BOOTSTRAP BASED SEQUENTIAL PROBABILITY RATIO TESTS

Fiky Y. Suratman and Abdelhak M. Zoubir

Signal Processing Group,Technische Universität Darmstadt
Merckstrasse 25, 64283, Darmstadt, Germany

ABSTRACT

We present a generalized sequential probability ratio test for composite hypotheses wherein the thresholds are updated in an adaptive manner based on the data recorded up to the current sample using the parametric bootstrap. The resulting test avoids the asymptotic assumption usually made in earlier works. The increase of the average sample number of the proposed method is not significant compared to the sequential probability ratio test which is based on known parameters, especially in a low SNR region. In addition, the probability of false alarm and the probability of missed detection are maintained below the preset values. A comparison shows that the thresholds based on the parametric bootstrap are in close agreement with the thresholds based on Monte-Carlo simulations.

Index Terms— Bootstrap, sequential probability ratio test, composite hypothesis, spectrum sensing, cognitive radio

1. INTRODUCTION

The sequential probability ratio test (SPRT) was pioneered by Abraham Wald in 1943 in a confidential report, which was then published in 1945 [1]. His initial work on the general theory of the cumulative sum [2] has provided a tool to derive the operating characteristic and the average sample number as the performance indices of the SPRT. The main idea behind the SPRT is to pre-specify the probability of detection errors (false alarm and missed detection) and minimize the average of required sample number [1, 3]. This scheme results in a random sample number with its average smaller than the corresponding best fixed sample number detector. The work of Bussgang and Middleton [4] applies Wald's SPRT to sequentially detect signals in noise. The case of correlated observation is also treated. Subsequently, some analytical results and modified procedures on the SPRT also have been introduced, such as in [5–7]. In addition, since the SPRT has an average sample number (ASN) larger than the corresponding fixed sample number detector when the parameter value lies between the two simple hypotheses, some methods to reduce the sample number in this condition have been proposed [8,9]. They are known as truncated SPRT. Implementation of sequential analysis to a distributed environment can be found, for example, in [10, 11].

Recently, the superiority of the sequential detector in reducing the sample number is an attractive choice in spectrum sensing for cognitive radio (CR) [12]. This is because a short duration of sensing time (a small sample number) is required in spectrum sensing to improve agility of the CR users, which successively increase the throughput of a CR network and reduce the degree of interference to the licensed network. Some examples of the implementation in this emerging research field are [13–17]. However, addressing sequential detections for composite hypotheses is only few in literature.

The bootstrap, a computer-based method for assigning measures of accuracy to statistical estimates [18], is used here. From a data manipulation point of view, the main idea encapsulated by the bootstrap is to simulate as much of the real world probability mechanism as possible, in which any unknowns are replaced with estimates from the observed data. The application of the bootsrap to various signal processing techniques can be found in [19, 20] and references therein. In addition, bootstrap-based detection methods have been successfuly implemented in [21–23]. However, the literature on bootstrapping methods for sequential detections is very scarce.

Some studies in [24, 25] on sequential detection for composite hypotheses have focused on asymptotic optimality and properties for the general non-iid case. In this setup, the probability of false alarm and the probability of missed detection are assumed to converge to zero or the constant thresholds are assumed to converge to infinity. The present work, in some sense, is based on the same framework as [16]. However, the authors in [16] mainly concentrate on finding a bound on the error term of the log-likelihood ratio in the asymptotic sense, such that the threshold margin for the upper and the lower thresholds, which are functions of the sample number, can be replaced by some positive constants. Eventually, constant thresholds are used in this earlier work. In general, the present paper differs from the previous work in several aspects. *First*, the SPRT is extended to cope with the composite hypothesis problem by combining the generalized SPRT and the parametric bootstrap to find the thresholds. *Second*, the thresholds are not constant but updated for each new sample recorded and also suitable for small sample number, not only in the asymptotic sense. Hence, a smaller average sample number is possible. *Last*, exhaustive simulations to determine the thresholds in worst case scenario such as in [16] are avoided.

The remainder of the paper is organized as follows. The SPRT for simple hypothesis is briefly explained in Section 2. Section 3 begins with signal model and discusses subsequently in detail the generalized SPRT and the proposed bootstrap based SPRT for composite hypothesis. An example and simulation results are presented in Section 4. Finally, the paper is concluded with Section 5.

