Distributed Nonparametric Sequential Spectrum Sensing under Electromagnetic Interference*

Sahasranand K. R. and Vinod Sharma
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore, India
Email: {sanandkr, vinod}@ece.iisc.ernet.in

ABSTRACT

We propose a distributed sequential algorithm for quick detection of spectral holes in a Cognitive Radio set up. Two or more local nodes make decisions and inform the fusion centre (FC) over a reporting Multiple Access Channel (MAC), which then makes the final decision. The local nodes use energy detection and the FC uses mean change detection in the presence of fading, heavy-tailed electromagnetic interference (EMI) and outliers. The statistics of the primary signal, channel gain and the EMI is not known. Different nonparametric sequential algorithms are compared to choose appropriate algorithms to be used at the local nodes and the FC. Modification of a recently developed random walk test is selected for the local nodes for energy detection as well as at the fusion centre for mean change detection. We show via simulations and analysis that the nonparametric distributed algorithm developed performs well in the presence of fading, EMI and outliers. The algorithm is iterative in nature making the computation and storage requirements minimal.

Key words: Nonparametric tests, sequential detection, distributed detection, energy detector, electromagnetic interference, heavy-tailed distributions, shadowing-fading, outliers.

I. INTRODUCTION

The need for reusing the sparsely available spectrum has motivated the widely studied paradigm of Cognitive Radio (CR) [1]. A spectrum remains idle when the primary user (licensee) is not using it. The secondary nodes need to quickly detect this spectral hole and make use of it for data transmission during this interval and stop transmitting once the primary starts transmitting. This is called Spectrum sensing in the CR scenario. Spectrum sensing is performed in a wide variety of ways depending upon the knowledge of the primary signalling and the channel gains [2], [3].

Detection of spectral holes has to be performed at very low SNRs (~ 20 dB) in the presence of shadowing and fading [4]. There is also a need to detect the presence of holes as early as possible to make efficient use of idle channel and to minimize interference to the primary users. Hence sequential procedures serve better which can reduce the expected number of samples required, by more than half, over the fixed sample procedures [5]. This also calls for distributed detection which exploits spatial diversity to mitigate fading and also can reduce detection time [2], [3]. Furthermore, the transmit power, channel gains, coding and modulations of the primary are unknown and hence standard algorithms such as matched filter or cyclostationarity detector [3] may not be available. Energy detection (or generalised energy detection [6]) is found to be the technique applicable in such scenarios. Lack of complete knowledge about the signal and the channel fading (shadowing) forces us to use nonparametric (or semiparametric) detection algorithms. Besides, the distribution of SINR may not be known and noise power could be time varying due to time varying electromagnetic interference (EMI). EMI is modelled using heavy-tailed distributions (7), (8) and outliers [9] could be present in the samples received at the local nodes as well as the fusion centre (FC) over a reporting Multiple Access Channel (MAC). Channel fading can have Rayleigh, Rician or Nakagami distribution and shadowing is modelled by Log Normal distribution [10], [11]. Hence robust tests which work well with heavy tailed noise and signals are required. In summary, it is desirable to have distributed, nonparametric, robust, sequential algorithms for spectrum sensing in a CR system which mitigate the effects of heavy tailed distributions also.

Spectrum sensing has been extensively studied in recent years. [1], [12] and the references therein give an overview of early work in spectrum sensing. See [2], [3], [13] for more recent contributions. Sequential procedures are found to be better suited for quick detection of spectral holes [14]. Distributed spectrum sensing has been a recent development in this direction [2], [13], [15], [16], [17], [18], [19] and the references therein). Various studies have suggested parametric ([20], [21]) as well as nonparametric ([22], [23]) solutions to this problem. None of these works studies the effect of EMI or outliers on the detection algorithm. Some of the issues in distributed detection are that the reporting channel (or decisions from the local nodes to the FC) should not require much bandwidth and the energy consumed and the delay in reporting the decisions should also be small [2]. Many of the works [2], [3], [21], [24] do not consider MAC noise or multipath fading in the reporting channel. However see [25] and the references therein for studies which consider shadowing and fading in reporting channels. Design of algorithms at the local nodes as well as the fusion centre are motivated by the various above considerations.

The contribution of this paper is in designing new sequential, nonparametric energy detection and mean change detection algorithms which perform well in the presence of slow-fast fading, heavy-tailed EMI and outliers. We are not aware of any other robust nonparametric scheme to mitigate the effects of EMI and outliers. These algorithms are then used in a distributed environment to provide an efficient spectrum sensing algorithm. Theoretic analysis of the distributed algorithm is also provided.

The paper is organized as follows. Section II provides the system model and the distributed set up. Section III presents several available (nonparametric) algorithms and their comparison via simulations. It also selects appropriate algorithms for the local nodes and FC for our distributed algorithm.

*This work was partially supported by a grant from ANRC.

978-1-4799-2361-8/14/$31.00 © 2014 IEEE
Section IV provides theoretical performance analysis of selected algorithms. Section V theoretically analyses the distributed algorithm. It also shows the effect of heavy tails on the system performance. Section VI provides the performance of the distributed algorithm for specific examples via simulations. Section VII concludes the paper.

II. SYSTEM MODEL AND DISTRIBUTED ALGORITHM

We consider a CR system where L CR (local) nodes are scanning the environment to detect if a primary user is transmitting or not. Based on their observations, the nodes make local decisions and transmit to the FC. The FC makes the final decision from the local decisions it receives from the secondary nodes. This is the most common distributed spectrum sensing architecture ([2], [3]).

