \mathbb{N}$, the sampling procedure is initiated at time $t = 1$ by setting $s_1 = 1$. Depending on the information accumulated up to time $t$, the sampling procedure takes one of the following three possible actions:

- $\text{A}_1$) *Detection:* stop taking further samples due to having enough confidence to declare that the sequence $s_t$ is generated according to $F_1$.
- $\text{A}_2$) *Observation:* take one more sample from the same sequence at time $t+1$ due to lack of confidence to make a decision. Hence, we have $s_{t+1} = s_t$.
- $\text{A}_3$) *Exploration:* discard sequence $s_t$ due to having enough confidence that this sequence is generated by $F_0$ and switch to its following sequence, i.e., we have $s_{t+1} = s_t + 1$.

In order to formalize the sampling procedure, we define $\tau$ as the stopping time, at which the sampling procedure is terminated and action $\text{A}_1$ is taken. At time $t \in \{1,\dots,\tau-1\}$, in order to determine which of the two actions $\text{A}_2$ or $\text{A}_3$ should be performed, we define the switching function $\psi : \{1,\dots,\tau-1\} \rightarrow \{0,1\}$, where $\psi(t) = 0$ indicates that we should continue taking one more sample from the current sequence, and $\psi(t) = 1$ indicates that the current sequence should be discarded and a sample from the next sequence should be taken, i.e.,

$$\psi(t) = \begin{cases} 0 & \text{action A}_2 \text{ and } s_{t+1} = s_t \\ 1 & \text{action A}_3 \text{ and } s_{t+1} = s_t + 1 \end{cases} \ . \tag{6}$$

### *D. Problem Formulation*

The optimal sampling procedure can be characterized uniquely by finding the optimal stopping time and switching sequence. In such sequential and data-adaptive sampling procedures there exists an inherent tension between decision quality and decision delay. Hence designing the optimal sampling strategy involves striking a balance between these two performance measures. In this paper, we focus on minimizing the expected decision delay subject to a controlled decision quality by solving

$$\begin{array}{ll} \inf_{\tau \in \mathcal{T}, \psi_\tau} & \mathbb{E}[\tau] \\ \text{s.t.} & \mathbb{P}(\mathsf{T}_{s_\tau} = \mathsf{H}_0) \leq \beta \end{array} \ . \tag{7}$$

And, in particular, the focus is on designing a finite-horizon sampling procedure in which a decision should be performed by a pre-specified time $T \in \mathbb{N}$.

## III. Optimal Finite-horizon Sampling

By using the discussions in [6] and [7], the solution for the canonical quickest search optimization problem formulated in (7) can be equivalently obtained by solving the following Bayesian formulation:

$$\inf_{\tau \in \mathcal{T}, \psi_\tau} \left[\mathbb{P}(\mathsf{T}_{s_\tau} = \mathsf{H}_0) + c_\beta \mathbb{E}[\tau]\right] \ , \tag{8}$$

where $c_\beta$ is a constant (function of $\beta$) through which the costs associated with decision delay and accuracy are integrated into one cost function. Next, given the information accumulated up to time $t$, we characterize the cost of the detection, observation, and exploration actions, based on which we can identify the best next action. By defining $\tilde{G}_t^T(\mathcal{F}_t)$ where $\mathcal{F}_t \triangleq \{Y_1,\dots,Y_t\}$, as the cost of the best action at time $t$, we have

$$\tilde{G}_t^T(\mathcal{F}_t) \triangleq \min\{\tilde{J}_{t;1}^T(\mathcal{F}_t), c_\beta + \min_{i=2,3}\{\tilde{J}_{t;i}^T(\mathcal{F}_t)\}\} \ , \tag{9}$$

where $\tilde{J}_{t;1}^T(\mathcal{F}_t)$, $c_\beta + J_{t;2}^T(\mathcal{F}_t)$, and $c_\beta + J_{t;3}^T(\mathcal{F}_t)$ are the costs pertinent to detection, observation, and exploration, respectively. Furthermore, we have the following recursive connections between $\{J_{t;i}^T(\mathcal{F}_t)\}_{i=1}^3$ and $\tilde{G}_t^T(\mathcal{F}_t)$:

$$\text{A}_1: \quad \tilde{J}_{t;1}^T(\mathcal{F}_t) \ = \ 1 - \pi_t \tag{10}$$

$$\text{A}_2: \quad \tilde{J}_{t;2}^T(\mathcal{F}_t) \ = \ \mathbb{E}[\tilde{G}_{t+1}^T(\mathcal{F}_{t+1}) | \mathcal{F}_t, \psi(t) = 0] \tag{11}$$

$$\text{A}_3: \quad \tilde{J}_{t;3}^T(\mathcal{F}_t) \ = \ \mathbb{E}[\tilde{G}_{t+1}^T(\mathcal{F}_{t+1}) | \mathcal{F}_t, \psi(t) = 1] \tag{12}$$