2. THE SIMPLE HYPOTHESIS CASE

The sequential probability ratio test (SPRT) for binary hypothesis testing is designed based on simple hypothesis $\mathcal{H}_0$ against simple alternative $\mathcal{H}_1$ [1,3]. Let $f(\vec{x}_N)$ denotes the density function of a random column vector $\vec{x}_N = [x(1)\ x(2)\cdots x(N)]^T$ which represents the iid observations of a signal up to the Nth sample. The density function $f_0(\vec{x}_N)$ under $\mathcal{H}_0$ and $f_1(\vec{x}_N)$ under $\mathcal{H}_1$ are assumed to be exactly known. Then, the SPRT is optimum for minimizing the sample number for targeted probability of false alarm α and probability of missed detection β among other tests [3]. The required sample number is not determined in advance, but it is random, depending on the statistics of the current observation. The test stops when either the upper constant-threshold A or the lower B is crossed for the first

time. In summary, the SPRT can be written as

$$\bar{Z}_N = \log \frac{f_1(\vec{x}_N)}{f_0(\vec{x}_N)} \begin{cases} \geq A, & \text{accept } \mathcal{H}_1 \\ \leq B, & \text{accept } \mathcal{H}_0, \\ A < \bar{Z}_N < B, & \text{continue to observe} \end{cases} \quad (1)$$

For the preset probability of false alarm α and missed detection β that are much smaller than 1, the thresholds can be calculated from Wald's approximation [3]

$$A = \log \frac{1}{\alpha}, \quad B = \log \beta, \quad (2)$$

to guarantee that the actual probability of false alarm $P_f \leq \alpha$ and missed detection $P_m \leq \beta$.

3. THE COMPOSITE HYPOTHESIS CASE

The simple hypothesis approach is not suitable in most practical scenarios, since unknown parameters might exist which enforce to use non-exact distributions. Therefore, we aim at extending the SPRT to the composite hypothesis case. Suppose that the detector receives a signal that is distributed according to

$$\begin{aligned} \mathcal{H}_0 : \vec{x}_N &\sim f_0(\vec{x}_N; \theta_0), \ \theta_0 \in \Theta_0 \\ \mathcal{H}_1 : \vec{x}_N &\sim f_1(\vec{x}_N; \theta_1), \ \theta_1 \in \Theta_1 \end{aligned} \quad (3)$$

where θ_0 and θ_1 are the unknown parameters in the respective hypothesis, and $\Theta_0, \Theta_1 \subset \Theta$ are disjoint parameter spaces, i.e. $\Theta_0 \cap \Theta_1 = \emptyset$.

3.1. The Generalized SPRT

Note that the two parameters θ_0 and θ_1 in (3) are unknown. Hence, we replace them by their maximum likelihood estimates (MLEs). As in the case of a fixed sample number detector [26], the generalized log-likelihood ratio (GLR) can then be written as

$$\hat{Z}_N = \log \frac{f_1(\vec{x}_N; \hat{\theta}_1^{(N)})}{f_0(\vec{x}_N; \hat{\theta}_0^{(N)})} \quad (4)$$

where

$$\hat{\theta}_i^{(N)} = \arg\max_{\theta_i \in \Theta_i} \log \left(f_i(\vec{x}_N; \theta_i) \right) \quad (5)$$

is the MLE assuming hypothesis $\mathcal{H}_i, i = 0, 1$ is true. Meanwhile, subscript or superscript N is to indicate the data has been recorded up to the Nth sample.

Furthermore, the two thresholds of the SPRT (2) cannot be used directly in this setup. A threshold modification is required to compensate the estimation errors, introduced by the MLEs. Otherwise, it will lead to early terminations that increase the actual probability of false alarm P_f and missed detection P_m. Note that we cannot just increase A or decrease B by some constants to improve P_f and P_m, since it will successively increase the average sample number. Meanwhile, the average of estimation errors is large for a small N and vice versa. Hence, to reflect this condition, an appropriate method is needed to set the thresholds adaptively based on N. In this way, they should be large to tolerate large estimation errors for N small and converge to (2) as N increases.

