Signal Processing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Spectrum sensing using distributed sequential detection via noisy reporting MAC[☆]

Jithin K. Sreedharan [a,*], Vinod Sharma [b]

[a] INRIA Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex, France
[b] Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India

ARTICLE INFO

Article history:
Received 16 March 2014
Received in revised form
10 July 2014
Accepted 12 July 2014

Keywords:
Cooperative spectrum sensing
Distributed detection
Sequential detection

ABSTRACT

This paper considers cooperative spectrum sensing algorithms for Cognitive Radios which focus on reducing the number of samples to make a reliable detection. We propose algorithms based on decentralized sequential hypothesis testing in which the Cognitive Radios sequentially collect the observations, make local decisions and send them to the fusion center for further processing to make a final decision on spectrum usage. The reporting channel between the Cognitive Radios and the fusion center is assumed more realistically as a Multiple Access Channel (MAC) with receiver noise. Furthermore the communication for reporting is limited, thereby reducing the communication cost. We start with an algorithm where the fusion center uses an SPRT-like (Sequential Probability Ratio Test) procedure and theoretically analyze its performance. Asymptotically, its performance is close to the optimal centralized test without fusion center noise. We further modify this algorithm to improve its performance at practical operating points. Later we generalize these algorithms to handle uncertainties in SNR and fading.

1. Introduction

Presently there is a scarcity of spectrum due to the proliferation of wireless services. Cognitive Radios (CRs) are proposed as a solution to this problem. They access the spectrum licensed to existing communication services (primary users) opportunistically and dynamically without causing much interference to the primary users. This is made possible via spectrum sensing by the Cognitive Radios (secondary users), to gain knowledge about the spectrum usage by the primary devices. However due to the strict spectrum sensing requirements [1] and the various inherent wireless channel impairments, spectrum sensing has become one of the main challenges faced by the Cognitive Radios.

Multipath fading, shadowing and hidden node problem cause serious problems in spectrum sensing. Cooperative (decentralized or distributed) spectrum sensing in which different cognitive radios interact with each other exploiting spatial diversity [1,2] is proposed as an answer to these problems. It also reduces the probability of false alarm and the probability of miss-detection. Cooperative spectrum sensing can be either centralized or distributed [1]. In the centralized algorithm a central unit gathers sensing data from the Cognitive Radios and identifies the spectrum usage [3]. On the other hand, in the distributed case each secondary user (SU) collects observations, makes a local decision and sends it to a fusion center (FC) to make the final decision. Centralized algorithms provide better performance but also have more communication overhead in transmitting all the data to the fusion node. In the distributed case, the information that is

---

[☆] Preliminary versions of this paper have been presented in NCC 2011, WCNC 2011 and Allerton 2011. This work is partly supported by a grant from ANRC.

[*] Corresponding author. Tel.: +33 4 92 38 78 46.
E-mail addresses: jithin.sreedharan@inria.fr (J.K. Sreedharan), vinod@ece.iisc.ernet.in (V. Sharma).

http://dx.doi.org/10.1016/j.sigpro.2014.07.009
0165-1684/© 2014 Elsevier B.V. All rights reserved.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

exchanged between the secondary users and the fusion node can be a soft decision (summary statistic) or a hard decision. Soft decisions can give better gains at the fusion center but also consume higher bandwidth at the control channels (used for sharing information among secondary users). However hard decisions provide as good a performance as soft decisions when the number of cooperative users increases [3].

Spectrum sensing problem can be formulated in different ways, two of them being Neyman–Pearson framework (fixed sample size detection) and sequential detection framework which reduces the average number of samples taken for deciding if a primary is transmitting or not [4]. Also, there are two types of sequential detection: one can consider detecting when a primary turns ON (or OFF) (change detection, see [5,6] and the references therein) or just testing the hypothesis whether the primary is ON or OFF ([7–9] and references therein). In [5], cooperative spectrum sensing under sequential change detection framework with no coordination between the secondary users is considered, and random broadcast policies and several improvements are proposed. In [6] a nonparametric framework is considered and performance is studied theoretically also. In sequential hypothesis testing one considers the case where the status of the primary channel is known to change very slowly, e.g., detecting occupancy of a TV transmission. Usage of idle TV bands by the Cognitive network is being targeted as the first application for cognitive radio. In this setup (minimising the expected sensing time with constraints on probability of errors) Walds' SPRT (Sequential Probability Ratio Test) provides the optimal performance for a single Cognitive Radio [4]. But the optimal solutions for cooperative setup are not available [10].

In this paper, we consider sequential hypothesis testing in cooperative setup. Feedback from the fusion node to the CRs can possibly improve the performance. However that also requires an extra signaling channel which may not be available and has its own cost. Therefore we do not consider feedback in our system. In sequential decentralized detection framework, optimization needs to be performed jointly over sensors and fusion center policies as well as over time. Unfortunately, this problem is intractable for most of the sensor configurations [10,11], specifically when there is no feedback from the fusion center and there is limited local memory, which is more relevant in practical situations. Recently [11] and [12] proposed asymptotically optimal (order 1 (Bayes) and order 2 respectively) decentralized sequential hypothesis tests for such systems with full local memory. But these models do not consider noise at the fusion center and assume a perfect communication channel between the CR nodes and the fusion center. Also, often asymptotically optimal tests do not perform well at a finite number of observations. Zou et al. and Yilmaz et al. [7,8] also proposed cooperative sequential algorithms for spectrum sensing, but neither of them deal with the fusion center noise and SNR uncertainty case.

Noisy channels between local nodes and fusion center are considered in [13] in the decentralized sequential detection framework. But optimality of the tests is not discussed and the paper is more focused on finding the best signalling schemes at the local nodes with the assumption of parallel channels between local nodes and the fusion center and perfect knowledge of local node probabilities of error.

We first propose a decentralized algorithm DualSPRT which uses SPRT at the local nodes and a SPRT-like test at the fusion center. Furthermore, we consider the receiver noise at the fusion center and allow multiple local nodes to transmit simultaneously their decisions to the fusion center to reduce the transmission time. This of course means that the fusion center does not know explicitly how many local nodes are transmitting at a time and certain fusion center decision rules, e.g., AND/OR/Majority [1,3] are ruled out in our setup. Moreover unlike some of the previous works on cooperative spectrum sensing using sequential testing (see [9,13] and references therein) we analyze this algorithm theoretically also.

We study asymptotic performance of DualSPRT, with fusion center noise. It is particularly important in the CR context because of detection in wireless channels at low SNR [14]. It can approach the optimal centralized sequential solution (in Bayes and frequentist sense), which does not consider noise at FC. We assume a MAC (Multiple Access Channel) as the reporting channel at the fusion center and the test is not based on the local node probability of error. Later we modify DualSPRT to improve its performance. The parameters of the modified algorithm are easier to fine tune also. Furthermore we introduce a new way of quantizing SPRT decisions of local nodes and extend this algorithm to cover SNR uncertainties and fading channels. We have seen via simulations that our algorithm works better than the algorithm in [11] and almost as well as the algorithm in [12] even when the fusion center noise is not considered and MAC layer transmission delays are ignored in [12,11]. Li and Evans [15] and Li et al. [16] consider distributed detection with MAC, but not in sequential detection framework. Banavar et al. [17,18] take into account MAC in the distributed estimation setup.

In addition, we generalize our algorithm to include uncertainty in the received Signal to Noise Ratio (SNR) at the CRs and fading channels between primary and CR. This requires a composite hypothesis testing extension to the decentralized sequential detection problem and is not considered in any of the above references (although [13] considers SNR uncertainty and fading between the CRs and the fusion center).

This paper is organized as follows. Section 2 presents the model. Section 3 provides the DualSPRT algorithm. An approximate theoretical performance of the algorithm is also provided. Section 4 studies the asymptotic performance of DualSPRT. In Section 5 we improve over DualSPRT. We compare the different versions so obtained and also compare them with existing asymptotically optimal decentralized sequential algorithms. Section 6 extends these algorithms to consider the effect of fading and SNR uncertainty. Section 7 concludes the paper.

2. System model

We consider a Cognitive Radio system with one primary transmitter and L secondary users. The L nodes sense the channel to detect the spectral holes. The decisions made by the secondary users are transmitted to a fusion node via a reporting MAC for it to make a final decision. This is the most common architecture for distributed detection and distributed

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

spectrum sensing [1,14]. In order to keep the traffic on the reporting channel low, we will ensure that a local node transmits a finite valued message to the fusion center.

Let $X_{k,l}$ be the observation made at secondary user l at time k. The $\{X_{k,l}, k \ge 1\}$ are independent and identically distributed (i.i.d.). It is assumed that the observations are independent across Cognitive Radios. Based on $\{X_{n,l}, n \le k\}$ the secondary user l transmits $Y_{k,l}$ to the fusion node. It is assumed that the secondary nodes are synchronized so that the fusion node receives $Y_k = \sum_{l=1}^{L} Y_{k,l} + Z_k$, where $\{Z_k\}$ is i.i.d. receiver noise. The fusion center uses $\{Y_k\}$ and makes a decision. The observations $\{X_{k,l}\}$ depend on whether the primary is transmitting (Hypothesis H_1) or not (Hypothesis H_0) as

$$\text{Under } H_0\text{: } X_{k,l} = \zeta_{k,l}, \quad k = 1, 2, \ldots,$$
$$\text{Under } H_1\text{: } X_{k,l} = h_l S_k + \zeta_{k,l}, \quad k = 1, 2, \ldots,$$

where h_l is the channel gain of the lth user, S_k is the primary signal and $\zeta_{k,l}$ is the observation noise at the lth user at time k. We assume that $\{\zeta_{k,l}, k \ge 1\}$ are i.i.d. Let N be the time to decide on the hypothesis by the fusion node. We assume that N is much less than the coherence time of the channel so that the slow fading assumption is valid. This means that h_l is random but remains constant during the spectrum sensing duration.

The general problem is to develop a distributed algorithm in the above setup which solves the problem:

$$\begin{aligned} &\min \quad E_{DD} \overset{\Delta}{=} E_i[N] \\ &\text{subject to} \quad P_1(\text{Reject } H_1) \le \alpha_1 \ \& \ P_0(\text{Reject } H_0) \le \alpha_0, \end{aligned} \tag{1}$$

where P_i is the probability measure and E_i the expectation when H_i is the true hypothesis, $i \in \{0, 1\}$, and $0 \le \alpha_0, \alpha_1 \le 1$. We will separately consider $E_1[N]$ and $E_0[N]$. It is well known that for a single node case ($L=1$) Wald's SPRT performs optimally in terms of reducing $E_1[N]$ and $E_0[N]$ for given probability of errors. Motivated by the optimality of SPRT for a single node (and DualCUSUM in [6]), we propose using DualSPRT in the next section and study its performance.

We use P_{MD} for $P_1(\text{reject } H_1)$ and P_{FA} for $P_0(\text{reject } H_0)$. In case of E_{DD}, hypothesis under consideration can be understood from the context.

3. Decentralized sequential tests: dualSPRT

In this section we develop DualSPRT algorithm for decentralized sequential detection and also study its performance.