At time k, node l senses $\tilde{X}_{kl}$ (at baseband level) where

$$X_{kl} = H_{kl}S_k + N_{kl}$$

if a primary is transmitting (Hypothesis $H_1$). Here, at time k, $H_{kl}$ is the channel gain from the primary to the local node l, $S_k$ is the symbol transmitted by the primary and $N_{kl}$ is the $l$th receiver noise with possibly some EMI. If the primary is not transmitting at time k (Hypothesis $H_0$) then

$$\tilde{X}_{kl} = N_{kl}.$$  

We assume that $\{S_k, k \geq 1\}$ and $\{N_{kl}, k \geq 1\}$ are independent identically distributed (i.i.d.) and independent of each other. In the following this assumption will be slightly generalized. Also, $\{N_{kl}\}$ are assumed independent sequences for different nodes l.

For $\{H_{kl}, k \geq 1\}$, we either assume that $H_{kl} \equiv H_1$, a random variable, possibly unknown (this is a commonly made assumption [10], [25]), representing slow fading, or an i.i.d. sequence, representing fast fading. $H_{kl}$ represents multipath fading as well as shadowing. For shadowing, log normal distribution is considered a good approximation [10], while for multipath fading, Rayleigh, Rician and Nakagami distributions are considered suitable [11]. Thus $H_{kl}$ could possibly have a heavy-tailed component (due to log normal distribution) and a light-tailed component (due to the fast fading component) [27]. Often the combined effect of these is approximated by a K-distribution [28] which has a heavy tail.

If sensing is done at times of primary symbol transmission then assuming $\{S_k\}$ to be i.i.d. is realistic which will often take values in a finite alphabet depending on the modulation scheme used by the primary. The secondary may not know the coding and modulation used by the primary. Also, different primary users may be using the same channel and a primary can change its modulation and coding with time. Thus, we will not assume that the local nodes know the signalling of the primary. This is a common assumption in the CR literature.

As a result of unknown $H_{kl}, S_k$ statistics, it is usually recommended to use energy detection at the local nodes ([2], [3]). Thus, we consider the energy samples

$$X_{kl} = \sum_{i=(k-1)M+1}^{MK} (\tilde{X}_{ij})^2$$

at each local node l where M is a constant decided as part of the sensing algorithm. Taking square of $\tilde{X}_{kl}$ in (1) provides the usual energy detector and is shown to be optimal for Gaussian noise in the absence of $S_k$ statistics. However, it has been shown [6] that for non-Gaussian noise, instead of 2, some other power $p$ of $|\tilde{X}_{kl}|$ may perform better. In the following we will keep $p = 2$ but allow the possibility of other powers when EMI is significant (see below).

In the following we will only assume $\{X_{kl}, k \geq 1\}$ to be i.i.d. independent sequences under $H_0$ and $H_1$ allowing $\{\tilde{X}_{kl}, M+k+1 \leq l \leq M(k+1)\}$ to have arbitrary dependence. This provides flexibility in modelling fading and sensing versus signalling duration.

The receiver noise is usually distributed as Gaussian, mean 0 and variance (say) $\sigma^2$ (denoted as $N(0, \sigma^2)$). However, in wireless channels there can often be a significant component of EMI [7]. EMI is modelled by Gaussian mixtures (which are light-tailed) and symmetric alpha-stable distributions (which are heavy-tailed for $\alpha < 2$) ([8]). Thus $N_{kl}$ will often not be Gaussian and can possibly be heavy-tailed. Of course, as a result of squaring $\tilde{X}_{kl}$, the noise distribution will no longer be symmetric.

Now we consider the hypothesis testing problem one encounters for energy detection with samples (1). We will denote by $P_e[X]$ and $Var[X]$, the distribution, the mean and the variance of $X$ under the hypothesis $H_i$, $i = 0, 1$. For simplicity, we take $\{\tilde{X}_{kl}, k \geq 1\}$ i.i.d. in this paragraph. If $N_{kl}$ has a general distribution with mean 0 and variance $\sigma^2$, under $H_0$, $\mathbb{E}_0[X_{kl}] = M \sigma^2$ and $\text{Var}_0[X_{kl}] = M(\mathbb{E}_0[(X_{kl})^2] - \sigma^2)$. Also, under $H_1$, $\mathbb{E}_1[X_{kl}] = M \sigma_1^2 + E_{al}$ and $\text{Var}_1[X_{kl}] = M(\mathbb{E}_1[(X_{kl})^2] - \sigma_1^2 + \mathbb{E}_1[(H_{kl}S_k)^2])$ where $E_{al} = \mathbb{E}_1[H_{kl}^2S_k^2]$, the received energy at node l.

If $\sigma_1^2 > E_{al}$ and $\sigma_1^2, E_{al}$ are known but the distributions of $N_{kl}, S_k$ are not known, we can consider it as a nonparametric mean detection problem with $H_0 : \mu = \mu_0 = M \sigma_2^2$ vs $H_1 : \mu = \mu_1 = M \sigma_1^2 + E_{al}$. It is a simple hypothesis testing problem with equal known variance under both hypotheses. If $E_{al}$ is not known but we know that $E_{al}$ is lower bounded by $E_{al}$ then the testing problem is $H_0 : \mu = \mu_0 = M \sigma_2^2$ vs $H_1 : \mu = M \sigma_2^2 + E_{al} \geq M \sigma_1^2 + E_{al} = \mu_1$. Now $H_1$ is a composite hypothesis. If $\sigma_2^2$ is also not known but we know that $\sigma_2^2 < \sigma_1^2 < \sigma_{al}^2$ then the problem is $H_0 : \mu = M \sigma_1^2$ vs $H_1 : \mu = M \sigma_2^2 + E_{al} \geq M \sigma_1^2 + E_{al} = \mu_1$. Now the variance under the two hypotheses are the same but unknown. The most general situation arises when the low SNR assumption is also violated and now the unknown variances under the two composite hypotheses are not the same.