The generalized SPRT for composite hypotheses can now be defined as [16]

$$\hat{Z}_N \begin{cases} \geq A_N, & \text{accept } \mathcal{H}_1 \\ \leq B_N, & \text{accept } \mathcal{H}_0, \\ A_N < \hat{Z}_N < B_N, & \text{continue to observe} \end{cases} \quad (6)$$

where $\hat{Z}_N$ is the GLR from (4). The upper threshold A_N and the lower B_N are functions of the sample number N,

$$A_N = A + \Delta_N^0, \quad B_N = B - \Delta_N^1, \quad (7)$$

where A and B are determined from (2), and $\Delta_N^i \geq 0, i = 0, 1$, is the modified value for the threshold at the Nth stage under $\mathcal{H}_i$. The determination of Δ_N^i will be elaborated in the following subsection.

3.2. The Bootstrap Based SPRT (B-SPRT)

The log-likelihood ratio $\hat{Z}_N$ in (6) can be considered as an estimate of the log-likelihood ratio $\bar{Z}_N^i$ in (1) under hypothesis $\mathcal{H}_i, i = 0, 1$,

$$\begin{aligned} \mathcal{H}_0 : \hat{Z}_N &= \log \frac{f_1(\vec{x}_N; \bar{\theta}_1)}{f_0(\vec{x}_N; \theta_0)} + \log \left(\frac{f_1(\vec{x}_N; \hat{\theta}_1^{(N)})}{f_1(\vec{x}_N; \bar{\theta}_1)} \frac{f_0(\vec{x}_N; \theta_0)}{f_0(\vec{x}_N; \hat{\theta}_0^{(N)})} \right) \\ &= \bar{Z}_N^0 + \Delta Z_N^0 \\ \mathcal{H}_1 : \hat{Z}_N &= \log \frac{f_1(\vec{x}_N; \theta_1)}{f_0(\vec{x}_N; \bar{\theta}_0)} + \log \left(\frac{f_1(\vec{x}_N; \hat{\theta}_1^{(N)})}{f_1(\vec{x}_N; \theta_1)} \frac{f_0(\vec{x}_N; \bar{\theta}_0)}{f_0(\vec{x}_N; \hat{\theta}_0^{(N)})} \right) \\ &= \bar{Z}_N^1 + \Delta Z_N^1, \end{aligned} \quad (8)$$

where ΔZ_N^i is the error term under $\mathcal{H}_i, i = 0, 1$. The MLEs

$$\begin{aligned} \mathcal{H}_0 : \hat{\theta}_0^{(N)} &\xrightarrow{\text{P}} \theta_0 \quad \hat{\theta}_1^{(N)} \xrightarrow{\text{P}} \bar{\theta}_1, \quad \theta_0 \in \Theta_0, \bar{\theta}_1 \in \Theta_1 \\ \mathcal{H}_1 : \hat{\theta}_0^{(N)} &\xrightarrow{\text{P}} \bar{\theta}_0 \quad \hat{\theta}_1^{(N)} \xrightarrow{\text{P}} \theta_1, \quad \bar{\theta}_0 \in \Theta_0, \theta_1 \in \Theta_1 \end{aligned} \quad (9)$$

in which

$$\bar{\theta}_i = \arg\min_{\check{\theta}_i \in \Theta_i} KL(f_j(\vec{x}_N; \theta_j) || f_i(\vec{x}_N; \check{\theta}_i)), \ i, j = 0, 1, i \neq j, \ (10)$$

and $KL(f_j || f_i)$ denotes Kullback-Leibler distance between distributions f_j and f_i. Here $\xrightarrow{\text{P}}$ denotes convergence in probability as $N \to \infty$. At stopping time $N = N_T$, the probability of false alarm P_f and missed detection P_m can then be written as

$$\begin{aligned} P_f &= P(\hat{Z}_{N_T} \geq A_{N_T}, \bar{Z}_{N_T}^0 \geq A | \mathcal{H}_0) \\ &\quad + P(\hat{Z}_{N_T} \geq A_{N_T}, \bar{Z}_{N_T}^0 < A | \mathcal{H}_0) \\ &\leq P(\bar{Z}_{N_T}^0 \geq A | \mathcal{H}_0) + P(\hat{Z}_{N_T} \geq A_{N_T}, \bar{Z}_{N_T}^0 < A | \mathcal{H}_0) \\ &\leq \alpha + \Delta P_f \quad (11) \\ P_m &= P(\hat{Z}_{N_T} \leq B_{N_T}, \bar{Z}_{N_T}^1 \leq B | \mathcal{H}_1) \\ &\quad + P(\hat{Z}_{N_T} \leq B_{N_T}, \bar{Z}_{N_T}^1 > B | \mathcal{H}_1) \\ &\leq P(\bar{Z}_{N_T}^1 \leq B | \mathcal{H}_1) + P(\hat{Z}_{N_T} \leq B_{N_T}, \bar{Z}_{N_T}^1 > B | \mathcal{H}_1) \\ &\leq \beta + \Delta P_m. \quad (12) \end{aligned}$$