3.1. DualSPRT algorithm

To explain the setup and analysis we start with the simple case, where the channel gain, $h_l = 1$ for all l's. We will consider fading in Section 6. In this algorithm, S_k is assumed to be fully known. This assumption can be weakened by using averaged energy samples discussed later in Section 6.1. DualSPRT is as follows:

1. Secondary node l, computes at step k,

   $$W_{0,l} = 0,$$
   $$W_{k,l} = W_{k-1,l} + \log[f_{1,l}(X_{k,l})/f_{0,l}(X_{k,l})], \quad k \ge 1,$$

   where $f_{1,l}$ is the density of $X_{k,l}$ under H_1 and $f_{0,l}$ is the density of $X_{k,l}$ under H_0 (w.r.t. a common distribution).
2. Secondary node l transmits a constant b_1 at time k if $W_{k,l} \ge \gamma_{1,l}$ or transmits b_0 when $W_{k,l} \le -\gamma_{0,l}$. When $W_{k,l}$ does not cross the interval $(-\gamma_{0,l}, \gamma_{1,l})$, node l does not transmit anything, i.e.,

   $$Y_{k,l} = b_1 \, \mathbb{1}\{W_{k,l} \ge \gamma_{1,l}\} + b_0 \, \mathbb{1}\{W_{k,l} \le -\gamma_{0,l}\}$$

   where $\gamma_{0,l}, \gamma_{1,l} > 0$ and $\mathbb{1}\{A\}$ denotes the indicator function of set A. Parameters $b_1, b_0, \gamma_{1,l}, \gamma_{0,l}$ are chosen appropriately.
3. Physical layer fusion is used at the fusion centre, i.e., $Y_k = \sum_{l=1}^{L} Y_{k,l} + Z_k$, where $\{Z_k\}$ is the i.i.d. noise at the fusion node.
4. Finally, fusion center calculates

   $$\begin{aligned} &F_k = F_{k-1} + \log[g_{\mu_1}(Y_k)/g_{-\mu_0}(Y_k)], \\ &F_0 = 0, \quad \mu_0 > 0, \quad \mu_1 > 0, \end{aligned} \tag{2}$$

   where $g_{-\mu_0}$ is the density of $Z_k - \mu_0$ and g_{μ_1} is the density of $Z_k + \mu_1$, μ_0 and μ_1 being positive constants appropriately chosen.
5. The fusion center decides about the hypothesis at time N where

   $$N = \inf\{k: F_k \ge \beta_1 \text{ or } F_k \le -\beta_0\}$$

   and $\beta_0, \beta_1 > 0$. The decision at time N is H_1 if $F_N \ge \beta_1$, otherwise H_0.

Performance of this algorithm depends on $(\gamma_{1,l}, \gamma_{0,l}, \beta_1, \beta_0, b_1, b_0, \mu_1, \mu_0)$. In particular these parameters should be chosen such that the overall probabilities of error are less than α_1 and α_0 respectively. Any prior information available about H_0 or H_1 can be used to decide constants (via, say, formulating this problem in the Bayesian framework; we will comment on this again). Also we choose these parameters such that the probability of false alarm/miss-detection, P_{fa}/P_{md} at local nodes is higher than P_{FA}/P_{MD}. A good set of parameters for given SNR values can be obtained from our analysis below.

Deciding at local nodes and transmitting decisions to the fusion node reduces the transmission rate and transmit energy used by the local nodes in communication with the fusion node. Also, physical layer fusion in Step 3 reduces transmission time, but requires synchronization of different local nodes. This assumption has been made in other distributed detection/estimation studies also [6,17]. If synchronization is not possible in a given system, then some other MAC algorithm, e.g., TDMA can be used with channel coding. But this will incur extra delay.

Using sequential tests at SUs and at FC (without physical layer synchronization and fusion receiver noise) has been shown to perform well in [7,9]. In the next subsection we analyze the performance under our setup.

3.2. Performance analysis

We first provide the analysis for the mean detection delay E_{DD} and then for P_{FA}. The exact analysis is intractable as for other distributed detection algorithms. Thus we provide an approximate analysis. We also study the asymptotic performance of the algorithm in Section 4 which shows that our algorithm is asymptotically optimal. However we will see that the approximate analysis of this section is more accurate for finite values of parameters than the asymptotic analysis.

KL-divergence of two probability distributions P and Q on the same measurable space $(\Omega, \mathcal{F})$ is defined as

$$D(P \| Q) = \begin{cases} \int \log \dfrac{dP}{dQ}\, dP & \text{if } P \ll Q, \\ \infty & \text{otherwise,} \end{cases} \qquad (3)$$

where $P \ll Q$ denotes that P is absolutely continuous w.r.t. Q. More explicitly, at node l, let

$$\delta_{i,l} = E_i\left[\log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})}\right], \quad \rho_{i,l}^2 = Var_{H_i}\left[\log \frac{f_{1,l}(X_{k,l})}{f_{0,l}(X_{k,l})}\right].$$

Then $\delta_{1,l} = D(f_{1,l} \| f_{0,l})$ and $\delta_{0,l} = -D(f_{0,l} \| f_{1,l})$. We will assume $\delta_{i,l}$ finite throughout this paper. Sometimes we will also need $\rho_{i,l}^2 < \infty$. When the true hypothesis is H_1, by Jensen's Inequality, $\delta_{1,l} > 0$ and when it is H_0, $\delta_{0,l} < 0$. At secondary node l, SPRT sum $\{W_{k,l}, k \geq 0\}$ is a random walk with drift given by $\delta_{i,l}$ under the true hypothesis H_i.

Let $N_l = \inf\{k: W_{k,l} \notin (-\gamma_{0,l}, \gamma_{1,l})\}, N_l^1 = \inf\{k: W_{k,l} \geq \gamma_{1,l}\}$ and $N_l^0 = \inf\{k: W_{k,l} \leq -\gamma_{0,l}\}$. Then $N_l = \min\{N_l^0, N_l^1\}$. Also let $N^0 = \inf\{k: F_k \leq -\beta_0\}$ and $N^1 = \inf\{k: F_k \geq \beta_1\}$. Then stopping time of DualSPRT, $N = \min(N^1, N^0)$.

For simplicity in the rest of this section, we take $\gamma_{1,l} = \gamma_{0,l} = \gamma$, $\beta_1 = \beta_0 = \beta$, $b_1 = -b_0 = b$ and $\mu_1 = \mu_0 = \mu$. Of course the analysis will carry over for the general case.

For convenience we summarize the important notation used in this paper in Table 1. Notation specific to some algorithms are also mentioned separately in parenthesis.

3.2.1. E_{DD} analysis

At the fusion node F_k crosses β under H_1 when a sufficient number of local nodes transmit b_1. The dominant event occurs when the number of local nodes transmitting are such that the mean value of the increments of the sum F_k will just have turned positive. In the following we find the mean time to this event and then the time to cross β after this. The E_{DD} analysis is same under hypothesis H_0 and H_1. Hence we provide the analysis for H_1.

The following lemmas provide justification for considering only the events $\{N_l^i\}$ and $\{N^i\}$ for analysis of $E_{DD} = E_i[N]$.

Lemma 1. *For* $i = 0,1$, $P_i(N_l = N_l^i) \to 1$ *as* $\gamma \to \infty$ *and* $P_i(N = N^i) \to 1$ *as* $\gamma \to \infty$ *and* $\beta \to \infty$.

Proof. From random walk results [19, Chapter IV] we know that if a random walk has negative drift then its maximum is finite with probability one. This implies that $P_i(N_l^j < \infty) \to 0$ as $\gamma \to \infty$ for $i \neq j$ but $P_i(N_l^i < \infty) = 1$ for any $\gamma < \infty$. Thus $P_i(N_l = N_l^i) \to 1$ as $\gamma \to \infty$. This also implies that as $\gamma \to \infty$, the mean of increments of F_k is positive for H_1 and negative for H_0. Therefore, $P_i(N = N^i) \to 1$ as $\gamma \to \infty$ and $\beta \to \infty$. □

Lemma 2. *Under* H_i, $i = 0,1$ *and* $j \neq i$,

(a) $|N_l - N_l^i| \to 0$ *a.s. as* $\gamma \to \infty$ *and* $\lim_{\gamma \to \infty} N_l/\gamma = \lim_{\gamma \to \infty} N_l^i/\gamma = 1/D(f_{i,l} \| f_{j,l})$ *a.s. and in* L^1.

(b) $|N - N^i| \to 0$ *a.s. and* $\lim N/\beta = \lim N^i/\beta$ *a.s. and in* L^1, *as* $\gamma \to \infty$ *and* $\beta \to \infty$.

Table 1
List of important notations. 1 – DualSPRT; 2 – SPRT-CSPRT; 3 – GLR-SPRT; 4 – GLR-CSPRT.

Notation	Meaning
L	Number of CRs
$X_{k,l}$	Observation at CR l at time k
$Y_{k,l}$	Transmitted value from CR l to FC at time k
Y_k	FC observation at time k
h_l	Channel gain of the lth CR
$\zeta_{k,l}$	Observation noise at CR l at time k
Z_k	FC MAC noise at time k
$f_{i,l}$, g_μ	PDF of $X_{1,l}$ under H_i, PDF of $Z_k + \mu$
$W_{k,l}$	Test statistic at CR l at time k
F_k	Test statistic at FC at time k (1)
F_k^1, F_k^0	Test statistics at FC (2)
ξ_k	LLR at FC (1)
$\{\xi_k^*, k \geq 1\}$	i.i.d with distribution of ξ^*, defined in (10), (1)
F_n^*, $\widehat{F}_n^*$	$\sum_{k=1}^n \xi_k^*$, $\sum_{k=1}^n \lvert\xi_k^*\rvert$ (1)
$\mathcal{A}^i$, $\Delta(\mathcal{A}^i)$	{all CRs transmit b_i under H_i}, $E_i[\xi_k \lvert \mathcal{A}^i]$ (1)
$\gamma_{1,l}, \gamma_{0,l}$	Thresholds at CR l (1,2)
$g(t)$	Threshold at CR (3, 4)
β_1, β_0	Thresholds at FC
μ_1, μ_0	Design parameters in FC LLR
b_1, b_0	Transmitting values to the FC at CR (1, 3)
b_j^i	Transmitting values to the FC at CR (2, 4)
N	First time F_k crosses $(-\beta_0, \beta_1)$ (1)
N^1, N^0	First time F_k crosses β_1, crosses $-\beta_0$ (1)
N_l, N_l^1, N_l^0	Corresponding values of N, N^1, N^0 at CR l (1)
$N_l(g, c)$	First time $W_{n,l}$ crosses $g(nc)$ at CR l (3, 4)
τ_β, T_β	First time F_k^0 crosses $-\beta_0$, F_k^1 crosses β_1 (2)
$\delta_{i,l}, \rho_{i,l}^2$	Mean and variance of LLR at CR l under H_i
$\delta_{i,FC}^j$	Mean of LLR at FC under H_i when j CRs transmit
t_j	Time epoch when $\delta_{i,FC}^{j-1}$ changes to $\delta_{i,FC}^j$ (1)
T_j	Time epoch when $\delta_{i,FC}^{j-1}$ changes to $\delta_{i,FC}^j$ (2)
$\overline{F_j}$	$E[F_{t_j - 1}]$
D_{tot}^0, D_{tot}^1	$\sum_{l=1}^L D(f_{0,l} \| f_{1,l})$, $\sum_{l=1}^L D(f_{1,l} \| f_{0,l})$
r_l, ρ_l	$D(f_{0,l} \| f_{1,l})/D_{tot}^0$, $D(f_{1,l} \| f_{0,l})/D_{tot}^1$
$\tau_l(c)$	Last time RW with drift $\delta_{0,l}$ will be above $-\lvert \log c \rvert$
$\tau(c)$	$\max_{1 \leq l \leq L} \tau_l(c)$
R_i	$\min_{1 \leq l \leq L} - \log \inf_{t \geq 0} E_i\left[\exp\left(-t \log \frac{f_{1,l(X_{1,l})}}{f_{0,l}(X_{1,l})}\right)\right]$
G, g	CDF of $\lvert\xi_1^*\rvert$, MGF of $\lvert\xi_1^*\rvert$
$\Lambda(\alpha)$	$\sup_\lambda(\alpha\lambda - \log g(\lambda))$
α^+	ess sup$\lvert\xi_1^*\rvert$
$\mathcal{R}_c(\delta)$	Bayes Risk of test δ with cost c
$\nu(a)$	First time RW $\left\{\log \frac{g_{\mu_1(Z_k)}}{g_{-\mu_0(Z_k)}}\right\} + \left(\Delta\left(\mathcal{A}^0\right) - E_0\left[\log \frac{g_{\mu_1(Z_k)}}{g_{-\mu_0(Z_k)}}\right]\right.$ $k \geq \tau(c) + 1\}$ crosses a.