As a consequence of the above comments, for a local node to make a decision, nonparametric statistical techniques which do not require complete knowledge of the distributions of observations $X_{kl}$ under $H_0$ and $H_1$ are suitable for energy detection. To make quick decisions, local nodes will use sequential detection. Thus node l will make its decision at a random time based on its local observations $\{X_{kl}, k \geq 1\}$. In the next section we compare several nonparametric sequential algorithms for energy detection and pick the best.

If node l decides $H_1$ at time k, it will transmit $+b_l$ to the FC. If it decides $H_0$, it transmits $-b_0$. If the node has not made a decision at a time, it transmits nothing. Thus, at time k, FC receives $Y_k = \sum_{l=1}^{L} G_{kl} Y_{kl} + Z_k$ where $Y_{kl}$ is the transmission from node l, $G_{kl}$ is the corresponding channel gain and $Z_k$ is the superposition of the receiver noise (which will often have a distribution $N(0, \sigma^2)$) and EMI. Thus, $Z_k$ will be a summation of Gaussian noise and Gaussian mixtures and/or alpha-stable EMI. Also, there is fading and shadowing on the wireless channels from the local nodes to the FC. Thus, we need at the FC a nonparametric sequential algorithm but unlike at the local nodes, the signalling $(+b_l$ or $-b_0)$ is known to the FC. Furthermore, unlike at the local nodes, we can use
partially coherent detection (we may be able to estimate the phase; in particular, the sign of $G_{kl}$ although not necessarily the magnitude of the channel gains). Then the local node multiplies its transmission $Y_{kl} (+b_1$ or $-b_0)$ by the sign of $G_{kl}$ and transmits. Thus, $Y_k = \sum_{l=1}^L |G_{kl}|Y_{kl} + Z_k$. Therefore we do not need an energy detector (actually in our setup we may not be able to use the energy detector at the FC) but in fact a nonparametric detector which performs well for mean change detection with symmetric noise will be a suitable choice.

As discussed above, at the local nodes as well as at the FC, due to possibly significant EMI, the noise may be heavy-tailed. Such a scenario in CR has been considered in [7]. But the impact of heavy-tailed noise has not been specifically studied. In [17], this was considered in the context of change detection and it was shown that heavy tails can degrade the performance significantly. In this paper, for the distributed hypothesis testing algorithm also, we show that heavy-tailed distributions can significantly impact the performance. Then we will modify the algorithms so that their impact along with that of the outliers which are also present, can be mitigated.

Often the reporting (MAC) channel from the local nodes to the FC is considered noiseless (2, 21, 3, 24). However, as mentioned above, like any other wireless channel, it does experience EMI, outliers and receiver noise. One implication of this is that the decisions transmitted by local nodes may not reach the FC without error making the use of standard Fusion centre rules - AND, OR, majority etc. [3] less accurate and/or difficult to implement.

Now we describe our basic distributed algorithm which has been shown to be asymptotically optimal and performs well at practical parameter values ([18], [23]). It also makes an efficient use of the reporting MAC. An optimal algorithm in this setting is not known [16]. We will complete this algorithm by choosing appropriate detection algorithms for the local nodes and the FC in the next sections. We will also study the performance of the overall algorithm so developed especially under the influence of EMI and outliers.

Distributed Algorithm

- Each local node $l$ receives observation $X_{kl}$ at time $k$.
- Each node $l$ uses a sequential algorithm to compute $T_{kl} = f(X_{kl}, X_{(k-1)l}, ..., X_{1l})$ and makes a decision at time $N_l$ where
  \[ N_l = \inf \{ n : T_{kl} \notin (-\gamma_0, \gamma_1) \}, \]
  $\gamma_0$, $\gamma_1$ are appropriately chosen constants and the decision is $H_0$ if $T_{Nkl} \leq -\gamma_0$ and $H_1$ if $T_{Nkl} \geq \gamma_1$. It transmits $Y_{kl}$ to the FC at time $k$ where
  \[ Y_{kl} = b_1 \mathbb{1}\{T_{kl} \geq \gamma_1\} - b_0 \mathbb{1}\{T_{kl} \leq -\gamma_0\}. \]

Node $l$ will keep transmitting till the FC makes a decision.
- At time $k$, FC receives
  \[ Y_k = \sum_{l=1}^L Y_{kl} + Z_k \]
  and computes $W_k$ based on an algorithm to be decided. At time
  \[ N = \inf \{ n : W_n \notin (-\beta_0, \beta_1) \}, \]
  it decides $H_1$ if $W_N \geq \beta_1$ and $H_0$ if $W_N \leq -\beta_0$ where $\beta_0$, $\beta_1$ are appropriately specified. After $N$, all nodes stop transmitting.

The energy detection algorithm to be used by the local nodes and the mean change detection to be used at the FC will be chosen in the next section.

One of the advantages of our distributed algorithm is that the local node $l$ which has a good channel gain $H_{kl}$ from the primary will make a decision faster and will influence the FC decision more. Also, since each local node keeps transmitting its decision till the FC decides, if a local node has made a wrong decision, most likely it will soon change it and hence wrong local decisions will have minimal effect on the FC decision, especially when $P_{FA}$ and $P_{MD}$ are small.

### III. Algorithms for Single Node

In this section we consider sequential nonparametric single node algorithms with their statistics denoted by $T_n$, which can be used by the local nodes and the FC for energy detection and mean change detection respectively. Optimal tests for single nodes also do not exist. We will not use the node index $l$ in this section.

#### A. Rank test

Rank test (Wilcoxon rank test) is a location test [5] for location $\mu$ of a distribution $F(x - \mu)$ which is symmetric around $\mu$. For testing $\mu \leq \mu_0$ vs $\mu \geq \mu_1$, $\mu_1 > \mu_0$, its statistics is defined as follows.

i. Let $Y_i = X_i - \frac{\mu_0 + \mu_1}{2}$, where $X_i$'s are the observations.
ii. Calculate $R_i$, the rank of $Y_i$ in $Y_1, ..., Y_n$ when these are arranged in ascending order of their absolute values.
iii. Test statistic $T_n = \sum_{i=1}^n sgn(Y_i) \frac{R_i}{n+1}$ where $sgn(x) = \frac{x}{|x|}$ for $x \neq 0$ and $0$ for $x = 0$.