Hence, we aim at minimizing ΔP_f and ΔP_m by finding appropriate values for Δ_N^0 and Δ_N^1 of (7), respectively. Note that the thresholds should be calculated at each Nth stage since the stopping time N_T is not known in advance. Therefore, let the last terms of (11) and (12) be expressed as functions of N

$$\begin{aligned} P(\hat{Z}_N &\geq A_N, \bar{Z}_N^0 < A | \mathcal{H}_0) \\ &\leq P(|\hat{Z}_N - \bar{Z}_N^0| \geq A_N - A | \mathcal{H}_0) \\ &= P(|\Delta Z_N^0| \geq \Delta_N^0 | \mathcal{H}_0) \quad (13) \\ P(\hat{Z}_N &\leq B_N, \bar{Z}_N^1 > B | \mathcal{H}_1) \\ &\leq P(|\hat{Z}_N - \bar{Z}_N^1| \geq B - B_N | \mathcal{H}_1) \\ &= P(|\Delta Z_N^1| \geq \Delta_N^1 | \mathcal{H}_1). \quad (14) \end{aligned}$$

Now, suppose that the two probabilities are pre-specified as tolerance values to represent ΔP_f and ΔP_m

$$P(|\Delta Z_N^0| \geq \Delta_N^0 | \mathcal{H}_0) = \epsilon_0, \quad P(|\Delta Z_N^1| \geq \Delta_N^1 | \mathcal{H}_1) = \epsilon_1. \quad (15)$$

Accordingly, $\Delta_N^i, i = 0, 1$, can be found by using the parametric bootstrap of (15) to finally get (A_N, B_N) in (7).

In our setup, the thresholds at the $(N + 1)$th stage is calculated based on the observed data up to the Nth sample. Hence, upon receiving the $(N + 1)$th sample, the log-likelihood ratio $\hat{Z}_{N+1}$ is computed and then compared to the calculated thresholds (A_{N+1}, B_{N+1}). In principle, the parametric bootstrap replaces any unknown parameters with estimates and re-uses data to provide a way to approximate distributional information, using the measured sample as a basis. Suppose that N samples have been recorded and are represented by $\vec{x}_N$. Let F_θ be the distribution of $\vec{x}_N$, parameterized by θ. B_s bootstrap replications of size $N + 1$ are generated from parametric estimates of the population $\hat{F}_{\hat{\theta}_i^{(N)}}$ by assuming $\mathcal{H}_i$ is true, i.e.

$$\mathcal{H}_0 : \hat{F}_{\hat{\theta}_0^{(N)}} \rightarrow \vec{x}_{N+1}^{0b} = [x_1^{0b} \cdots x_{N+1}^{0b}]^T$$

$$\mathcal{H}_1 : \hat{F}_{\hat{\theta}_1^{(N)}} \rightarrow \vec{x}_{N+1}^{1b} = [x_1^{1b} \cdots x_{N+1}^{1b}]^T, \ b = 1, \cdots, B_s \quad (16)$$

parameterized by $\hat{\theta}_0^{(N)}$ and $\hat{\theta}_1^{(N)}$ which are the MLEs from (5). Note that we generate the bootstrap replications for $\mathcal{H}_0$ and $\mathcal{H}_1$ to get the upper threshold A_N and the lower threshold B_N, respectively. Furthermore, the log-likelihood ratios $\hat{Z}_{N+1}$ and $\bar{Z}_{N+1}$ are computed for all bootstrap replications $b = 1, \cdots, B_s$ of each assumed hypothesis,