Proof. Under H_0,

$$N_l^0\, \mathbb{1}\{N_l^0 < N_l^1\} \leq N_l \leq N_l^0, \qquad (4)$$

and since $P_0[N_l^0 < N_l^1] \to 1$ as $\gamma \to \infty$, $|N_l^0 - N_l| \to 0$ a.s. as $\gamma \to \infty$. Also from Random Walk results [19, p. 88], $N_l^0/\gamma \to 1/D(f_{0,l} \| f_{1,l})$ a.s. and $E[N_l^0]/\gamma \to 1/D(f_{0,l} \| f_{1,l})$. Thus we also obtain $N_l/\gamma \to 1/D(f_{0,l} \| f_{1,l})$ a.s. and in L^1. Similarly the corresponding results hold for N and N^0 as γ and $\beta \to \infty$. (4) holds in the expected sense also. □

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

Thus when γ is large, we can approximate $E_1[N_l]$ by $\gamma/D(f_{1,l}\|f_{0,l})$. Also under H_1, by central limit theorem for the first passage time N_l^1 (Theorem 5.1, Chapter III in [19]),

$$N_l^1 \sim \mathcal{N}\left(\frac{\gamma}{\delta_{1,l}}, \frac{\rho_{1,l}^2 \gamma}{\delta_{1,l}^3}\right), \tag{5}$$

where $\mathcal{N}(a,b)$ denotes Gaussian distribution with mean a and variance b. From Lemma 2, we can use this result for N_l also. Similarly we can obtain the results under H_0 and at the fusion node. Let $\delta^j_{i,FC}$ be the mean of increments of the fusion center test sum F_k, under H_i, when j local nodes are transmitting. Let t_j be the point at which the mean of increments of F_k changes from $\delta^{j-1}_{i,FC}$ to $\delta^j_{i,FC}$ and let $\overline{F_j} = E[F_{t_j - 1}]$, the mean value of F_k just before transition epoch t_j. The following lemma holds.

Lemma 3. *Under H_i, $i=0,1$, as $\gamma \to \infty$, P_i (Decision at time t_k is H_i and t_k is the kth order statistics of $N_1^i, N_2^i, \ldots, N_L^i$) $\to 1$.*

Proof. From Lemma 1, P_i (Decision at time t_k is H_i and t_k is the kth order statistics of $N_1^i, N_2^i, \ldots, N_L^i$) $\geq P_i(N_l^i < N_l^j, j \neq i, l = 1, \ldots, L) \to 1$, as $\gamma \to \infty$. □

We use Lemmas 1–3 and Eq. (5) in the following to obtain an approximation for E_{DD} when γ and β are large. Large γ and β are needed for small probability of error. Then we can assume that the local nodes are making correct decisions. Although F_k is a random walk before t_1, it is not so between t_j and t_{j+1} for $j=1,\ldots,L$. But we assume that in the following approximation.

Let

$$l^* = \min\left\{ j: \delta^j_{1,FC} > 0 \text{ and } \frac{\beta - \overline{F_j}}{\delta^j_{1,FC}} < E[t_{j+1}] - E[t_j] \right\}.$$

$\overline{F_j}$ can be iteratively calculated as

$$\overline{F_j} = \overline{F}_{j-1} + \delta^j_{1,FC}(E[t_j] - E[t_{j-1}]), \quad \overline{F}_0 = 0. \tag{6}$$

Note that $\delta^j_{1,FC}$ $(0 \leq j \leq L)$ can be found by assuming $E_1[Y_k]$ as b_j and t_j as the jth order statistics of $\{N_l^1,\ 0 \leq l \leq L\}$. The Gaussian approximation (5) can be used to calculate the expected value of the order statistics using the method given in [20]. This implies that $E[t_j]'s$ and hence $\overline{F_j}s$ are available offline. By using these values E_{DD} $(\approx E_1(N^1))$ can be approximated as

$$E_{DD} \approx E[t_{l^*}] + \frac{\beta - \overline{F_{l^*}}}{\delta^{l^*}_{1,FC}}, \tag{7}$$

where the first term on R.H.S. is the mean time till the mean of increments becomes positive at the fusion node while the second term indicates the mean time for F_k to cross β from t_{l^*} onward.

3.2.2. P_{MD}/P_{FA} analysis

We provide analysis under H_1. P_{FA} analysis is same as that of P_{MD} analysis with obvious changes. When the thresholds at local nodes are reasonably large, according to Lemma 3, with a large probability local nodes are making the right decisions and t_k can be taken as the order statistics assuming that all local nodes make the right decisions. Then for missed detection the dominant event is $P_1(N^0 < t_1)$. Also for reasonable performance we should select thresholds such that $P_1(N^1 < t_1)$ is small. Then

$$P_{MD} = P_1(N^0 < N^1) \geq P_1(N^0 < t_1, N^1 > t_1) \approx P_1(N^0 < t_1). \tag{8}$$

Under the above conditions, this lower bound should give a good approximation. In the following, we get an approximation for this.

Let $\xi_k = \log[g_{\mu_1}(Y_k)/g_{-\mu_0}(Y_k)]$. Then $F_k = \xi_1 + \xi_2 + \cdots + \xi_k$ and if we assume that ξ_k before t_1, has mean zero and has distribution symmetric about zero (e.g., $\sim \mathcal{N}(0,\sigma^2)$) then

$$\begin{aligned}
&P_1(\text{reject } H_1 \text{ before } t_1)\\
&\approx \sum_{k=1}^{\infty} P_1\left[\{F_k < -\beta\} \bigcap_{n=1}^{k-1} \{F_n > -\beta\} | t_1 > k\right] P[t_1 > k]\\
&= \sum_{k=1}^{\infty} \left(P_1\left[F_k < -\beta | \bigcap_{n=1}^{k-1} \{F_n > -\beta\}\right] P_1\left[\bigcap_{n=1}^{k-1} \{F_n > -\beta\}\right]\right)(1 - \Phi_{t_1}(k))\\
&\overset{(A)}{=} \sum_{k=1}^{\infty} \left(P_1[F_k < -\beta | F_{k-1} > -\beta] P_1\left[\inf_{1 \leq n \leq k-1} F_n > -\beta\right]\right)(1 - \Phi_{t_1}(k))\\
&\overset{(B)}{\geq} \sum_{k=1}^{\infty} \left(\int_{c=0}^{\infty} P_1[\xi_k < -c] f_{F_{k-1}}\{-\beta + c\}\, dc \right)\\
&\quad (1 - 2P_1[F_{k-1} \leq -\beta])(1 - \Phi_{t_1}(k)),
\end{aligned}$$

where Φ_{t_1} is the Cumulative Distribution Function of t_1. Since we are considering only $\{F_k, k \leq t_1\}$, we remove the dependencies on t_1. In the above equations (A) is because of the Markov property of the random walk and (B) is due to the following lemma. This lemma can be obtained from [21, p. 525].

Lemma 4. *If ξ_1 has mean zero and distribution symmetric about zero*

$$P\left[\inf_{1 \leq n \leq k-1} F_n > -\theta\right] \geq 1 - 2P[F_{k-1} \leq -\theta].$$

Similarly we can write an upper bound by replacing $P[\cap_{n=1}^{k-1}\{F_n > -\theta\}]$ with $P[F_{k-1} > -\theta]$. We can make the lower bound tighter if we do the same analysis for the random walk between t_1 and t_2 with appropriate changes and add to the above bounds.

3.2.3. Example 1

We apply the DualSPRT on the following example and compare the E_{DD} and P_{FA} via analysis provided above with the simulation results. We assume that f_0 and f_1 are Gaussian with different means. This model is relevant when the noise and interference are log-normally distributed [2], and when $X_{k,l}$ is the sum of energy of a large number of observations at the secondary nodes at a low SNR.

Parameters used for simulation are as follows: $L=5$, $f_0 \sim \mathcal{N}(0,1)$ and $f_1 \sim \mathcal{N}(1,1)$. Also $f_0 = f_{0,l}$ and $f_1 = f_{1,l}$ for $1 \leq l \leq L$, and $b=1$. We plot P_E $(=P_{FA}$ under H_0 and P_{MD} under $H_1)$ and E_{DD} $(=E_1[N]$ or $E_0[N])$ versus β in Figs. 1 and 2 respectively. Here γ, μ and b are fixed for ease of calculation and they are chosen to provide good performance for the given P_{MD}/P_{FA}. The figures also contain the results obtained via analysis. We see a good match in theory and simulations. For comparison, Fig. 2 also contains asymptotic results which are presented in Section 4 below.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

The above example is for the case when $X_{k,l}$ have the same distribution for different l under the hypothesis H_0 and H_1. The next example considers the general case. Now the order statistics t_{l^*} in (7) needs to be appropriately computed.

3.2.4. Example 2

There are five secondary nodes with primary to secondary channel gain being 0, -1.5, -2.5, -4 and -6 dB respectively (corresponding post change means are 1, 0.84, 0.75, 0.63, 0.5). $f_0 \sim \mathcal{N}(0,1), f_0 = f_{0,l}$ for $1 \le l \le L$. Figs. 3 and 4 respectively provide the P_{FA} and E_{DD} via analysis and simulations. We see a good match.

4. Asymptotic properties of DualSPRT

In this section we prove asymptotic properties of DualSPRT.

We use the following notation:

$$D_{tot}^0 = \sum_{l=1}^{L} D(f_{0,l} \| f_{1,l}), \quad D_{tot}^1 = \sum_{l=1}^{L} D(f_{1,l} \| f_{0,l})$$

$$r_l = \frac{D(f_{0,l} \| f_{1,l})}{D_{tot}^0}, \quad \rho_l = \frac{D(f_{1,l} \| f_{0,l})}{D_{tot}^1}.$$

Let $\mathcal{A}^i$ be the event that all the secondary users transmit b_i when the true hypothesis is H_i. Also let $\Delta(\mathcal{A}^i)$ be the mean of increments of F_k when $\mathcal{A}^i$ happens, i.e., $\Delta(\mathcal{A}^i) = E_i[(\log g_{\mu_1}(Y_k)/g_{-\mu_0}(Y_k))|\mathcal{A}^i]$.

In the rest of this section, local node thresholds are $\gamma_{0,l} = -r_l|\log c|, \gamma_{1,l} = \rho_l|\log c|$ and fusion center thresholds are $\beta_0 = -|\log c|, \beta_1 = |\log c|$.

We will also need

$$\tau_l(c) \triangleq \sup\{n \ge 1 : W_{n,l} \ge -r_l|\log c|\}, \ \tau(c) \triangleq \max_{1 \le l \le L} \tau_l(c). \tag{9}$$

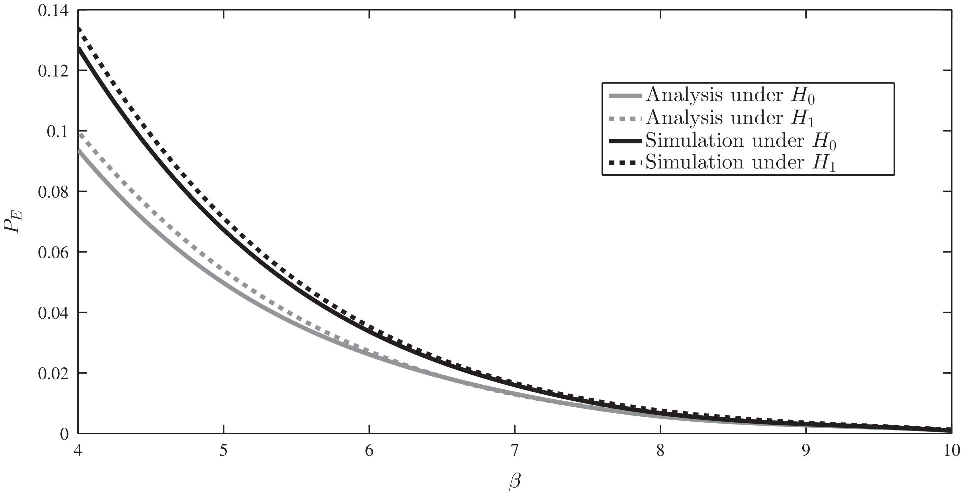

Fig. 1. DualSPRT-comparison between theory and simulation of probability of error.