We will use this statistic in our sequential set up. This statistic is distribution free for symmetric distributions [5].

#### B. Sequential t test

We use the usual $t$ test [29] extended to make it a two sided test. The test statistic is given by,

\[ T_n = n \frac{\bar{X}_n - \frac{\mu_0 + \mu_1}{2}}{s_n} \]

where $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ is the sample mean, and $s_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_k - \bar{X}_n)^2}$ is the sample variance.

#### C. Random walk

Its test statistic is obtained by modifying the above $t$ test statistic:

\[ T_n = \sum_{i=1}^n (X_i - \frac{\mu_0 + \mu_1}{2}). \]

The statistic is a simple random walk and we refer to this algorithm as random walk.

The above three tests are primarily designed for detection of mean change $H_0: \mu \leq \mu_0$ vs $H_1: \mu \geq \mu_1$, but can also be used for testing some other functional of the distributions. Unlike sequential $t$ test and rank test, random walk test is iterative. Thus it is simpler to compute the statistic and does not require storing the whole data.
D. Mitigating effects of outliers, heavy tailed and fading

The sample mean and the sample variance used in the \( t \) test and random walk are not robust to outliers. This gets reflected in the performance of these tests (compare Figures 3 and 4 below; see also Figure [1]). From Figures 3 and 4 we also see that the rank test is quite robust to outliers although may not perform the best. This motivates the use of robust versions of the random walk and \( t \) tests [2]. Robust tests are obtained by replacing the sample mean (and sample variance) in these tests by their robust versions.

- \( M - t \) test is obtained by applying a cut-off function \( \psi \) to obtain a robust sample mean (modified sample variance is in the denominator of \( T_n \) below) and obtain the statistics of \( t \) test as

\[
T_n = \frac{\sum_{i=1}^{n} \psi (X_i - \mu_0 + \mu_1)}{\left( \sum_{i=1}^{n} \psi^2 (X_i - \bar{X}_n) \right)^{\frac{1}{2}}},
\]

where \( \psi : \mathcal{R} \rightarrow \mathcal{R} \) is a non decreasing, continuous, odd and bounded function. For \( N(0, 1) \), a recommended \( \psi \) [9] is

\[
\psi_0(z) = \begin{cases} 
K, & \text{if } z > K, \\
0, & \text{if } |z| \leq K, \\
-K, & \text{if } z < -K.
\end{cases}
\]

for a given \( K < \infty \).

- Applying the \( \psi \) function on the random walk, we get a robust version called \( M \)-random walk via the statistic

\[
T_n = \frac{\sum_{i=1}^{n} \psi (X_i - \mu_0 + \mu_1)}{2},
\]

This statistic is iterative, unlike \( \alpha \)-trimmed random walk/t test or \( M - t \) test.

It is known that the \( t \) test is not efficient for heavy-tailed distributions [29]. One expects this behaviour for the random walk test also (see Figure [1] below). On the other hand, the rank test is quite efficient for heavy tailed distributions also.

We will also see that the Huber’s \( M \) function \( \psi \) not only robustifies \( t \) and random walk tests but also makes them more efficient w.r.t. heavy-tailed distributions. We will confirm these findings from simulations and the theory in Section [15].

In very heavy-tailed case (SoS with \( \alpha < 1 \) or for energy detection with \( \alpha < 2 \), the mean of the sample \( X_k \) is infinity. Thus, random walk and \( t \) test will not work. The rank test can possibly still work. Even the above robust versions of random walk and \( t \) test (3) and (4) will not work directly because \( \mu_0 \) and \( \mu_1 \) will be infinity. Thus, we consider samples

\[
\bar{X}_i = \psi_1 (X_i)
\]

where \( \psi_1 \) is from the class of functions mentioned below equation (3), and use \( M \)-random walk test on it with \( \mu_0 \) and \( \mu_1 \) corresponding to the means of \( \bar{X}_i \). We call this \( M^2 \)-random walk test. We will see below via simulations that \( M^2 \)-random walk test works for SoS with \( \alpha < 2 \) while \( M \)-random walk, random walk, \( t \), \( M - t \) and \( M - t \) based on samples (7) do not work at all.

Choice of \( \psi \) in (3), (6) and \( \psi_1 \) in (7) affects the performance of the algorithm (see [9] for different \( \psi \) in parametric set up). In our nonparametric setup we will simply use \( \psi_0 \) defined in (3) with different \( K \) values. Our aim of using \( \psi \) for heavy-tailed case is to create light-tailed samples (7). In our simulations below for energy samples, we will take \( K \) large for \( \psi_1 (\approx 200) \) but small (\( \leq 5 \)) for \( \psi \) in (3) and (6).

It has been known that slow fading can significantly degrade the performance of a detection algorithm (25). We will see that this happens for the above algorithms also. This is because in slow fading, \( X_k = H S_k + N_k \) and for usual fading distributions e.g., Rayleigh, \( H \) can be small with a large probability. In this case, applying the \( \psi \) function does not help. Then if we do not make a decision when \( |H| \leq \delta \) for a small \( \delta \), it can significantly improve the performance if we take \( E_H [\mathbb{E}_n [N(H)]] \) as the performance measure for given \( P_{T_A} \) and \( P_{S.D} \). The constant \( \delta \) needs to be chosen carefully depending on the desired probabilities of error. We will study the effect of this operation via simulation and theory in the following.

When both EMI and slow fading are present, then we should combine the above two operations: not make a decision if \( |H| \leq \delta \) and when we do make, we use (3), (6) and (7). We will call the corresponding algorithms, \( \delta \)-random walk, \( M - \delta \)-random walk and \( M^2 - \delta \)-random walk.