$$\hat{Z}_{N+1}^{ib} = \log \frac{f_1(\vec{x}_{N+1}^{ib}; \hat{\theta}_1^{(N+1)ib})}{f_0(\vec{x}_{N+1}^{ib}; \hat{\theta}_0^{(N+1)ib})}, \ \bar{Z}_{N+1}^{0b} = \log \frac{f_1(\vec{x}_{N+1}^{0b}; \bar{\theta}_1)}{f_0(\vec{x}_{N+1}^{0b}; \hat{\theta}_0^{(N)})},$$

$$\bar{Z}_{N+1}^{1b} = \log \frac{f_1(\vec{x}_{N+1}^{1b}; \hat{\theta}_1^{(N)})}{f_0(\vec{x}_{N+1}^{1b}; \bar{\theta}_0)}, \ i = 0, 1, \ b = 1, \cdots, B_s \quad (17)$$

where $\hat{\theta}_0^{(N+1)ib}$ and $\hat{\theta}_1^{(N+1)ib}$ are the MLEs based on the bootstrap data $\vec{x}_{N+1}^{ib}$ from (16). Therefore, we have B_s times $|\Delta Z_{N+1}^{ib}| = |\hat{Z}_{N+1}^{ib} - \bar{Z}_{N+1}^{ib}|$ from which $\Delta_N^i, i = 0, 1$, is calculated. Note that the resulting values $|\Delta Z_{N+1}^{ib}|, b = 1, \cdots, B_s$, are first ranked into increasing order to obtain $|\Delta Z_{N+1}^{i1^*}| \leq |\Delta Z_{N+1}^{i2^*}| \leq \cdots \leq |\Delta Z_{N+1}^{iB_s^*}|$ for each hypothesis $\mathcal{H}_i, i = 0, 1$. Therefore, $\Delta_N^i, i = 0, 1$, in (15) are estimated by $\hat{\Delta}_N^i, i = 0, 1$, such that

$$\frac{\#(|\Delta Z_{N+1}^{0b^*}| \geq \hat{\Delta}_N^0 | \mathcal{H}_0)}{B_s} = \epsilon_0, \quad \frac{\#(|\Delta Z_{N+1}^{1b^*}| \geq \hat{\Delta}_N^1 | \mathcal{H}_1)}{B_s} = \epsilon_1, \quad (18)$$

where $\#(x)$ indicates cardinality. At this point, the modified thresholds A_N and B_N in (7) can be determined. The algorithm of the bootstrap based sequential probability ratio test (B-SPRT) is summarized in Table 1.

4. EXAMPLE

As an example, we consider the following binary composite hypothesis problem,

$$\mathcal{H}_0 : \vec{x}_N \sim \mathcal{CN}\left(0, \sigma_0^2 \mathbf{I}\right), \qquad 0 < \sigma_0^2 \leq 1$$

$$\mathcal{H}_1 : \vec{x}_N \sim \mathcal{CN}\left(0, \sigma_1^2 \mathbf{I}\right), \qquad 1.1 \leq \sigma_1^2 < \infty \quad (19)$$

Table 1. Algorithm of B-SPRT

Step 1) Initialize α, β, A and B from (2), B_s, ϵ_0, ϵ_1
Step 2) Draw $N = N_S$ samples and obtain initial MLEs $\hat{\theta}_1^{(N)}$ and $\hat{\theta}_0^{(N)}$ from (5)
REPEAT
Step 3) Update the thresholds using the bootstrap:
$\rightarrow$ Generate B_s bootstrap replications from (16)
$\rightarrow$ For $i = 0, 1, b = 1, \cdots, B_s$, compute $\quad |\Delta Z_{N+1}^{ib}| = |\hat{Z}_{N+1}^{ib} - \bar{Z}_{N+1}^{ib}|$
$\rightarrow$ Rank $|\Delta Z_{N+1}^{ib}|, i = 0, 1, b = 1, \cdots, B_s$, into increasing order: $|\Delta Z_{N+1}^{i1^*}| \leq |\Delta Z_{N+1}^{i2^*}| \leq \cdots \leq |\Delta Z_{N+1}^{iB_s^*}|$
$\rightarrow$ Calculate $\hat{\Delta}_{N+1}^i, i = 0, 1$ from (18) and update the thresholds by (7)
Step 4) Draw next sample $N \leftarrow N + 1$ to obtain MLEs $\hat{\theta}_1^{(N)}$ and $\hat{\theta}_0^{(N)}$
Step 5) Calculate the test statistic $\hat{Z}_N$ from (6)
UNTIL $\hat{Z}_N \geq A_N$ or $\hat{Z}_N \leq B_N$
Step 6) If $\hat{Z}_N \geq A_N$, accept $\mathcal{H}_1$ and if $\hat{Z}_N \leq B_N$, accept $\mathcal{H}_0$