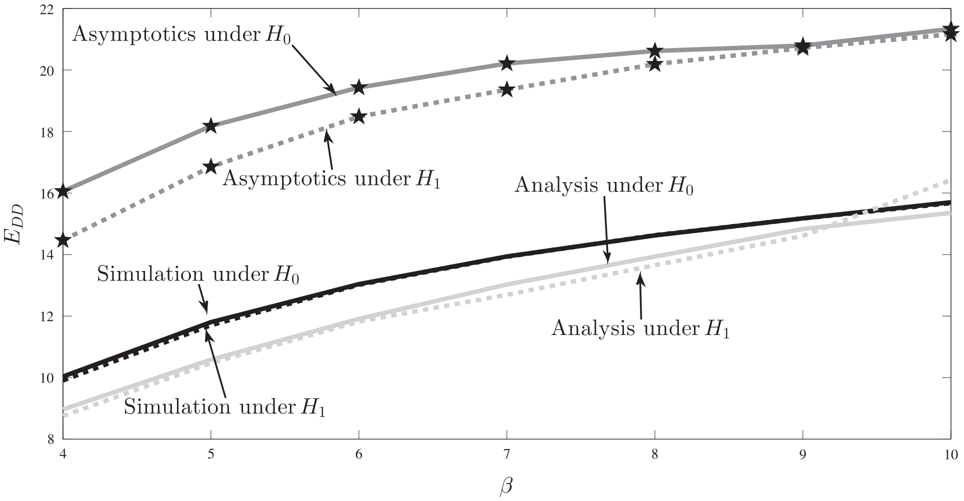

Fig. 2. DualSPRT-comparison between theory and simulation of expected detection delay.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

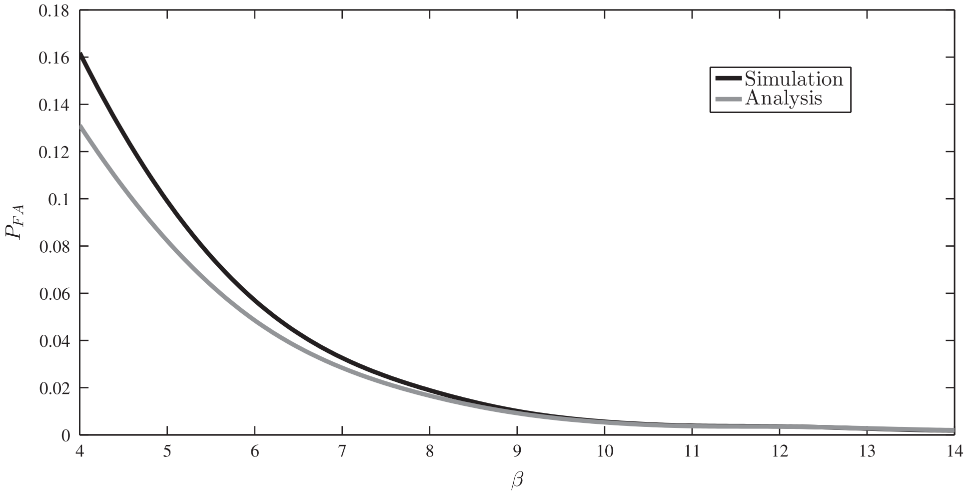

Fig. 3. DualSPRT-comparison between theory and simulation of probability of false alarm.

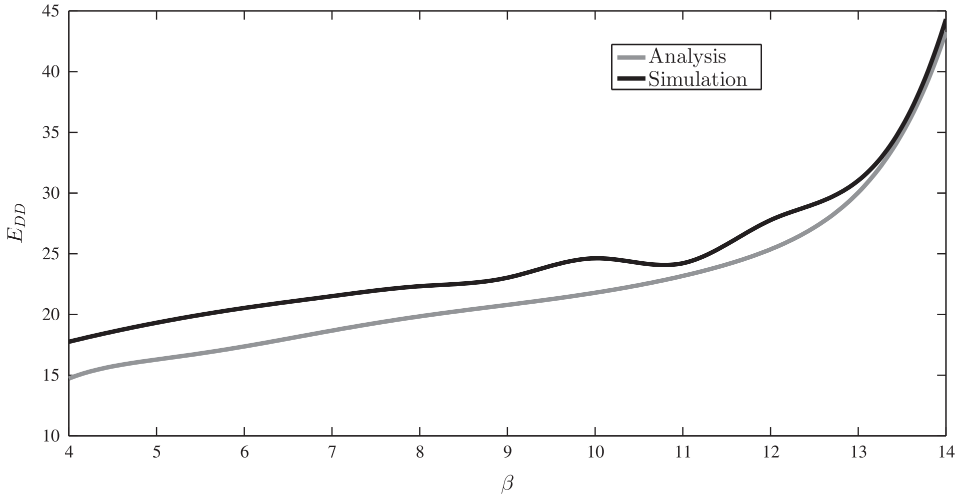

Fig. 4. DualSPRT-comparison between theory and simulation of expected detection delay.

Let

$$\xi^* = \max_{\theta \in \Theta} \log \left[\frac{g_{\mu_1}(Z_1+\theta)}{g_{-\mu_0}(Z_1+\theta)} \right], \tag{10}$$

where $\Theta = \{k_1 b_1 + k_0 b_0, k_1, k_0 = 0, 1, \ldots, L,\ k_1 + k_0 \leq L\}$. Also let $\{\xi_k^*, k \geq 1\}$ be i.i.d. with the distribution of ξ^*. Then for $\alpha > 1$, $E[(\xi^*)^{\alpha+1}] < \infty$ if $E[(\log g_{\mu_1}(Z_1+\theta)/g_{-\mu_0}(Z_1+\theta))^{\alpha+1}] < \infty$ for all $\theta \in \Theta$.

Theorem 1. *For all l and for some $\alpha > 1$, let $E_i[|\log f_{1,l}(X_{1,l})/f_{0,l}(X_{1,l})|^{\alpha+1}] < \infty$ and $E[(\xi_1^*)^{\alpha+1}] < \infty$, i=0,1. Then, under H_i,*

$$\overline{\lim_{c \to 0}} \frac{N}{|\log c|} \leq \frac{1}{D_{tot}^i} + M_i \text{ a.s. and in } L_1,$$

where $M_i = C_i/\Delta(\mathcal{A}^i)$, $C_0 = -[1+E[\xi_1^*]/D_{tot}^0]$ and $C_1 = [1+E[\xi_1^*]/D_{tot}^1]$.

Proof. See Appendix A.

Fig. 2 compares the asymptotic upper bounds of E_{DD} in Theorem 1 with the approximations provided in Section 3.2 and simulations. We see that the approximate analysis of Section 3.2 provides much better approximation at threshold values of practical interest in Cognitive Radio. Perhaps this is the reason, the asymptotically optimal schemes do not necessarily provide very good performance at operating points of practical interest.

Next we consider the asymptotics of P_{FA} and P_{MD}. Let

$$R_i = \min_{1 \leq l \leq L} \left(-\log \inf_{t \geq 0} E_i \left[\exp\left(-t \log \frac{f_{i,l(X_{1,l})}}{f_{j,l}(X_{1,l})} \right) \right] \right), \quad j = 1-i.$$

Let G be the distribution of $|\xi_1^*|$. Also let g be the corresponding moment generating function. Let $\Lambda(\alpha) = \sup_\lambda(\alpha\lambda - \log g_i(\lambda))$ and $\alpha^+ = \text{ess sup}|\xi_1^*|$. Let

$$s(\eta) = \begin{cases} \dfrac{\eta}{\alpha^+} & \text{if } \eta \geq \Lambda(\alpha^+), \\ \dfrac{\eta}{\Lambda^{-1}(\eta)} & \text{if } \eta \in (0, \Lambda(\alpha^+)). \end{cases} \tag{11}$$

Theorem 2. *Let $g(\lambda) < \infty$ in a neighborhood of zero. Then,*

(a) $\lim_{c \to 0} P_{FA}/c = 0$ *if for some* $0 < \eta < R_0$, $s(\eta) > 1$.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

(b) $\lim_{c\to 0} P_{MD}/c = 0$ *if for some* $0 < \eta < R_1$, $s(\eta) > 1$.

Proof. See Appendix B.

Remark 1. When $\alpha^+ = \infty$ which is generally true, $\Lambda(\alpha^+) = \infty$ [22] and in Theorem 2(a) and 2(b) we need to consider only $R_i < \Lambda(\alpha^+)$.

Remark 2. In [6, Lemma 1 Appendix A], it is proved that log likelihood ratio converts a large class of distributions into light tailed distributions and then $g(\lambda)$ is finite in a neighborhood of zero.

We compare the asymptotic results obtained in Theorems 1 and 2 with that of SPRT with all the data available at the local nodes centrally without noise. Let N_{ct} be the stopping time of such an SPRT. Then, from [23, Theorem 2.11.1 and 2.11.2],

$$\lim_{c\to 0} \frac{E_i[N_{ct}]}{|\log c|} = \frac{1}{D_{tot}^i} \tag{12}$$

$$\lim_{c\to 0} \frac{\log 1/P_{FA}}{|\log c|} \to 1, \quad \lim_{c\to 0} \frac{\log 1/P_{MD}}{|\log c|} \to 1. \tag{13}$$

Theorem 2 implies the asymptotics (13) on P_{FA} and P_{MD} for DualSPRT. Comparing Theorem 1 with (12), we see that the rates of convergence of DualSPRT are optimal. For the limits to equal, we need M_0 and M_1 to be zero. In Section 4.1 we compute M_0 and M_1 for Gaussian fusion center noise.

We can consider the asymptotic performance in the Bayesian framework also. Then the two hypotheses H_0 and H_1 are assumed to have known prior probabilities π and $1-\pi$ respectively. A cost c (≥ 0) is assigned to each time step taken for decision. Let $W_i > 0, i = 0, 1$ be the cost of falsely rejecting H_i. Then Bayes risk of a test δ with stopping time N is defined as

$$\begin{aligned}\mathcal{R}_c(\delta) = \pi[cE_0(N) + W_0 P_0\{\text{reject } H_0\}]\\ + (1-\pi)[cE_1(N) + W_1 P_1\{\text{reject } H_1\}].\end{aligned} \tag{14}$$

Optimizing (14) makes sense even when one does not have prior π (i.e., within the frequentist framework) because then taking π and W_i appropriately, one can think of selecting a decision rule that asymptotically minimizes a weighted sum of $E_i[N]$ and $P_i[\text{reject } H_i], i = 1, 2$.

Let $\mathcal{R}_c(\delta_{cent.})$ and $\mathcal{R}_c(\delta_{DualSPRT})$ be the Bayes's Risk of the optimal centralized SPRT without considering fusion center noise and of DualSPRT respectively. Then [11, p. 2076],

$$\lim_{c\to 0} \frac{\mathcal{R}_c(\delta_{cent.})}{c|\log c|} = \left(\frac{\pi}{D_{tot}^0} + \frac{1-\pi}{D_{tot}^1}\right).$$

From Theorems 1, 2(a) and 2(b), using (14), for DualSPRT with fusion center noise,

$$\lim_{c\to 0} \frac{\mathcal{R}_c(\delta_{DualSPRT})}{c|\log c|} = \left(\frac{\pi}{D_{tot}^0} + \frac{1-\pi}{D_{tot}^1} + C\right),$$

where $C = M_0\pi + M_1(1-\pi)$. The constant C can be made arbitrarily small by making M_0 and M_1 small.

4.1. Example-Gaussian distribution

In the following we apply Theorems 1 and 2 when the fusion center noise is Gaussian $\mathcal{N}(0, \sigma_{FC}^2)$. We take $\mu_1 = \mu_0 = \mu > 0$ and $b_1 = -b_0 = b > 0$. For Theorem 1, $\Delta(\mathcal{A}^0) = -2\mu Lb/\sigma_{FC}^2$ and $\Delta(\mathcal{A}^1) = 2\mu Lb/\sigma_{FC}^2$. Therefore M_0 and M_1 in Theorem 1 $\to 0$ if $L \to \infty$ and/or $b \to \infty$. This also happens if $\sigma_{FC}^2 \to 0$.