E. Simulation Results

We compare the above algorithms for mean change detection when the channels may experience slow/fast fading with shadowing and SoS EMI and outliers. This scenario can be useful for energy detection at low SNR. We have taken \( \alpha = 1.8 \) for the SoS distribution (8) and fading is Rayleigh distributed with parameter \( P \) where \( P \sim \log N(0, 0.36) \) represents shadowing (27). The receiver noise \( Z \sim N(0, \sigma^2) \) and \( SNR = 10 \log_{10} \frac{E[H^2][|g - m_0|^2]}{\sigma^2 E[\mathbb{E}_n [N]|g]} \). The X-axis shows \( P_{T_A} + P_{S.D} \) and the Y-axis shows \( E[\mathbb{E}_n [N]|g] + E_c[|X|] \). For slow fading we keep the channel gains constant till the decisions are made. The simulations were run 10,000 times and averaged to obtain the probabilities of error and the mean time to sense.

Figures [1](5) show the simulations for various algorithms with various combinations of fast/slow fading, SoS EMI and outliers. We draw the following conclusions.

- From Figures [2] [5] we see that the random walk test always performs better than the \( t \) test and the rank test.
- From Figures [2] [3] [4] comparing the top part of each figure (for fast fading) with the bottom part (for slow fading), for each algorithm, slow fading performs much worse. The effect of heavy-tailed EMI is somewhat like that of fast fading.
- From Figure [1] we see that for random walk, slow fading has the most devastating effect on performance. This can be seen for other algorithms also from other figures. Next major damage is done by outliers. We see that heavy-tailed EMI also degrades the performance significantly.
- From Figures [3] [5] we observe that when there is only Gaussian noise and fast/slow fading \( M \)-random walk does not improve the performance over random walk. This is expected because the operation of \( \psi \) is used only to improve the performance with respect to outliers and heavy-tailed EMI. We will see in the next section that degradation via (slow) fading is mainly due to the channel gain \( H \) being low very often. Also see comments below.
- That \( M \)-random walk and \( M^2 \)-random walk are very effective in mitigating the effects of heavy-tailed EMI and outliers can be seen from Figures [5] [4]. From these we can conclude that outliers can cause major damage (for random walk, \( t \) test) but are effectively handled by \( M \)-random walk. The rank test is not affected so much. In case of energy detection with SoS EMI and fast fading, the only algorithm (among the algorithms considered) that works at all is \( M^2 \)-random walk. Other algorithms do not provide probability of error \( \leq 0.3 \).
Performance of $M^2$-random walk test with EMI is presented in Figure 8 along with that of the distributed algorithm. From Figure 8 we also see that unlike in Figure 8 the outliers are helping the performance in the energy detection case. This is because we consider outliers only when there is signal ($H_1$) and not under $H_0$ unlike in Figure 8 where $H_1$ and $H_0$ both have signal.

As mentioned above, slow fading causes maximum degradation. This is because, for Rayleigh fading, the channel gain $H$ is low with a large probability. In that case, not making a decision when $H$ is very small is the sensible thing to do. Thus our algorithm $\delta$-random walk actually improves the performance significantly in this case (see Figure 9).

$M^2$-random walk can be made to work close to $M$-random walk if we take $K$ in $\phi_1$ large. Therefore we could as well use $M^2$-random walk for mean change detection also. Asymptotic analysis of the random walk test is provided in [30]. In the next section we briefly present that and also include the effects of heavy tailed noise and fading which was not discussed in [30]. This will explain why $M$-random walk and $M^2$-random walk perform better under heavy-tailed EMI and outliers and using truncation on $H$ improves performance in the presence of slow fading.

![Fig. 1: Effect of different factors on the performance of random walk.](image)

**IV. ANALYSIS FOR RANDOM WALK**

Based on Section III we have decided to use the $M$-random walk at the FC and the $M^2$-random walk test at the local nodes. Asymptotic analysis of the random walk test is provided in [30]. In this section we briefly present that and also include the effects of heavy tailed noise which was not discussed in [30]. This will explain why $M$-random walk and $M^2$-random walk perform better under heavy-tailed EMI and outliers. Then we will include the effect of fast and slow fading and show that $\delta$-random walk can improve the performance.

First we consider the scenario of mean change detection. Here, the noise can have heavy tail due to Gaussian and symmetric $\alpha$-stable (or other heavy tailed) distribution. The overall noise distribution is symmetric. Furthermore, the fading distribution can also be heavy tailed.

The family of $\alpha$-stable laws is denoted by $S_{\alpha}(\sigma, \beta, \mu)$ with $0 < \alpha \leq 2$ its index, $\sigma > 0$ a scale parameter, $-1 \leq \beta \leq 1$ its skewness and $-\infty < \mu < +\infty$ its location. When $\alpha = 2$ then it becomes $N(\mu, 2\sigma^2)$. All $\alpha$-stable laws have continuous, positive, uni-modal probability density function. A random variable $X$ with $\alpha$-distribution, $0 < \alpha < 2$ satisfies $P[X > x] \sim x^{-\alpha} \sim P[X < -x]$ and $E[|X|^p] < \infty$ for $0 < p < \alpha$ and $E[|X|^p] \sim \infty$ for $p \geq \alpha$.

We also allow for the possibility of Gaussian mixture for EMI, which is light-tailed. Also, for $M$-random walk and $M - t$

![Fig. 2: Mean change detection at FC in the presence of Gaussian noise. Top: Log $N$ shadowing - Rayleigh fast fading. Bottom: Log $N$ shadowing - Rayleigh slow fading.](image)

test, due to bounded Huber $\psi$ function, all distributions become light-tailed.

We will use the following terminology. For CDF $F$, $\bar{F}(x) = 1 - F(x)$, $F^{-2}$ is convolution of $F$ with itself and $F^{-2}(x) = 1 - F(2x)$.