To proceed, let us define $\pi_t$ as the posterior probability that the sequence observed at time $t$, i.e., sequence $s_t$, is generated according to $F_1$. Hence, we have

$$\pi_t \triangleq \mathbb{P}(\mathsf{T}_{s_t} = \mathsf{H}_1 \mid \mathcal{F}_t) \ . \tag{13}$$

By using the correlation structure described in Subsection II-B, the posterior probability at time $t+1$, for $t \in \mathbb{N}$, can be obtained recursively in terms of $\pi_t$ according to

$$\begin{aligned} \pi_{t+1} & \\ = & \ \frac{\pi_t f_1(Y_{t+1})}{\pi_t f_1(Y_{t+1}) + (1-\pi_t) f_0(Y_{t+1})} \cdot \mathbb{I}(\psi(t) = 0) \\ + & \ \frac{\bar{\pi}_t f_1(Y_{t+1})}{\bar{\pi}_t f_1(Y_{t+1}) + (1-\bar{\pi}_t) f_0(Y_{t+1})} \cdot \mathbb{I}(\psi(t) = 1) \ , \end{aligned} \tag{14}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function, and we have defined

$$\bar{\pi}_t \triangleq \pi_t(\epsilon_1 - \epsilon_0) + \epsilon_0 \ . \tag{15}$$

Furthermore, for $t = 1$, we have

$$\pi_1 \ = \ \frac{\epsilon f_1(Y_1)}{\epsilon f_1(Y_1) + (1-\epsilon) f_0(Y_1)} \ . \tag{16}$$

By taking into account the recursive form in (14), and by following the same line of argument as in [6], the following lemma shows that the optimal decision rule for stopping or switching at time $t$ is related to $\mathcal{F}_t$ only through $\pi_t$.

*Lemma 1:* The functions $\{\tilde{J}^T_{t;i}(\mathcal{F}_t)\}_{i=1}^3$ depend on $\mathcal{F}_t$ only through $\pi_t$ and can be rewritten as functions of $\pi_t$ which we denote by $\{J^T_{t;i}(\pi_t)\}_{i=1}^3$.

*Proof:* By definition we have $\tilde{J}^T_{t;1}(\mathcal{F}_t) = 1-\pi_t$, which establishes the claim for $i=1$. Furthermore, at time $T$, the minimal cost function $\tilde{G}^T_T(\mathcal{F}_T)$ is dominated by that of the detection action, and we immediately have $\tilde{G}^T_T(\mathcal{F}_T)=1-\pi_T$. By using backward induction with starting point $t=T$, we next show that if $\tilde{G}^T_{t+1}(\mathcal{F}_{t+1})$ depends on $\mathcal{F}_{t+1}$ only through $\pi_{t+1}$, then the functions $\{\tilde{J}^T_{t;i}(\mathcal{F}_t)\}_{i=2}^3$ and $\tilde{G}^T_t(\mathcal{F}_t)$ depend on $\mathcal{F}_t$ only through $\pi_t$. For this purpose, we denote $\tilde{G}^T_{t+1}(\mathcal{F}_{t+1})$ by $G^T_{t+1}(\pi_{t+1})$, based on which from (11) and (14) we obtain

$$\begin{aligned}\tilde{J}^T_{t;2}(\mathcal{F}_t) &= \mathbb{E}[G^T_{t+1}(\pi_{t+1})|\mathcal{F}_t,\psi(t)=0]\\ &= \int G^T_{t+1}\left(\frac{f_1(Y_{t+1})\pi_t}{f_1(Y_{t+1})\pi_t+f_0(Y_{t+1})(1-\pi_t)}\right)\\ &\quad\times\left(f_1(Y_{t+1})\pi_t+f_0(Y_{t+1})(1-\pi_t)\right)\,dY_{t+1}\ . \end{aligned} \tag{17}$$