Using Remark 1, the condition in Theorem 2(a) is $\sigma_{FC}^2 \eta/(4\mu^2\sqrt{2\eta} + 2\mu Lb) > 1$ for some $0 < \eta < R_0$ and that for Theorem 2(b) is the same for some $0 < \eta < R_1$. Combining these two, it is sufficient to satisfy the later condition with $0 < \eta < \min(R_0, R_1)$. For Gaussian input observations at the local nodes, assuming $f_{1,l} = f_1, f_{0,l} = f_0$ for $1 \leq l \leq L$, we get $\delta_{i,l} = \delta_i$ and $\rho_{i,l} = \rho_i$, $R_i = \delta_i^2/2\rho_i^2$.

Remark 3. From the analysis provided in Section 3.2, it is possible to provide beforehand, atleast approximately, the set of values for thresholds to achieve desired error probabilities and these can be used to design the test. Alternatively the asymptotic analysis provided in this section can also be used to design the parameters. Now we fix L. The conditions in Theorem 2 provide bounds for the choice of μ and b (e.g., in the above Gaussian example it provides upper-bounds). The thresholds $\gamma_{0,l}, \gamma_{1,l}, \beta_0$ and β_1 are taken as $-r_l|\log c|, \rho_l|\log c|, -|\log c|$ and $|\log c|$ respectively as functions of c complying to the analysis in this section. According to Theorem 2, for lower values of c, c can be considered as an upper bound for P_{FA} and P_{MD}. Such a c for meeting the specified P_{FA} and P_{MD} can then be used to calculate the threshold values.

5. Improved decentralized sequential tests: SPRT-CSPRT

This section considers some improvements over Dual SPRT. The performance of the improved algorithms is compared with existing decentralized schemes.

New algorithms: SPRT-CSPRT and DualCSPRT. In DualSPRT presented in Section 3.1, observations $\{Y_k\}$ to the fusion center are not always identically distributed. Till the first transmission from secondary nodes, these observations come from i.i.d. noise distribution, but not after that. Since the non-asymptotic optimality of SPRT is known for i.i.d. observations only [4], using SPRT at the fusion center is not optimal.

We improve DualSPRT with the following modifications. Steps (1)–(3) (corresponding to the algorithm run at the local nodes) are same as in DualSPRT. The steps (4) and (5) are replaced by:

4. Fusion center runs two algorithms:

$$F_k^1 = (F_{k-1}^1 + \log[g_{\mu_1}(Y_k)/g_Z(Y_k)])^+, \quad F_0^1 = 0, \tag{15}$$

$$F_k^0 = (F_{k-1}^0 + \log[g_Z(Y_k)/g_{-\mu_0}(Y_k)])^-, \quad F_0^0 = 0, \tag{16}$$

   where $(x)^+ = \max(0, x)$, $(x)^- = \min(0, x)$, μ_1 and μ_0 are positive constants, g_Z is the pdf of i.i.d. noise $\{Z_k\}$ at the fusion center and g_μ is the pdf of $\mu + Z_k$.
5. The fusion center decides about the hypothesis at time

$$\inf\{k : F_k^1 \geq \beta_1 \text{ or } F_k^0 \leq -\beta_0\}$$

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

and $\beta_0, \beta_1 > 0$. The decision is H_1 if $F_k^1 \geq \beta_1$ and H_0 if $F_k^0 \leq -\beta_0$.

The following discussion provides motivation for this test.

1. If the SPRT sum defined in (2) goes below zero it delays in crossing the positive threshold β_1. Hence if we keep SPRT sum at zero whenever it goes below zero, it reduces E_{DD}. This happens in CUSUM [24]. Similarly one can use a CUSUM statistic under H_0 also. These ideas are captured in (15) and (16).
2. The proposed test is also capable of reducing false alarms caused by noise Z_k before first transmission at t_1 from the local nodes. For F_k^1 and F_k^0 to move away from zero, the mean of increments should be positive and negative respectively. Let $\widehat{\mu}_k = E[Y_k]$ at time k. Then,

   $$E_{\widehat{\mu}_k}\left[\log \frac{g_{\mu_1}(Y_k)}{g_Z(Y_k)}\right] = D\left(g_{\widehat{\mu}_k} \| g_Z\right) - D\left(g_{\widehat{\mu}_k} \| g_{\mu_1}\right). \qquad (17)$$

   Hence before t_1, positive mean value of increments is not possible. After t_1 under H_1 (assuming the local nodes make correct decisions, the justification for which is provided in Section 3), the mean of increments becomes more positive. Similarly for F_k^0. But in case of DualSPRT, SPRT sum at the fusion center has the increments given by $\log g_{\mu_1}(Y_k)/g_{-\mu_0}(Y_k)$. This is difficult to keep zero only before t_1 and thus creates more errors due to noise Z_k.
3. Even though the problem under consideration is hypothesis testing, this is essentially a change detection problem at the fusion center. The observations at the fusion center have the distribution of noise before t_1 and after t_1 the mean changes. But in our scenario, this is a composite sequential change detection problem with the observations that are not i.i.d. and we look for change in both directions. Thus, it is difficult to use existing algorithms available for sequential change detection. Nevertheless our test (15)–(16) provides a guaranteed performance in this scenario.

We consider one more improvement. When a local Cognitive Radio SPRT sum crosses its threshold, it transmits b_1/b_0. This node transmits till the fusion center SPRT sum crosses the threshold. If it is not a false alarm, then its SPRT sum keeps on increasing (decreasing). But if it is a false alarm, then the sum will eventually move towards the other threshold. Hence instead of transmitting $b_1/\ b_0$ the Cognitive Radio can transmit a higher/lower value in an intelligent fashion. This should improve the performance. Thus we modify step (3) in DualSPRT as

$$\begin{aligned} Y_{k,l} = \sum_{i=1}^{4} & b_i^1 \mathbb{1}\{W_{k,l} \in [\gamma_1 + (i-1)\Delta_1, \gamma_1 + i\Delta_1)\} \\ & + b_i^0 \mathbb{1}\{W_{k,l} \in [-\gamma_0 - (i-1)\Delta_1, -\gamma_0 - i\Delta_0)\}, \qquad (18) \end{aligned}$$

where Δ_1 and Δ_0 are the parameters to be tuned at the Cognitive Radio. $4\Delta_1$ and $4\Delta_0$ are taken as ∞. The drift under H_1 (H_0) is a good choice for Δ_1 (Δ_0).

We call the algorithm with the above two modifications as SPRT-CSPRT (with 'C' as an indication about the motivation from CUSUM).

If we use CSPRT at both the secondary nodes and the fusion center with the proposed quantization methodology (we call it DualCSPRT) it works better as we will show via simulations in Section 5.1. This test permits the local nodes to make decisions faster than normal SPRT due to clipping at zero level, but with more errors. The influence of the increase in number of false transmissions from local nodes is nullified by the repeated use of CUSUM-like statistic at the fusion node with appropriate choice of parameters.

5.1. Performance comparison

Throughout the rest of this section we use $\gamma_{1,l} = \gamma_{0,l} = \gamma$, $\beta_1 = \beta_0 = \beta$ and $\mu_1 = \mu_0 = \mu$ for the simplicity of simulations and analysis.

We compare DualSPRT, SPRT-CSPRT and DualCSPRT via simulations.

We have used the following parameters. There are 5 nodes (L=5) and $f_{0,l} \sim \mathcal{N}(0,1)$, for $1 \leq l \leq L$. Primary to secondary channel gains are 0, -1.5, -2.5, -4 and -6 dB respectively (the corresponding post change means of Gaussian distribution with variance 1 are 1, 0.84, 0.75, 0.63 and 0.5). We assume $Z_k \sim \mathcal{N}(0,5)$ and the mean of increments of DualSPRT and SPRT-CSPRT at the fusion center is taken as $2\mu Y_k$, with μ being 1. We also take $D_0 = D_1 = 0$, $\{b_1^1, b_2^1, b_3^1, b_4^1\} = \{1,2,3,4\}$, $\{b_1^0, b_2^0, b_3^0, b_4^0\} = \{-1,-2,-3,-4\}$ and $b_1 = -b_0 = 1$ (for DualSPRT). Parameters γ and β are chosen from a range of values to achieve a particular P_{FA}. Fig. 5 provides the E_{DD} and P_{MD} via simulations. We see a significant improvement in E_{DD} compared to DualSPRT. The difference increases as P_{MD} decreases. The performance under H_0 is similar.

Performance comparisons with the asymptotically optimal decentralized sequential algorithms which do not consider fusion center noise (DSPRT [12], Mei's SPRT [11]) are given in Fig. 6. Note that DualSPRT and SPRT-CSPRT include fusion center noise. Here we take $f_{0,l} \sim \mathcal{N}(0,1)$, $f_{1,l} \sim \mathcal{N}(1,1)$ for $1 \leq l \leq L$ and $Z_k \sim \mathcal{N}(0,1)$. We find that the performance of SPRT-CSPRT is close to that of DSPRT (which is second order asymptotically optimal) and better than Mei's SPRT (which is first order asymptotically optimal). Similar comparisons were obtained with other data sets.

6. Unknown received SNRs and fading

This section considers the extensions of DualSPRT and SPRT-CSPRT to take care of the SNR uncertainty and the slow fading between the primary user and a Cognitive Radio. Since the transmissions from CR to FC are in CR network, we assume reporting channel to FC as AWGN only. This assumption is commonly made [1,2].

6.1. Different and unknown SNRs

We consider the case where the received signal power from the PU to a CR node is fixed but not known to the local Cognitive Radio nodes. This can happen if the transmit power of the primary is not known and/or there is unknown shadowing. Now we limit ourselves to the

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

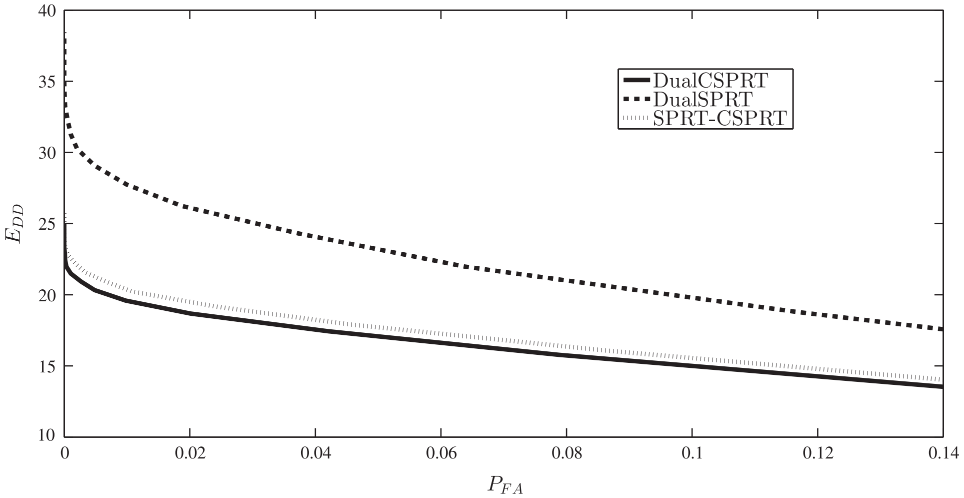

Fig. 5. Comparison among DualSPRT, SPRT-CSPRT and DualCSPRT for different SNR's between the primary and the secondary users, under H_1.