**Definition [31]:** $F$ is light-tailed if $\int_{-\infty}^{\infty} e^{\alpha x} dF(x) < \infty$ for all $\alpha$ with $0 \leq |\alpha| < \alpha_1 < \infty$; otherwise it is heavy-tailed. $F$ is long-tailed ($F \in L$) if $\lim_{x \to \infty} \frac{\bar{F}(x+y)}{\bar{F}(x)} = 1$ for all finite $y$. $F$ is sub-exponential ($F \in S$) if $\lim_{x \to \infty} \frac{\bar{F}(x+y)}{\bar{F}(x)} = 2$ where $B$ is the distribution of $\max\{0, X\}$ while $X$ has the distribution of $F$. $F$ is regularly varying of index $-\alpha$, $\alpha > 0$, (denoted by $F \in R(-\alpha)$), if $\bar{F}(x) = l(x)e^{-\alpha x}$, where $l$ is a slowly varying function, i.e., for all $\lambda > 0$, $\frac{\bar{F}(\lambda x)}{\bar{F}(x)} \to 1$ as $x \to \infty$. $F \in S^*$ if $\lim_{x \to \infty} \frac{\bar{F}(x+y)}{\bar{F}(x)} = 2 \int_{-\infty}^{\infty} \bar{F}(y)dy$.

A long-tailed distribution is heavy-tailed. Also, $S^* \subset S \subset L$ and $R(-\alpha) \subset S$. If $F \in R(-\alpha)$ and it also has a finite mean, then it is in $S^*$. Gaussian, exponential and Laplace distributions are light-tailed while Pareto, Log Normal and Weibull distributions are sub-exponential. For $\alpha < 2, S_{\alpha}(\sigma, \beta, \mu) \sim R(-\alpha)$.

When $S_k$ takes values in a finite set and $H_k$ is light-tailed then $H_{k}S_k$ is light-tailed; if $H_k$ is heavy-tailed then $H_{k}S_k$ is heavy-tailed, if $H_k \in R(-\alpha)$ then $H_{k}S_k \in R(-\alpha)$. If independent random variables $X$ and $Y$ are light-tailed then $X + Y$ is light-tailed. If any of $X$ and $Y$ is heavy-tailed so is $X + Y$. If $F \in S$, $G(x) = O(F(x))$, then $F \ast G \in S$. If $X, Y$ are long-tailed then $X + Y$ is long-tailed. If $X \in R(-\alpha)$, $Y \in R(-\alpha)$ then $(X + Y) \in R(-\min\{\alpha_1, \alpha_2\})$. If $X \in L$, then $X^2 \in L$. If $X \in R(-\alpha)$, then $X^2 \in R(-\alpha/2)$.

The above results provide us the tail behaviour of $N_k + H_{k}S_k$, $N_k + H_{k}b_1$ and $N_k - H_{k}b_0$ in terms of tail behaviour of $N_k$ and $H_k$ where $b_0$ and $b_1$ are positive constants.

Consider the random walk statistics [3] or the robustified
random walk $\psi$ with Huber function $\psi$.

We write it as $T_n = \sum_{k=1}^{\infty} Y_k$ where $Y_k = (X_k - \frac{\ln(2\pi i)}{\sqrt{2\pi} i})$ or $Y_k = \psi(X_k - \frac{\ln(2\pi i)}{\sqrt{2\pi} i})$. We choose $\psi$ such that $\theta_1 \triangleq \mathbb{E}_0[Y_1] < 0$ and $\theta_1 \triangleq \mathbb{E}_1[Y_1] > 0$. Implications for $M^2$-random walk directly follow.

The sequential test for the random walk statistics stops at $N = \inf \{n: Y_n \notin (-t_0, t_1)\}$ where $t_0, t_1 > 0$. We will pick $t_0, t_1$ independently of the distribution of $Y_k$. Since the statistic $Y_k$ is independent of distribution of $Y_k$, this will make the test completely nonparametric. We will discuss picking $t_0$ and $t_1$ later on. Once $t_0, t_1$ are fixed, the actual performance of the test does depend on the distribution of $Y_1$ and we study that now. Define, for $t > 0$,

$$N_1(t) = \inf \{n: T_n > t\}, N_0(-t) = \inf \{n: T_n < -t\}.$$

We consider $\mathbb{E}_0[N]$. The results will similarly hold for $\mathbb{E}_1[N]$. Let $S_n = \sum_{k=1}^{n} Y_k$ and $M = \sup_{n \geq 0} S_n$.

Under $H_0$, $\mathbb{E}_0[Y_k] = \theta_0 < 0$. Thus $N_0(-t) < \infty$ a.s. for all $t > 0$ and $N_1(t) = +\infty$ a.s. for all large $t > 0$ (since $M < \infty$ a.s.). Consider $N(t) = \min\{N_0(t), N_1(t)\}$. Thus,

$$\lim_{t \to \infty} \mathbb{P}_0[N(-t) = N_0(-t)] = 1, \quad \text{and} \quad \lim_{t \to \infty} \frac{N_1(t)}{t} = \lim_{t \to \infty} \frac{N_0(-t)}{t}, \quad \text{a.s.}$$

Since we want to design algorithms with small probabilities of error, we will work with $t$ where $P[N(-t) = N_0(-t)]$ is large. Thus, we consider $N_0(-t_0)$. From random walk theory \cite{32}, the following results hold. We have $\lim_{t \to \infty} \frac{N_0(-t_0)}{t} = \frac{t_0}{\theta_0}$ a.s. and in $L_1$ even when $\theta_0 = -\infty$ (then the limit is 0). For $r \geq 1$, if $\mathbb{E}[|Y_1|^r] < \infty$ then $\mathbb{E}[N(t_0)^r] < \infty$ and if $Y_1$ has finite moment generating function in a neighbourhood of 0 then $N(t)$ also has. Here and in the following $Y_1^- = \min\{0, Y_1\}$ and $Y_1^+ = \max\{0, Y_1\}$. Also $F$ denotes the distribution of $Y_1$.