where $\mathbf{I}$ is the $N \times N$ identity matrix. Therefore, the MLEs are

$$\hat{\sigma}_0^{2(N)} = \min\left\{\frac{1}{N}\vec{x}_N^H \vec{x}_N, 1\right\},$$

$$\hat{\sigma}_1^{2(N)} = \max\left\{\frac{1}{N}\vec{x}_N^H \vec{x}_N, 1.1\right\}, \quad (20)$$

and it is easy to verify that

$$\mathcal{H}_0 : \hat{\sigma}_0^{2(N)} \xrightarrow{P} \sigma_0^2 \qquad \hat{\sigma}_1^{2(N)} \xrightarrow{P} \bar{\sigma}_1^2 = 1.1$$

$$\mathcal{H}_1 : \hat{\sigma}_0^{2(N)} \xrightarrow{P} \bar{\sigma}_0^2 = 1 \ \ \hat{\sigma}_1^{2(N)} \xrightarrow{P} \sigma_1^2. \quad (21)$$
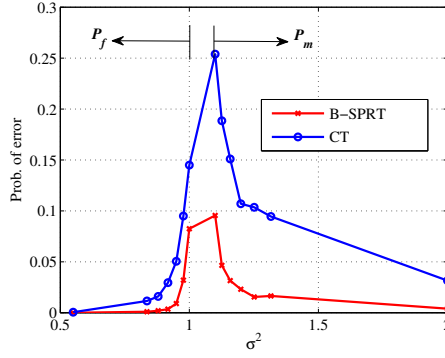
The log-likelihood ratio from (6) now becomes

$$\hat{Z}_N = \left(\frac{1}{\hat{\sigma}_0^{2(N)}} - \frac{1}{\hat{\sigma}_1^{2(N)}}\right)\vec{x}_N^H \vec{x}_N + N \log \frac{\hat{\sigma}_0^{2(N)}}{\hat{\sigma}_1^{2(N)}}. \quad (22)$$

For all simulations, the preset probability of false alarm α is chosen to be equal to the preset probability of missed detection $\beta = \alpha = 0.1$ and $N_S = 1$. We pre-specify the tolerance values in (18) to be $\epsilon_0 = \epsilon_1 = 0.05$, and bootstrap replications are set to $B_s = 100$. Hence, the $\hat{\Delta}_N^i$ is the 5th largest of the ranked $|\Delta Z_{N+1}^{ib^*}|, b = 1, \cdots, B_s$. All the results are generated using 2×10^3 Monte-Carlo runs. We quantify the underlying variance σ^2 for $0.5 < \sigma^2 = \sigma_0^2 \leq 1$ (under $\mathcal{H}_0$) and for $1.1 \leq \sigma^2 = \sigma_1^2 \leq 2$ (under $\mathcal{H}_1$). In terms of signal to noise ratio (SNR), the worst case scenario, $SNR = -10 \ dB$, happens when σ^2 lies at the boundary of the parameter space, i.e. at $\sigma_1^2 = 1.1$ under $\mathcal{H}_1$ and at $\sigma_0^2 = 1$ under $\mathcal{H}_0$.

We apply the central limit theorem on the test statistic (22) under $\mathcal{H}_0$ and $\mathcal{H}_1$ to calculate the minimum required sample number N_f of the fixed sample number detector in Fig. 2, when N_f is large. Then, we arrive at the following expression,

$$N_f \approx \left(\frac{\mathcal{Q}^{-1}(\beta)\left(\frac{\sigma_1^2}{\sigma_0^2} - 1\right) + \mathcal{Q}^{-1}(\alpha)\left(1 - \frac{\sigma_0^2}{\sigma_1^2}\right)}{\frac{\sigma_1^2}{\sigma_0^2} + \frac{\sigma_0^2}{\sigma_1^2} - 2}\right), \quad (23)$$

Fig. 1. The performance improvement of B-SPRT on P_f and P_m compared to the generalized SPRT using constant thresholds A and B (CT).