The integrand of (17) clearly depends on $\mathcal{F}_t$ only through $\pi_t$, which we denote by $J^T_{t;2}(\pi_t)$. Similarly, based on (12) and (14) we have

$$\begin{aligned}\tilde{J}^T_{t;3}(\mathcal{F}_t) &= \mathbb{E}[G^T_{t+1}(\pi_{t+1})|\mathcal{F}_t,\psi(t)=1]\ ,\\ &= \int G^T_{t+1}\left(\frac{f_1(Y_{t+1})\bar{\pi}_t}{f_1(Y_{t+1})\bar{\pi}_t+f_0(Y_{t+1})(1-\bar{\pi}_t)}\right)\\ &\quad\times\left(f_1(Y_{t+1})\bar{\pi}_t+f_0(Y_{t+1})(1-\bar{\pi}_t)\right)\,dY_{t+1}\ . \end{aligned} \tag{18}$$

By recalling that $\bar{\pi}_t=\pi_t(\epsilon_1-\epsilon_0)+\epsilon_0$, it is concluded that $\tilde{J}^T_{t;3}(\mathcal{F}_t)$ depends on $\mathcal{F}_t$ through $\pi_t$, denoted by $J^T_{t;3}(\pi_t)$. Next by recalling that

$$\tilde{G}^T_t(\mathcal{F}_t)=\min\{\tilde{J}^T_{t;1}(\mathcal{F}_t),c_\beta+\min\{\tilde{J}^T_{t;2}(\mathcal{F}_t),\tilde{J}^T_{t;3}(\mathcal{F}_t)\}\}\ ,$$

we immediately have

$$\tilde{G}^T_t(\mathcal{F}_t)=\min\{J^T_{t;1}(\pi_t),c_\beta+\min\{J^T_{t;2}(\pi_t),J^T_{t;3}(\pi_t)\}\}\ ,$$

which indicates that $\tilde{G}^T_t(\mathcal{F}_t)$ depends on $\mathcal{F}_t$ through $\pi_t$, denoted by $G^T_t(\pi_t)$ ∎

By leveraging the result of Lemma 1, we next establish that functions $J^T_{t;2}(\pi_t)$ and $J^T_{t;3}(\pi_t)$ are concave in $\pi_t$.

*Lemma 2:* The functions $J^T_{t;2}(\pi_t)$ and $J^T_{t;3}(\pi_t)$ are non-negative concave functions of $\pi_t$ for $\pi_t\in[0,1]$.

*Proof:* Non-negativity of these functions follows from the fact that at the stopping time the minimal cost is $G^T_T(\pi_T)=1-\pi$, which is non-negative, in conjunction with the recursive connection between $\{J^T_{t;i}(\pi_t)\}_{i=1}^3$ and $G^T_{t+1}(\pi_{t+1})$ provided in (11) and (12).

The concavities of these functions can be established through backward induction by proving that the minimal cost function $G^T_{t+1}(\pi_{t+1})$ being concave leads to concave structures for $J^T_{t;2}(\pi_t)$ and $J^T_{t;3}(\pi_t)$, and consequently a concave structure for $G^T_t(\pi_t)$. For this purpose, at time $T$ we have $G^T_T(\pi_T)=1-\pi_T$, confirming that the function $G^T_t(\pi_T)$ is concave at the starting point of the inductive argument. By assuming that $G^T_{t+1}(\pi_{t+1})$ is concave, we next show that the functions $J^T_{t;2}(\pi_t)$ and $J^T_{t;3}(\pi_t)$ are concave. For this purpose, for any two arbitrary probability terms $\pi^1_t$ and $\pi^2_t$ we define $\pi^3_t$ as their convex combination for an arbitrary $\lambda\in[0,1]$, i.e.,

$$\pi^3_t \triangleq \lambda\pi^1_t+(1-\lambda)\pi^2_t\ , \tag{19}$$

and aim to show that for $i\in\{2,3\}$

$$\lambda J^T_{t;i}(\pi^1_t)+(1-\lambda)J^T_{t;i}(\pi^2_t)\ \geq\ J^T_{t;i}(\pi^3_t)\ . \tag{20}$$

By using (17)-(18) and expanding the left hand of (20) we obtain

$$\begin{aligned}&\lambda J^T_{t;i}(\pi^1_t)+(1-\lambda)J^T_{t;i}(\pi^2_t)\\ &\quad=\int[\mu_i G^T_{t+1}(\pi^1_{t+1})+(1-\mu_i)G^T_{t+1}(\pi^2_{t+1})]\\ &\qquad\times[\lambda q_i(\pi^1_t)+(1-\lambda)q_i(\pi^2_t)]\,dY_{t+1}\ ,\end{aligned} \tag{21}$$

where we have defined

$$\mu_i \triangleq \frac{\lambda q_i(\pi^1_t)}{\lambda q_i(\pi^1_t)+(1-\lambda)q_i(\pi^2_t)}\ , \tag{22}$$

$$q_2(\pi)\triangleq \pi f_1(Y_{t+1})+(1-\pi)f_0(Y_{t+1})\ , \tag{23}$$

$$\begin{aligned}\text{and,}\ \ q_3(\pi) &\triangleq (\pi(\epsilon_1-\epsilon_0)+\epsilon_0)f_1(Y_{t+1})\\ &\quad+[1-\pi(\epsilon_1-\epsilon_0)-\epsilon_0]f_0(Y_{t+1})\ .\end{aligned} \tag{24}$$