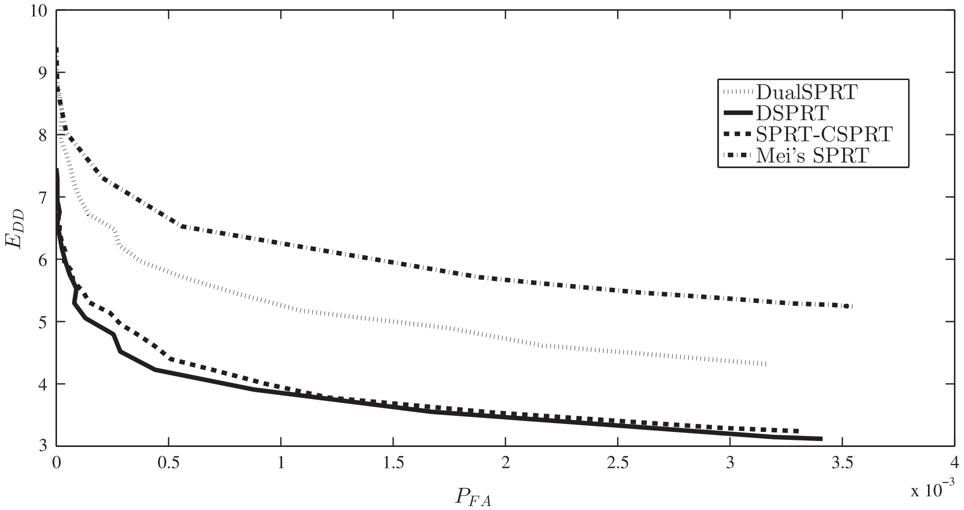

Fig. 6. Comparison among DualSPRT, SPRT-CSPRT, Mei's SPRT and DSPRT under H_1.

energy detector where the observations $X_{k,l}$ are average energy of M samples received by the lth Cognitive Radio node. Then for somewhat large M, the distributions of $X_{k,l}$ under H_0 and H_1 can be approximated by Gaussian distributions: $f_{0,l} \sim \mathcal{N}(\sigma_l^2, 2\sigma_l^4/M)$ and $f_{1,l} \sim \mathcal{N}(P_l+\sigma_l^2, 2(P_l+\sigma_l^2)^2/M)$, where P_l is the received power and σ_l^2 is the noise variance at the lth CR node. Under low SNR conditions $(P_l+\sigma_l^2)^2 \approx \sigma_l^4$ and hence $X_{k,l}$ are Gaussian distributed with mean change under H_0 and H_1. Now taking $X_{k,l} - \sigma_l^2$ as the data for the detection algorithm at the lth node, since P_l is unknown we can formulate this problem as a sequential hypothesis testing problem with

$$H_0: \theta = 0; \; H_1: \theta \geq \theta_1, \tag{19}$$

where θ is P_l under H_1 and θ_1 is appropriately chosen.

The problem

$$H_0: \theta \leq \theta_0; \; H_1: \theta \geq \theta_1, \tag{20}$$

subject to

$$P_\theta\{\text{reject } H_0\} \leq \alpha \text{ for } \theta \leq \theta_0 \text{ and } P_\theta\{\text{reject } H_1\} \leq \beta \text{ for } \theta \geq \theta_1,$$

for exponential family of distributions is well studied in [25]. The following algorithm of Lai [25] is asymptotically Bayes optimal and hence we use it at the local nodes instead of SPRT. Let $\theta \in A = [a_1, a_2]$. Define

$$W_{n,l} = \max\left[\sum_{k=1}^{n} \log\frac{f_{\hat{\theta}_n}(X_k)}{f_{\theta_0}(X_k)}, \sum_{k=1}^{n} \log\frac{f_{\hat{\theta}_n}(X_k)}{f_{\theta_1}(X_k)}\right],$$

$$N_l(g,c) = \inf\{n : W_{n,l} \geq g(nc)\},$$

where $g()$ is a time varying threshold and $c > 0$ is a design parameter. The function g satisfies $g(t) \approx \log(1/t)$ as $t \to 0$ and is the boundary of an associated optimal stopping problem for the Wiener process [25]. $\hat{\theta}_n$ is the Maximum-Likelihood estimate of θ bounded by a_1 and a_2. For Gaussian f_0 and f_1, $\hat{\theta}_n = \max\{a_1, \min[S_n/n, a_2]\}$. At time

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

$N_l(g,c)$ decide upon H_0 or H_1 according as $\hat{\theta}_{N_l(g,c)} \le \theta^*$ or $\hat{\theta}_{N_l(g,c)} \ge \theta^*$, where θ^* is obtained by solving $D(f_{\theta^*} \| f_{\theta_0}) = D(f_{\theta^*} \| f_{\theta_1})$.

For our case where $H_0 : \theta = 0$, unlike in (20) where $H_0 : \theta \le 0$, $E_0[N_l(g,c)]$ largely depends upon the value θ_1. As θ_1 increases, $E_0[N_l(g,c)]$ decreases and $E_1[N_l(g,c)]$ increases. If $P_l \in [\underline{P}, \overline{P}]$ for all l then a good choice of θ_1, is $(\overline{P} - \underline{P})/2$.

6.1.1. GLR-SPRT

First we modify DualSPRT. In the distributed setup with the received power at the local nodes unknown, the local nodes will use the Lai's algorithm mentioned above while the fusion node runs the SPRT. All other details remain same. We call this algorithm GLR-SPRT.

6.1.2. GLR-CSPRT

This is a modified version of SPRT-CSPRT. Here, we modify GLR-SPRT to GLR-CSPRT with appropriate change in quantization and using CSPRT at the fusion center instead of SPRT. The quantization (18) is changed in the following way: if $\hat{\theta}_N \ge \theta^*$, let $\mathcal{I}_1 = [g(kc), g(kc3\,\Delta))$, $\mathcal{I}_2 = [g(kc3\Delta), g(kc2\,\Delta))$, $\mathcal{I}_3 = [g(kc2\Delta), g(kc\Delta))$ and $\mathcal{I}_4 = [g(kc\Delta), \infty)$. $Y_{k,l} = b_n^1$ if $W_{k,l} \in \mathcal{I}_n$ for some n. If $\hat{\theta}_N \le \theta^*$ we will transmit from $\{b_1^0, b_2^0, b_3^0, b_4^0\}$ under the same conditions. Here, Δ is a tuning parameter and $0 \le 3\Delta \le 1$.

The performance comparison of GLR-SPRT and GLR-CSPRT for the example in Section 5.1 (with $Z_k \sim \mathcal{N}(0,1)$) is given in Figs. 7 and 8. Here $\Delta = 0.25$. As the performance under H_1 and H_0 are different, we give the values under both. We can see that GLR-SPRT is always inferior to GLR-CSPRT. For E_{DD} under H_1, interestingly GLR-CSPRT has lesser values than that of SPRT-CSPRT for $P_{FA} > 0.02$ (note that SPRT-CSPRT has complete knowledge of the SNRs), while under H_0 it has higher values than SPRT-CSPRT.

6.2. Channel with fading

In this section we consider the system where the channels from the primary transmitter to the secondary

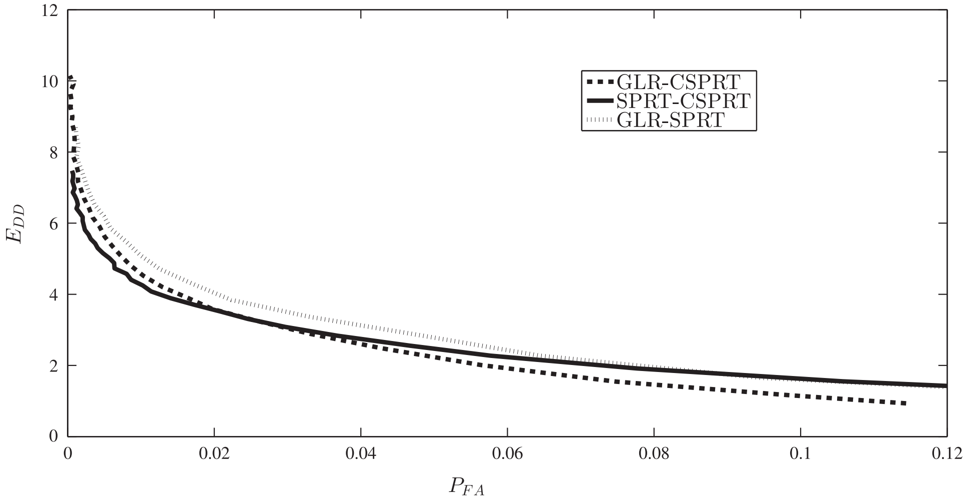

Fig. 7. Comparison among SPRT-CSPRT, GLR-SPRT and GLR-CSPRT for different SNR's between the primary and the secondary users under H_1.

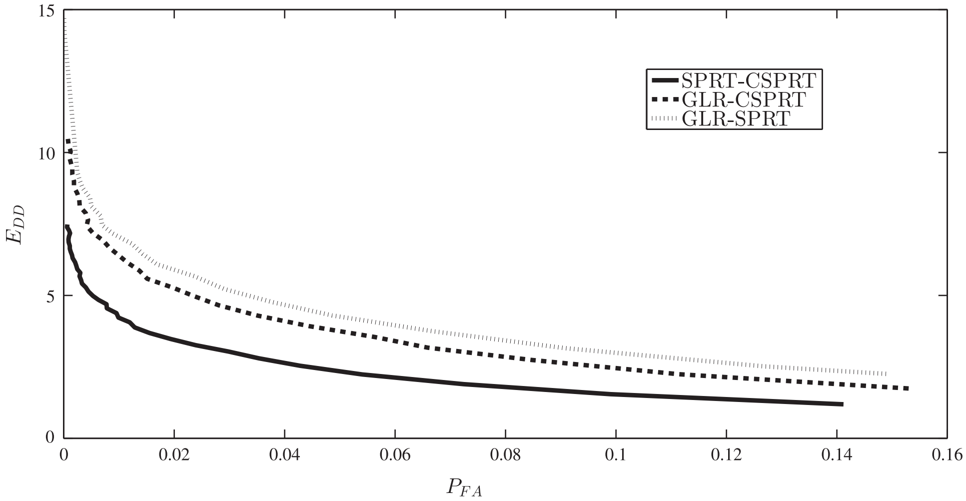

Fig. 8. Comparison among SPRT-CSPRT, GLR-SPRT and GLR-CSPRT for different SNR's between the primary and the secondary users under H_0.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

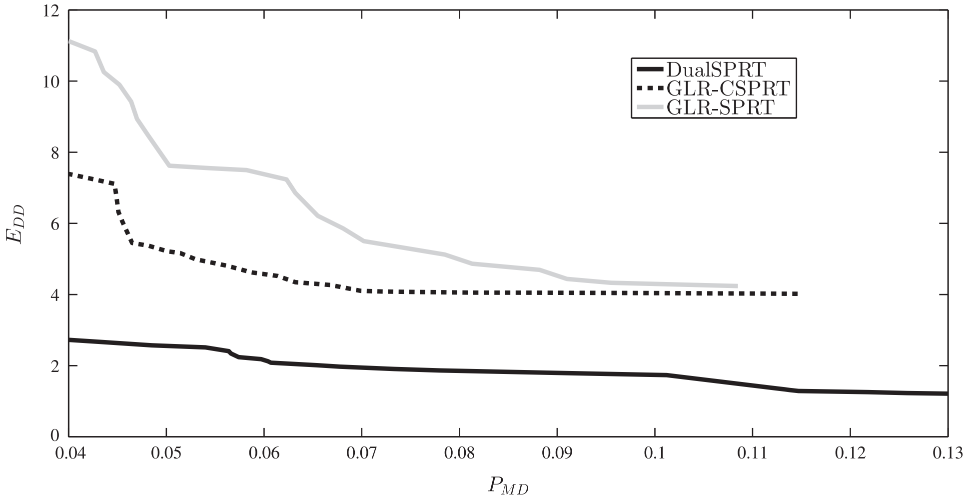

Fig. 9. Comparison among DualSPRT, GLR-SPRT and GLR-CSPRT with slow fading between the primary and the secondary users under H_1.