where $\mathcal{Q}^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution function. For this detector type, N_f is determined in advance and cannot be changed over the unknown σ^2 values in the parameter spaces. Hence, it should be calculated such as to fulfill the requirement of $\beta = \alpha = 0.1$ under the possible worst case scenario, i.e. $\sigma_0^2 = 1$ and $\sigma_1^2 = 1.1$. In this case, (23) yields $N_f \approx 724$ as depicted in Fig. 2.
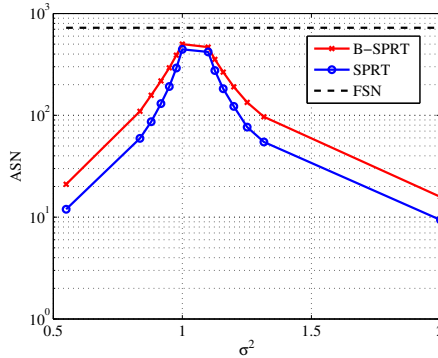
Fig. 2. The ASN of the B-SPRT compared to the SPRT and its counterpart of the fixed sample number detector (FSN).

To quantify the improvement of the B-SPRT in terms of the actual probability of false alarm P_f and probability of missed detection P_m, we compare it with a generalized SPRT whose two thresholds are held constant, i.e. $A_N = A$ and $B_N = A$. The result is depicted in Fig.1 which indicates that for all permitted σ^2 values under each hypothesis, P_f and P_m are smaller than $\alpha = 0.1$ and $\beta = 0.1$, respectively. The performance improvement is significant over the generalized SPRT with the constant thresholds, especially when σ^2 is near the boundaries of the parameter spaces (low SNR region). P_f and P_m of both detectors decrease as σ^2 departs from the boundaries, which is a common symptom in sequential detections [3]. Therefore, setting the thresholds as functions of sample number using the parametric bootstrap leads to convincing results in sequential detections for composite hypotheses.

Fig.2 shows that the average sample number (ASN) of the B-SPRT is less than that of the fixed sample number detector (FSN). It also shows that the ASN of the B-SPRT is still convincing, relative to the ASN of the SPRT which is based on known parameters. It is approximately $1.1\times$ the ASN of the SPRT when σ^2 approaches the boundaries and approximately $1.8\times$ the ASN for $\sigma^2 \ll 1$ or $\sigma^2 \gg 1.1$. For further validation, we compare the average of the upper threshold A_N and the lower B_N of the B-SPRT for $\sigma^2 = 1$ ($\mathcal{H}_0$) and $\sigma^2 = 1.1$ ($\mathcal{H}_1$), to the two thresholds calculated based on Monte-Carlo simulations. Fig.3 shows that their thresholds are relatively close and converge to the constant thresholds (CT) as the sample number N approaches the ASN under each hypothesis. Note that for $\sigma^2 = 1$ and $\sigma^2 = 1.1$, the ASNs of the B-SPRT are approximately 500 and 468, respectively, as can be extracted from Fig.2.
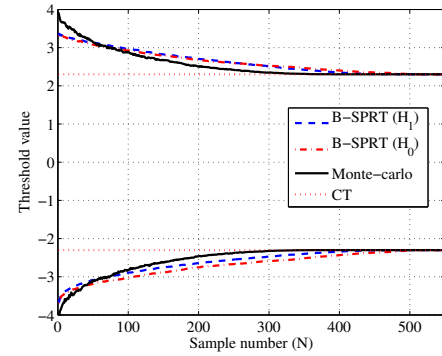
Fig. 3. Thresholds A_N and B_N of B-SPRT under $\mathcal{H}_0$ and $\mathcal{H}_1$ compared to the thresholds calculated based on Monte-Carlo simulations.

The increase of computational cost due to the bootstrap in sequential detection is inevitable. Therefore, finding a more suitable implementation for the bootstrap in a sequential test with a limited increase of computational cost is essential and is subject of our future work. However, in an era of exponentially declining computational costs, bootstrap-based methods such as in the presented application are becoming a bargain and more attractive. Based on these results, the proposed method is a promising technique. It is suitable for the situation where the time limitation is important such as in spectrum sensing for cognitive radio, i.e. to solve the problem of the sensing-throughput trade-off [27].