By using the concavity of $G^T_{t+1}(\pi_{t+1})$ we obtain the following lower bound on (21):

$$\begin{aligned}&\lambda J^T_{t;i}(\pi^1_t)+(1-\lambda)J^T_{t;i}(\pi^2_t)\\ &\quad\geq\int G^T_{t+1}(\mu_i\pi^1_{t+1}+(1-\mu_i)\pi^2_{t+1})\\ &\qquad\times[\lambda q_i(\pi^1_t)+(1-\lambda)q_i(\pi^2_t)]\,dY_{t+1}\ .\end{aligned} \tag{25}$$

Next, we remark that it can be readily verified that

$$\pi^3_{t+1}=\mu_i\pi^1_{t+1}+(1-\mu_i)\pi^2_{t+1}\ , \tag{26}$$

$$\text{and ,}\quad q_i(\pi^3_t)=\lambda q_i(\pi^1_t)+(1-\lambda)q_i(\pi^2_t)\ . \tag{27}$$

Hence, we can re-write (25) as

$$\begin{aligned}&\lambda J^T_{t;i}(\pi^1_t)+(1-\lambda)J^T_{t;i}(\pi^2_t)\\ &\quad\geq\int G^T_{t+1}(\pi^3_{t+1})q_i(\pi^3_t)\,dY_{t+1}=J^T_{t;i}(\pi^3_t)\ ,\end{aligned} \tag{28}$$

which proves the concavity of $\{J^T_{t;i}(\pi_t)\}_{i=2}^3$ in $\pi_t$. ∎

Given lemmas above, the following theorems establish the optimal stopping time and the switching rule at time $t$ as functions of $\pi_t$.

*Theorem 1 (Stopping Time):* For the finite-horizon quickest search problem in (8), the optimal stopping time is $\tau^*_T = \inf\{t\ :\ \pi_t\geq\pi^T_U\}$ where $\pi^T_U$ is a solution of

$$1-\pi^T_U\ =\ c_\beta+\min_{i=2,3}J^T_{t;i}(\pi^T_U)\ . \tag{29}$$

*Proof:* According to (9), the procedure stops taking further samples when the cost associated with terminating the procedure falls below those associated with the observation and exploration actions, i.e., $J^T_{t;1}(\mathcal{F}_t)$ becomes less than $c_\beta+J^T_{t;2}(\pi_t)$ or $c_\beta+J^T_{t;3}(\pi_t)$. In other words,

$$\tau^*_T=\inf\{t\ :\ \pi_t\geq 1-c_\beta-\min_{i=2,3}J^T_{t;i}(\pi_t)\}\ . \tag{30}$$

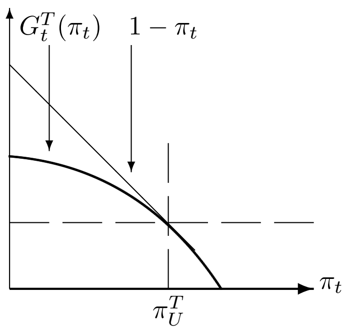


Fig. 1. An illustration of $G_t^T(\pi_t)$.

This characterization of the stopping time implies that the stopping rule can be equivalently cast as comparing the probability term $\pi_t$ with a threshold. By taking into account that functions $\{J_{t;i}^T(\pi_t)\}_{i=2}^3$ are concave in $\pi_t$, the relative structure of the minimal dynamic cost $G_t^T(\pi)$ and stopping cost function $1-\pi_t$ is depicted in Fig. 1. Based on this figure, and the optimal form of the stopping time given in (29), the optimal stopping rule can be equivalently stated as in Theorem 1 with $\pi_U^T$ marked in Fig. 1 and defined in (29). ■

*Theorem 2 (Switching Rule):* For the finite-horizon quickest search problem in (8), the optimal switching rule at time $t$ is to switch to a new sequence if and only if $J_{t;2}^T(\pi_t) > J_{t;3}^T(\pi_t)$.