Fig. 3: Mean change detection at FC in the presence of Gaussian and symmetric $\alpha$-stable noise. Top: Log $N$ shadowing - Rayleigh fast fading. Bottom: Log $N$ shadowing - Rayleigh slow fading.

Fig. 4: Mean change detection at FC in the presence of Gaussian and symmetric $\alpha$-stable noise and 5% $N(0, 20)$ outliers. Top: Log $N$ shadowing - Rayleigh fast fading. Bottom: Log $N$ shadowing - Rayleigh slow fading.

Fig. 5: Energy detection in the presence of Gaussian noise. Top: Without fading. Bottom: Under block Log $N$ shadowing - Rayleigh fast fading.

For $1 < r < 2$, if $\mathbb{E}[|Y_1|^r] < \infty$ then $\mathbb{E}_0[N(t_0)] = \frac{t_0}{\theta_0} + o(t^{2-r})$. If $\mathbb{E}[|Y_1|^2] < \infty$,

$$\frac{t_0}{\theta_0} \leq \mathbb{E}[N_0(\theta_0)] \leq \frac{t_0}{\theta_0} + \frac{\mathbb{E}[|Y_1|^2]}{2\theta_0} + o(1). \quad (8)$$

Similar results hold for $\mathbb{E}_1[N_1(t_1)]$ with condition on $\mathbb{E}[|Y_1|^r]$.
Similarly we get

\[ \mu_0 = t \theta_0 \]  

moments of \( N(0,5) \) noise and shadowing-slow fading, then we define

\[ \Gamma(x) = \int_0^x \log(1 + y) dy \]  

Now we consider the case where \( 1 - F_0(t_1) \sim t_1^{-\alpha_1} \) and \( 1 - F_1(t_1) \sim t_1^{-\alpha_1} \). Then \( \alpha = P_{FA} \sim (1 - F_0(t_1))E[N(t_0)] \sim t_1^{-\alpha_1} \) \( \theta_0 \) \( \beta = P_{MD} \sim t_1^{-\alpha_1} \) \( \theta_1 \) and hence

\[ E_0[N] \sim \theta_0 \left( \frac{1}{t_1^{-1}} \right)^{-\alpha} = \theta_0 \left( \frac{1}{t_1^{-1}} \right)^{-\alpha_1} = \theta_0 \left( \frac{1}{t_1^{-1}} \right)^{-\alpha_1} \]  

(13)

\[ E_1[N] \sim \theta_1 \left( \frac{1}{t_1^{-1}} \right)^{-\alpha} = \theta_1 \left( \frac{1}{t_1^{-1}} \right)^{-\alpha_1} = \theta_1 \left( \frac{1}{t_1^{-1}} \right)^{-\alpha_1} \]  

(14)

This shows that the performance of the random walk algorithm depends quite strongly on the tail behaviour of \( F \) and with heavy tails the performance can really deteriorate.

Now we briefly comment on the performance of the (robust) random walk for mean change detection under: \( H_0, Y_k = N_k \sim b_0 H_k \) and under \( H_1, Y_k = N_k + b_1 H_k \).

Initially assume that there is no fading, i.e., \( H_k \equiv 1 \). It is then a mean change detection with \( \mu_0 = -b_0 \) and \( \mu_1 = b_1 \). Now the above analysis directly provides the effect of light and heavy-tailed \( N_k \). Also, we see that by applying Huber function \( \psi \) we can substantially gain in case of heavy-tailed \( N_k \). For light-tailed case if we pick \( K \) small, then it can make \( \mu_0 \) and \( \mu_1 \) smaller and hence one may see worse performance.

Next we consider the case of slow fading: \( H_k \equiv H \). Now, it is realistic to assume that \( H \) has been estimated and the receiver knows it (coherent detection case). Then we can consider observations \( Y_k = \frac{Y_k}{H_k} + b_0 + (\frac{Y_k}{H_k} - b_0) \). Since \( N_k \) is zero mean, symmetric, independent of \( H, N_k/H \) stays zero mean. Also given \( H = b_0, b_1 \) will be heavy/light-tailed if \( N_k \) is. Thus, it becomes the case considered in the previous paragraph.

Denoting by \( P_{FA}(t_0, t_1, h), P_{Msd}(t_0, t_1, h), E_0[N_0(t_0, h), E_1[N_1(t_1, h), \text{the corresponding quantities, } E_0[N_0(t_0, h)], E_1[N_1(t_1, h),] \text{ and } E_2[N_2(t_2, h)] \text{,and so on.} \)

For light-tailed case, \( E_0[N_0(t_0)] \approx \left\{ \left( \frac{t_0}{h} \right)^{-\alpha} \right\} \) for \( 0 < \alpha < 1 \) and Rayleigh fading \( \frac{1}{t_1^{-1}} \). This is reflected in a significant performance degradation seen in the simulation results in Section III-E.

If \( P[N_k > t] \sim t^{-\alpha} \) then \( P[|N_k| > t_0] = t_0^{-\alpha} \sim (t_1 + b_0)H t_1^{-\alpha_1} \) and by (14) we get asymptotics for \( E_0[N_0(t)] \) and \( E_1[N_1(t), H] \). We can further take expectation over \( H \) to get the dependence on distribution of \( H \).