5. CONCLUSIONS

In this paper we have investigated how to apply the parametric bootstrap in sequential detection for composite hypothesis testing. The bootstrap is used to update the thresholds in an adaptive manner based on current observations. Simulation results have shown that the proposed method improves the probability of false alarm and missed detection of the generalized sequential probability ratio test. In addition, the ASN is comparable to the ASN of the SPRT, which is based on known parameter values, especially in low SNR region. Hence, the proposed method is an attractive candidate to be implemented in spectrum sensing for cognitive radio where the time constraint is a critical issue.

6. REFERENCES

[1] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, pp. 117–186, 1945.

[2] A. Wald, "On cumulative sums of random variables," *The Annals of Mathematical Statistics*, pp. 283–296, 1944.

[3] A. Wald, *Sequential analysis*, Dover Publications, 2004.

[4] J. Bussgang and D. Middleton, "Optimum sequential detection of signals in noise," *IRE Transactions on Information Theory*, vol. 1, no. 3, pp. 5–18, 1955.

[5] S. Tantaratana and J. Thomas, "Relative efficiency of the sequential probability ratio test in signal detection," *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 22–31, 1978.

[6] S. Tantaratana and H. Poor, "Asymptotic efficiencies of truncated sequential tests," *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 911–923, 1982.

[7] C. Lee and J. Thomas, "A modified sequential detection procedure," *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp. 16–23, 1984.

[8] J. Bussgang and M. Marcus, "Truncated sequential hypothesis tests," *IEEE Transactions on Information Theory*, vol. 13, no. 3, pp. 512–516, 1967.

[9] S. Tantaratana and J.B. Thomas, "Truncated sequential probability ratio test," *Information Sciences*, vol. 13, no. 3, pp. 283–300, 1977.

[10] A.M. Hussain, "Multisensor distributed sequential detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 698–708, 1994.

[11] V.V. Veeravalli, "Sequential decision fusion: theory and applications," *Journal of the Franklin Institute*, vol. 336, no. 2, pp. 301–322, 1999.

[12] E. Axell, G. Leus, E.G. Larsson, and H.V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Processing Magazine*, vol. 29, no. 3, pp. 101–116, 2012.

[13] S. Chaudhari, V. Koivunen, and H.V. Poor, "Autocorrelation-based decentralized sequential detection of ofdm signals in cognitive radios," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2690–2700, 2009.

[14] K.W. Choi, W.S. Jeon, and D.G. Jeong, "Sequential detection of cyclostationary signal for cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4480–4485, 2009.

[15] S.J. Kim and G.B. Giannakis, "Sequential and cooperative sensing for multi-channel cognitive radios," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4239–4253, 2010.

[16] Q. Zou, S. Zheng, and A.H. Sayed, "Cooperative sensing via sequential detection," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6266–6283, 2010.

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

[18] B. Efron and R.J. Tibshirani, *An introduction to the bootstrap*, vol. 57, Chapman & Hall/CRC, 1994.

[19] A.M. Zoubir and D.R. Iskander, *Bootstrap techniques for signal processing*, Cambridge University Press, 2004.

[20] A.M. Zoubir and D. Robert Iskander, "Bootstrap methods and applications," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 10–19, 2007.

[21] R.F. Brcich, A.M. Zoubir, and P. Pelin, "Detection of sources using bootstrap techniques," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 206–215, 2002.

[22] H.T. Ong and A.M. Zoubir, "Bootstrap-based detection of signals with unknown parameters in unspecified correlated interference," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 135–141, 2003.

[23] C. Debes, C. Weiss, A.M. Zoubir, and M.G. Amin, "Distributed target detection in through-the-wall radar imaging using the bootstrap," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*. IEEE, 2010, pp. 3530–3533.

[24] A.G. Tartakovsky, "An efficient adaptive sequential procedure for detecting targets," in *IEEE Aerospace Conference Proceedings*, 2002, vol. 4, pp. 4–1581 – 4–1596 vol.4.

[25] A.G. Tartakovsky, X.R. Li, and G. Yaralov, "Sequential detection of targets in multichannel systems," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 425–445, 2003.

[26] S.M. Kay, *Fundamentals of statistical signal processing, Volume II: Detection theory*, vol. 7, Upper Saddle River (New Jersey), 1998.

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