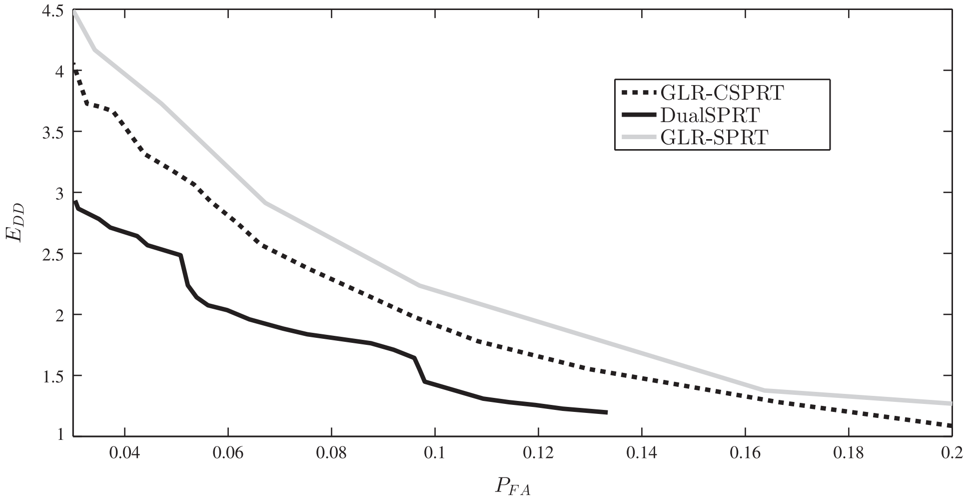

Fig. 10. Comparison among DualSPRT, GLR-SPRT and GLR-CSPRT with slow fading between the primary and the secondary users under H_0.

nodes have fading ($h_l \neq 1$). We assume slow fading, i.e., the channel coherence time is longer than the hypothesis testing time.

When the fading gain h_l is known to the lth secondary node then this case can be considered as the different SNR case as in the example given in Section 3.2.4. Thus we consider the case where the channel gain h_l is not known to the lth node.

We consider the energy detector setup of Section 6.1. However, now P_l, the received signal power at the local node l is random. If the fading is Rayleigh distributed then P_l has exponential distribution. The hypothesis testing problem becomes

$$H_0 : f_{0,l} \sim \mathcal{N}(0, \sigma^2); H_1 : f_{1,l} \sim \mathcal{N}(\theta, \sigma^2) \tag{21}$$

where θ is random with exponential distribution and σ^2 is the variance of noise. We will assume that σ^2 is known at the nodes.

We are not aware of this problem being handled via sequential hypothesis testing before. However we use Lai's algorithm in Section 6.1 where we take θ_1 to be the median of the distribution of θ, i.e., $P(\theta \geq \theta_1) = 1/2$ or mean of θ. These seem good choices for θ_1 as a compromise between $E_0[N]$ and $E_1[N]$.

We apply the technique on GLR-SPRT and GLR-CSPRT. We use an example where $\sigma^2 = 1, \theta \sim \exp(1)$, $\mathrm{Var}(Z_k) = 1$, $L = 5$ and θ_1 as the mean of θ. The performance of these algorithms are compared with that of DualSPRT (with perfect channel state information) in Figs. 9 and 10. The results are comparable even though the primary is completely unknown except the knowledge of fading distribution.

7. Conclusions

This paper presents fast algorithms for cooperative spectrum sensing satisfying reliability constraints. We

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

have presented and analyzed DualSPRT, a decentralized sequential hypothesis test. Simulation results corroborate the theoretical study of DualSPRT. Asymptotic properties of DualSPRT are also explored and its performance can approach asymptotically Bayes optimal tests. Improvement over DualSPRT using CUSUM statics for the fusion center test leads to another algorithm in which the selection of parameters is easy to choose apart from performance enhancement. Numerical experiments show that this algorithm performs as well as an asymptotic order-2 optimal algorithm without fusion center noise, proposed in the literature. We further extend our algorithms to cover the case of unknown SNR and channel fading and obtain satisfactory performance compared to perfect channel state information case.

Appendix A. Proof of Theorem 1

We will prove the theorem under H_0. The proof under H_1 will follow in the same way.

Let $\nu(a)$ be the stopping time when a random walk starting at zero and formed by the sequence $\{\log g_{\mu_1(Z_k)}/g_{-\mu_0(Z_k)}+ \ (\Delta(\mathcal{A}^0)-E_0[\log g_{\mu_1(Z_k)}/g_{-\mu_0(Z_k)}]), \ k \geq \tau(c)+1\}$ (with $-$ ve drift under H_0) crosses a. Then,

$$N \leq N^0 \leq \tau(c)+\nu(-|\log c|-F_{\tau(c)+1}).$$

Therefore,

$$\frac{N}{|\log c|} \leq \frac{\tau(c)}{|\log c|}+\frac{\nu(-|\log c|-F_{\tau(c)+1})}{|\log c|}. \tag{A.1}$$

We consider the first term on the R.H.S. of (A.1). From [19, Remark 4.4, p. 90] as $c \to 0$, $\tau_l(c) \to \infty$ a.s. and $\lim_{c\to 0}\tau_l(c)/|\log c| = -r_l/\delta_{0,l} = 1/D^0_{tot}$ a.s. Therefore,

$$\frac{\tau(c)}{|\log c|} \to \max_l \left\{ -\frac{r_l}{\delta_{0,l}} \right\} = \frac{1}{D^0_{tot}} \quad \text{a.s.} \tag{A.2}$$

Furthermore, from [26, proof of Theorem 1(i) ⇒ (ii) p. 871], it can be seen that $\{\tau_l(c)/|\log c|\}$ is uniformly integrable for each l. Therefore, $\{\tau(c)/|\log c|\}$ is also uniformly integrable and hence,

$$\frac{E_0[\tau(c)]}{|\log c|} \to \frac{1}{D^0_{tot}}. \tag{A.3}$$

The second term in R.H.S. of (A.1)

$$\frac{\nu(-|\log c|-F_{\tau(c)+1})}{|\log c|} \leq \frac{\nu(-|\log c|)}{|\log c|}+\frac{\nu(-F_{\tau(c)+1})}{|\log c|}. \tag{A.4}$$

We know, from [19, Chapter III], as $c \to 0$

$$\frac{\nu(-|\log c|)}{|\log c|} \to -\frac{1}{\Delta(\mathcal{A}^0)} \quad \text{a.s. and in } L_1. \tag{A.5}$$

Next consider $\nu(-F_{\tau(c)+1})$. Let F^*_k be a random walk formed from ξ^*_k. It can be shown that F^*_k stochastically dominates F_k and thus we can make $F^*_k \geq F_k$ a.s. for all $k \geq 0$. Then,

$$\frac{\nu(-F_{\tau(c)+1})}{|\log c|} \leq \frac{\nu(-F^*_{\tau(c)+1})}{|\log c|}.$$

Also,

$$\frac{F^*_{\tau(c)+1}}{|\log c|} = \frac{F^*_{\tau(c)+1}}{\tau(c)+1}\frac{\tau(c)+1}{|\log c|} \to E\left[\xi^*_1\right]\frac{1}{D^0_{tot}} \quad \text{a.s.}$$

Thus,

$$\frac{\nu(-F^*_{\tau(c)+1})}{|\log c|} = \frac{\nu(-F^*_{\tau(c)+1})}{F^*_{\tau(c)+1}}\frac{F^*_{\tau(c)+1}}{|\log c|} \to \frac{-1}{\Delta(\mathcal{A}^0)}\frac{E[\xi^*_1]}{D^0_{tot}} \quad \text{a.s.} \tag{A.6}$$

From (A.1) to (A.6),

$$\overline{\lim_{c\to 0}}\frac{N}{|\log c|} \leq \frac{1}{D^0_{tot}} - \frac{1}{\Delta(\mathcal{A}^0)}\frac{E[\xi^*_1]}{D^0_{tot}} \quad \text{a.s.}$$

Now we show L_1 convergence. For $\alpha > 1$,

$$\begin{aligned}\frac{E_0[\nu(-F^*_{\tau(c)+1})^\alpha]}{|\log c|^\alpha} &= \frac{1}{|\log c|^\alpha}\int_0^{|\log c|} E_0\left[\nu(-x)^\alpha \middle| F^*_{\tau(c)+1}=x\right]\, dP_{F^*_{\tau(c)+1}}(x) \\ &\quad + \frac{1}{|\log c|^\alpha}\int_{|\log c|}^{\infty} E_0\left[\nu(-x)^\alpha\right]\, dP_{F^*_{\tau(c)+1}}(x) \\ &\leq \frac{E_0[\nu(-|\log c|)^\alpha]}{|\log c|^\alpha} + \int_{|\log c|}^{\infty}\frac{E_0[\nu(-x)^\alpha]}{x^\alpha}\frac{x^\alpha}{|\log c|^\alpha}\, dP_{F^*_{\tau(c)+1}}(x).\end{aligned} \tag{A.7}$$

When $-$ve part of the increments of random walk of $\nu(t)$ has finite αth moment [19, Chapter 3, Theorem 7.1], $E_0[\nu(-x)^\alpha]/x^\alpha \to (-1/\Delta(\mathcal{A}^0))^\alpha$ as $x \to \infty$. Thus for any $\epsilon > 0$, $\exists$ M such that

$$\frac{E_0[\nu(-x)^\alpha]}{x^\alpha} \leq \left(\epsilon + \left(\frac{-1}{\Delta(\mathcal{A}^0)}\right)\right)^\alpha \quad \text{for } x > M.$$

Take c_1 such that $|\log c| > M$ for $c < c_1$. Then, for $c < c_1$,

$$\int_{|\log c|}^{\infty}\frac{E_0[\nu(-x)^\alpha]}{x^\alpha}\frac{x^\alpha}{|\log c|^\alpha}\, dP_{F^*_{\tau(c)+1}}(x) \leq \frac{\epsilon + \left(\frac{-1}{\Delta(\mathcal{A}^0)}\right)^\alpha}{|\log c|^\alpha}$$

$$\int_{|\log c|}^{\infty} x^\alpha\, dP_{F^*_{\tau(c)+1}}(x) \leq \frac{\epsilon + \left(\frac{-1}{\Delta(\mathcal{A}^0)}\right)^\alpha}{|\log c|^\alpha}, \quad E_0\left[(F^*_{\tau(c)+1})^\alpha\right]. \tag{A.8}$$

Since $\lim_{c\to 0}\tau(c)/|\log c| = 1/D^0_{tot}$ a.s. and $\{\tau(c)^\alpha/|\log c|^\alpha\}$ is uniformly integrable, when $E_0[(\log f_{1,l}(X_{1,l})/f_{0,l}(X_{1,l}))^{\alpha+1}] < \infty$, $1 \leq l \leq L$ and $E[(\xi^*_1)^{\alpha+1}] < \infty$, we get, [19, Remark 7.2, p. 42],

$$\lim_{c\to 0}\frac{E_0[(F^*_{\tau(c)+1})^\alpha]}{|\log c|^\alpha} = \frac{E[(\xi^*_1)^\alpha]}{D^0_{tot}},$$

and

$$\sup_{c>0}\frac{E_0[(F^*_{\tau(c)+1})^\alpha]}{|\log c|^\alpha} < \infty. \tag{A.9}$$

From (A.7) and (A.9), for some $1 > \delta > 0$,

$$\begin{aligned}\sup_{\delta > c > 0}\frac{E_0[\nu(-F^*_{\tau(c)+1})^\alpha]}{|\log c|^\alpha} &\leq \sup_{\delta > c > 0}\frac{E_0[\nu(-|\log c|)^\alpha]}{|\log c|^\alpha} \\ &\quad + \left[\epsilon + \left(\frac{-1}{\Delta(\mathcal{A}^0)}\right)^\alpha\right]\sup_{\delta > c > 0}\frac{E_0[(F^*_{\tau(c)+1})^\alpha]}{|\log c|^\alpha} < \infty.\end{aligned}$$

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

Therefore, $\{\nu(-F^*_{\tau(c)+1})/|\log c|\}$ is uniformly integrable and hence, from (A.6),

$$\lim_{c\to 0}\frac{E_0[\nu(-F^*_{\tau(c)+1})]}{|\log c|}\le -\frac{1}{\Delta(\mathcal{A}^0)}\cdot\frac{E[\xi^*_1]}{D^0_{tot}}.$$

This, with (A.1), (A.3), (A.4) and (A.5), implies that (since ϵ can be taken arbitrarily small),

$$\overline{\lim_{c\to 0}}\frac{E_0[N]}{|\log c|}\le\frac{1}{D^0_{tot}}+M_0,$$

where

$$M_0=-\frac{1}{\Delta(\mathcal{A}^0)}\left[1+\frac{E[\xi^*_1]}{D^0_{tot}}\right].$$

Similarly we can prove

$$\overline{\lim_{c\to 0}}\frac{E_1[N]}{|\log c|}\le\frac{1}{D^1_{tot}}+M_1,$$

where

$$M_1=\frac{1}{\Delta(\mathcal{A}^1)}\left[1+\frac{E[\xi^*_1]}{D^1_{tot}}\right].$$

Appendix B. Proof of Theorem 2

We prove the result for P_{FA}. For P_{MD} it can be proved in the same way.