*Proof:* According to (9), the exploration action is performed when the cost associated with switching to a new sequence falls below the one associated with the observation action, i.e., when $c_\beta + J_{t;3}^T(\pi_t)$ becomes less than $c_\beta + J_{t;2}^T(\pi_t)$. ■

## IV. Numerical Evaluations

In this section we present numerical results by placing the central focus on comparing the performance of the quickest search procedure in this paper, which takes into account the correlation structure, and that of the one in [6] that does not take into account the correlation structure. We consider the setting $\epsilon = 0.4$, $\epsilon_0 = 0.1$, $\epsilon_1 = 0.9$, and target at controlling the error probability below $\beta = 0.01$. It is also assumed that $F_0$ and $F_1$ are zero-mean Gaussian with variances $\sigma_0^2$ and $\sigma_1^2$, respectively. In Figures 2 and 3, the expected stopping time and error probability are, respectively, shown versus $\sigma_1^2/\sigma_0^2$. It is observed that while in both settings, the error probabilities are kept below the desired value $\beta$, ignoring the correlation structure leads to increased delay in reaching a decision, and therefore, the corresponding sampling strategy is suboptimal to the one characterized in this paper.

## V. Conclusions

We have characterized an optimal sequential sampling strategy for finite-horizon quickest search over a large number of correlated data streams that are generated by one of two possible statistical distributions. The generation of the data streams follows a pre-specified correlation structure in which the prior probability that a sequence is generated according to the distribution of interest is governed by the distribution of its preceding data stream. The proposed sampling strategy guarantees achieving the smallest expected delay in reaching a decision while an upper bound is enforced on the rate of erroneous decisions. The gains of this strategy over an existing one that does not take into consideration the correlation structure have been assessed numerically.

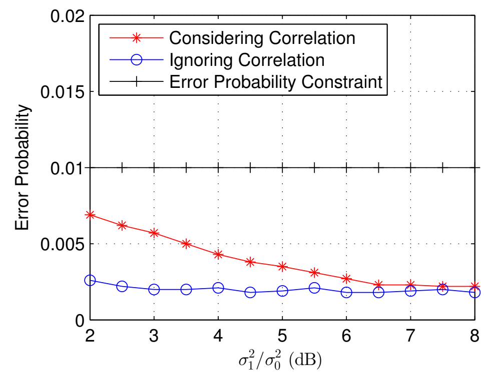


Fig. 2. Error probability versus $\sigma_1^2/\sigma_0^2$.

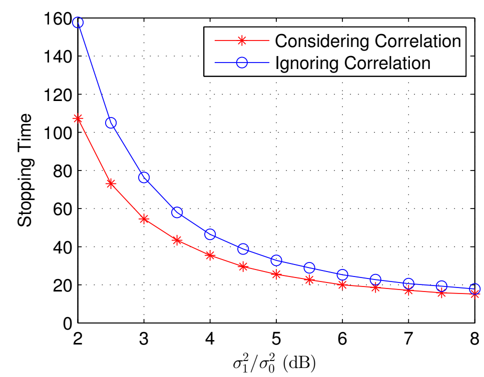


Fig. 3. Stopping time versus SNR.

## References


[1] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 677–688, Aug. 2001.

[2] R. J. Bolten and D. J. Hand, "Unsupervised profiling methods for fraud detection," in *Proc. Conference of Credit Scoring and Credit Control*, vol. VII, (Edingburgh, UK), 2001.

[3] A. Tajer, R. Castro, and X. Wang, "Adaptive sensing of congested spectrum bands," *IEEE Transactions on Information Theory*, vol. 58, pp. 6110– 6125, Sep. 2012.

[4] Y. Yilmaz, G. V. Moustakides, and X. Wang, "Cooperative sequential spectrum sensing based on level-triggered sampling," *IEEE Transactions on Signal Processing*, vol. 60, pp. 4509–4524, Sep. 2012.

[5] J. Lundén, S. A. Kassam, and V. Koivunen, "Robust nonparametric cyclic correlation-based spectrum sensing for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 58, pp. 38–52, Jan. 2010.

[6] L. Lai, H. V. Poor, Y. Xin, and G. Georgiadis, "Quickest search over multiple sequences," *IEEE Transactions on Information Theory*, vol. 57, pp. 5375–5386, Aug. 2011.

[7] A. N. Shiryaev, *Optimal Stopping Rules*, vol. 8. Springer, Nov. 2007.