Probability of False Alarm can be written as

$$P_0(\text{Reject } H_0)=P_0[\text{FA upto } \tau(c)]+P_0[\text{FA after } \tau(c)]. \tag{B.1}$$

Consider the first term in the R.H.S. of (B.1). Take F^*_k as in the proof of Theorem 1 with $F^*_k\ge F_k$ a.s. for all $k\ge 0$ and hence

$$\begin{aligned}P_0[\text{FA upto } \tau(c)]&\le P_0\left[\sup_{0\le k\le\tau(c)} F^*_k\ge|\log c|\right]\\&=P_0\left[\sum_{k=0}^{\tau(c)}|\xi^*_k|\ge|\log c|\right].\end{aligned} \tag{B.2}$$

From [27, Theorem 1.3], for $0<\eta<R^l_0=-\log\inf_{t\ge 0}E_0[e^{-t\log f_{1,l(X_{1,l})}/f_{0,l}(X_{1,l})}]$, $E_0[e^{\eta\,\tau_l(c)}]<\infty$. Combining this fact with $\tau(c)<\sum_{l=1}^L\tau_l(c)$ and the fact that $\tau_l(c)$ are independent of each other (see (9)) yields $E_0[e^{\eta\tau(c)}]<E_0[e^{\sum_{l=1}^L\eta\tau_l(c)}]<\infty$, for $0<\eta<R_0=\min_l R^l_0$. Therefore, from Markov inequality, with $k_1=E_0[e^{\eta\tau(c)}]$,

$$P[\tau(c)>t]\le k_1\exp(-\eta t). \tag{B.3}$$

Let $\widehat{F}^*_n=\sum_{k=1}^n|\xi^*_k|$. Then, with (B.3), the expected value of $|\xi^*_k|$ being positive and with exponential tail assumption of $G(t)$, from [22, Theorem 1, Remark 1], (B.2) is

$$P_0[\widehat{F}^*_{\tau(c)}>|\log c|]\le k_2\exp(-s(\eta)|\log c|), \tag{B.4}$$

for any $0<\eta<R_0$ where k_2 is a constant and $s(\eta)$ is defined in (11). Therefore,

$$\frac{P_0[\text{FA upto } \tau(c)]}{c}\le k_2\frac{c^{s(\eta)}}{c}\to 0, \tag{B.5}$$

if $s(\eta)>1$ for some η.

Now we consider the second term in (B.1),

$$P_0[\text{FA after } \tau(c)]=P_0[\text{FA after } \tau(c);\ \mathcal{A}^0]+P_0[\text{FA after } \tau(c);\ (\mathcal{A}^0)^c]$$

Since events {FA after $\tau(c)$} and $(\mathcal{A}^0)^c$ are mutually exclusive, the second term in the above expression is zero. Now consider $P_0[\text{FA after } \tau(c);\ \mathcal{A}^0]$. For $0<r<1$,

$$\begin{aligned}&P_0[\text{FA after } \tau(c);\ \mathcal{A}^0]\\&\le P_0\Big[\text{Random walk with drift } \Delta(\mathcal{A}^0) \text{ and initial value } F_{\tau(c)+1}\\&\qquad\text{crosses } |\log c|\Big]\\&\le P_0\Big[\text{Random walk with drift } \Delta(\mathcal{A}^0) \text{ and } F_{\tau(c)+1}\\&\qquad\le r|\log c| \text{ crosses } |\log c|\Big]\\&\qquad+P_0\Big[\text{Random walk with drift } \Delta(\mathcal{A}^0) \text{ and } F_{\tau(c)+1}\\&\qquad>r|\log c| \text{ crosses } |\log c|\Big]\\&\le P_0\Big[\text{Random walk with drift } \Delta(\mathcal{A}^0) \text{ and } F_{\tau(c)+1}\\&\qquad\le r|\log c| \text{ crosses } |\log c|\Big]\\&\qquad+P_0[F_{\tau(c)+1}>r|\log c|].\end{aligned} \tag{B.6}$$

Considering the first term in the above expression,

$$\begin{aligned}&\Big(P_0\Big[\text{Random walk with drift } \Delta\big(\mathcal{A}^0\big) \text{ and } F_{\tau(c)+1}\\&\qquad\le r\big|\log c\big| \text{ crosses } \big|\log c|\big]\Big)\Big/c\\&\le\Big(P_0\Big[\text{Random walk with drift } \Delta\big(\mathcal{A}^0\big) \text{ and } F_{\tau(c)+1}\\&\qquad=r\big|\log c\big| \text{ crosses } \big|\log c|]\Big)\Big/c\\&\overset{(A)}{\le}\frac{\exp(-(1-r)|\log c|s')}{c}=\frac{c^{(1-r)s'}}{c}\to 0,\end{aligned} \tag{B.7}$$

if $(1-r)s'>1$. Here (A) follows from [28, p. 7879],[1] where s' is positive and it is the solution of $E_0[e^{s'\log g_{\mu_1}(Y_k)/g_{-\mu_0}(Y_k)}|\mathcal{A}^0]=1$. We choose $s'>1$ and $0<r<1$ to satisfy $(1-r)s'>1$.

Consider the second term in (B.6). Using the stochastic dominance of $\{F_k\}$ by $\{\widehat{F}^*_k\}$,

$$P_0[F_{\tau(c)+1}>r|\log c|]\le P_0[\widehat{F}^*_{\tau(c)+1}>r|\log c|].$$

We have $P[\tau(c)+1>t]=P[\tau(c)>t-1]\le k'_1\exp(-\eta t)$, where $k'_1=e^{\eta}E_0[e^{\eta\tau(c)}]$. Therefore, following (B.4),

$$\frac{P_0[F_{\tau(c)+1}>r|\log c|]}{c}\le k'_2\frac{c^{rs(\eta)}}{c}\to 0,$$

if $rs(\eta)>1$ and k'_2 is a constant. We can choose $s(\eta)>1$ as in (B.5). Then $1/s(\eta)<r\le 1-1/s'$.

References

[1] I. Akyildiz, B. Lo, R. Balakrishnan, Cooperative spectrum sensing in cognitive radio networks: a survey, Phys. Commun. 4 (1) (2011) 40–62.

---

[1] For a random walk $W_n=\sum_{i=1}^n X_i$, with stopping times $T_a=\inf\{n\ge 1: W_n\le a\}$, $T_b=\inf\{n\ge 1: W_n\ge b\}$ and $T_{a,b}=\min(T_a,T_b)$, $a<0<b$, let s' be the non-zero solution to $M(s')=1$, where M denotes the M.G.F. of X_i. Then, $s'<0$ if $E[X_i]>0$, and $s'>0$ if $E[X_i]<0$ and $E[\exp(s'W_{T_{a,b}})]=1$ [28, p. 7879]. Then it can be shown that $P(W_{T_a})\le\exp(-s'a)$ when $E[X_i]>0$ and $P(W_{T_b})\le\exp(-s'b)$ when $E[X_i]<0$.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009

[2] J. Unnikrishnan, V.V. Veeravalli, Cooperative sensing for primary detection in cognitive radio, IEEE J. Sel. Top. Signal Process. 2 (1) (2008) 18–27.

[3] Z. Quan, S. Cui, H. Poor, A. Sayed, Collaborative wideband sensing for cognitive radios, IEEE Signal Process. Mag. 25 (6) (2008) 60–73.

[4] D. Siegmund, Sequential Analysis: Tests and Confidence Intervals, Springer-Verlag, Berlin, Heidelberg, 1985.

[5] H. Li, H. Dai, C. Li, Collaborative quickest spectrum sensing via random broadcast in cognitive radio systems, IEEE Trans. Wirel. Commun. 9 (7) (2010) 2338–2348.

[6] T. Banerjee, V. Sharma, V. Kavitha, A.K. JayaPrakasam, Generalized analysis of a distributed energy efficient algorithm for change detection, IEEE Trans. Wirel. Commun. 10 (1) (2011) 91–101.

[7] Q. Zou, S. Zheng, A.H. Sayed, Cooperative sensing via sequential detection, IEEE Trans. Signal Process. 58 (12) (2010) 6266–6283.

[8] Y. Yilmaz, G. Moustakides, X. Wang, Spectrum sensing via event-triggered sampling, in: Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2011, pp. 1420–1424.

[9] Y. Shei, Y.T. Su, A sequential test based cooperative spectrum sensing scheme for cognitive radios, in: IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2008.

[10] V.V. Veeravalli, Sequential decision fusion: theory and applications, J. Frankl. Inst. 336 (2) (1999) 301–322.

[11] Y. Mei, Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks, IEEE Trans. Inf. Theory 54 (5) (2008) 2072–2089.

[12] G. Fellouris, G.V. Moustakides, Decentralized sequential hypothesis testing using asynchronous communication, IEEE Trans. Inf. Theory 57 (1) (2011) 534–548.

[13] Y. Yilmaz, G.V. Moustakides, X. Wang, Channel-aware decentralized detection via level-triggered sampling, IEEE Trans. Signal Process. 61 (2) (2013) 300–315.

[14] J. Chamberland, V.V. Veeravalli, Wireless sensors in distributed detection applications, IEEE Signal Process. Mag. 24 (3) (2007) 16–25.

[15] F. Li, J. S. Evans, Optimal strategies for distributed detection over multiaccess channels, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2008, pp. 2417–2420.

[16] Y. Lin, B. Chen, L. Tong, Distributed detection over multiple access channels, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 3, 2007, pp. III–541–III–544.

[17] M.K. Banavar, C. Tepedelenlioglu, A. Spanias, Estimation over fading channels with limited feedback using distributed sensing, IEEE Trans. Signal Process. 58 (1) (2010) 414–425.

[18] M.K. Banavar, C. Tepedelenlioglu, A. Spanias, Distributed SNR estimation with power constrained signaling over gaussian multiple-access channels, IEEE Trans. Signal Process. 60 (6) (2012) 3289–3294.

[19] A. Gut, Stopped Random Walks: Limit Theorems and Applications, 2nd Edition, Springer, 2009.

[20] H.M. Barakat, Y.H. Abdelkader, Computing the moments of order statistics from nonidentical random variables, Stat. Methods Appl. 13 (1) (2004) 15–26.

[21] P. Billingsley, Probability and Measure, 2nd Edition, John Wiley & Sons, 1986.

[22] A.A. Borovkov, Unimprovable exponential bounds for distributions of sums of a random number of random variables, Theory Probab. Appl. 40 (2) (1995) 230–237.

[23] Z. Govindarajulu, Sequential Statistics, World Scientific Pub Co Inc, 2004.

[24] E. Page, Continuous inspection schemes, Biometrika 41 (1/2) (1954) 100–115.

[25] T.L. Lai, Nearly optimal sequential tests of composite hypotheses, Ann. Stat. (1988) 856–886.

[26] S. Janson, Moments for first-passage and last-exit times, the minimum, and related quantities for random walks with positive drift, Adv. Appl. Probab. (1986) 865–879.

[27] A. Iksanov, M. Meiners, Exponential moments of first passage times and related quantities for random walks, Electron. Commun. Probab. 15 (2010) 365–375.

[28] H.V. Poor, O. Hadjiliadis, Quickest Detection, Cambridge University Press, 2008.

Please cite this article as: J.K. Sreedharan, V. Sharma, Spectrum sensing using distributed sequential detection via noisy reporting MAC, Signal Processing (2014), http://dx.doi.org/10.1016/j.sigpro.2014.07.009