Above, we made the thresholds \( t_0 \) and \( t_1 \) dependent on \( H \) and ensured that for each \( h, P_{FA} \leq \alpha \) and \( P_{Msd} \leq \beta \). But this can often imply that \( E_0[N_1(t_1)] \) and/or \( E_1[N_1(t_1, H)] \) is infinite. A weaker requirement is to choose \( \theta_0 \) and \( \theta_1 \) independently of \( H \). In that case we can find positive constants \( \delta, \delta', \beta \) such that \( \delta < \min\{\alpha, \beta \} \) and \( P[H < \delta] \leq \delta \) with \( E_{FA}(H) \) \( H \geq \delta \leq \alpha \) \( E_{Msd}(H) \) \( H \geq \delta \leq \beta \) and

\[ E_{FA}(H) = P[H < \delta] + E_{FA}(H) (H \geq \delta) \leq \delta \leq \delta' \]  

(15)

and \( \delta + \beta'(1 - \delta) = \beta \). Now we do not make a decision when \( |H| > \delta \). For this case we can ensure that \( E_{H}[N_1(t_1)] < \infty \) for \( i = 0, 1 \). At least for Gaussian \( N_k \) and Rayleigh fading example above, \( \frac{1}{t_1^{-1}} |H| \geq \delta \) \( \leq \delta \).

If we assume that we only know the sign of \( H \) and not its magnitude (partial coherence knowing the phase only) then we define \( Y_k = \text{sgn}(H)Y_k = \text{sgn}(H)N_k + |H|Y_0 \) under \( H_1 \) (or \( \text{sgn}(H)N_k - |H|Y_0 \) \( H_0 \)). Now since \( N_k \) is zero mean, symmetric, \( \text{sgn}(H)N_k \) has the same distribution as \( N_k \). Also \( E_0[N_0(-t_0, h)] = \frac{1}{t_0^{-1}}, \ E_1[N_1(t_1, h)] = \frac{1}{t_1^{-1}} \) and if \( P[N_k > t] \sim t^{-\alpha} \) then we can get from (14), \( E_1[N_1(H)] \) and \( E_0[N(H)] \).
From the above two paragraphs, we can see the advantage of knowing magnitude $|H|$ at the receiver. Also, not knowing $|H|$ implies that we cannot decide when $|H| \geq \delta_1$ as needed in (13). Analysis of $P_{MD}$ follows in the same way.

If the phase of $H$ is also not known, then random walk algorithm is not the right choice for this problem because it will perform quite badly. Now we consider the fast fading case where $\{H_k\}$ is i.i.d. This is a less likely scenario but we briefly discuss it because it leads to some new results. As above, if we have a noncoherent case (no sign or magnitude of $H$ available) then we do not need any information about $H_k$. We do not need any information about $H$ and the above results can be presented. In this case particularly, since $|F_A|$ is light-tailed $H$ is expected that $\text{sgn}(H) = \text{sgn}(H_k)N_k + |H_k|b_1$ under $H_1$ and $\text{sgn}(H_k)N_k - |H_k|b_0$ under $H_0$, we obtain the following conclusions:

- If $N_k$ is light-tailed but $H_k$ is heavy-tailed, $P_t$ has a positive heavy tail and light negative tail and vice versa for $P_0$. Thus, system performance is not affected by the heavy-tailed $H_k$. One can see some beneficial effects because $E[N_i]$ will be somewhat shorter which is not captured by our analysis.
- If $N_k$ is heavy-tailed symmetric but $H_k$ is light-tailed then $P_0$ and $P_t$ both have heavy positive and negative tails. Thus, $P_{FA}$ and $P_{MD}$ both suffer.
- If $N_k$ and $H_k$ both are heavy-tailed then again $P_{FA}$ and $P_{MD}$ suffer.

Now we consider the system described in Section II. Under $H_0$, $X_k = N_k$ and under $H_1$, $X_k = H_kS_k + N_k$. As discussed, we use energy detection for this case by taking samples $X_k$ in $\mathbb{R}$. Then, from the results above, if $\{S_k\}$ is i.i.d. with values in a finite set and $\{H_k\}$ i.i.d. (fast fading) depending on the tail behaviour of $H_k$ and $N_k$, we know the tail behaviour of energy samples $X_k$. Also, under various SNR conditions, we know that the energy detection problem can be considered the mean change detection problem and the above results can be directly used. We do not need any information about $H_k$ itself; only the mean of $X_k$ under $H_1$ and $H_0$ must be required (at least for the low SNR case).

For slow fading case, $H_k \equiv h$, a constant in the sensing duration. Then, at low SNR, it is mean change detection with $\mu_0 = M\sigma^2$ and $\mu_1 = M\sigma^2 + 2E[S_k^2]$. Now, for given thresholds $t_0$ and $t_1$, $E[N_i(-t_0,h)] = \frac{\alpha_0}{\alpha_1}$ and $E[N_i(t_1,h)] = M\sigma^2 + E[S_k^2]$. Also, $P_{FA}(t_0,t_1,h)$ and $P_{MD}(t_0,t_1,h)$ can be approximated/bounded as above and the effect of heavy and light-tailed $N_k$ can be studied. Taking expectation over $H$ will provide the effects of tail of the distribution of $H$ as well.

If $H_k \equiv H$ (slow fading) and unknown, then let for $H = h$, $P_{FA}(h)$, $P_{MD}(h)$, $E[N_i(-t_0,h)]$, $E[N_i(t_1,h)]$ represent the corresponding probabilities of error and expected detection times. Then $E[N(-t_0)] \approx t_0/h$. If $N_k \in \mathbb{R}^*$, then $P_{FA}(t_1,h) \sim \left(1 - F_0(t_1 + h \delta_0)\right)\frac{\mu_1}{\mu_0}$ where $F_0$ is the cdf of $N_k$. Also, $E_p[P_{FA}(t_1,h)] \sim \int_0^{\infty} \left(1 - F_0(t_1 + h \delta_0)\right)\frac{dP_H(h)}{\mu_0}$ where $P_H$ is the distribution of $H$. Similarly one can study the case of $H$ being light-tailed. The analysis for $P_{MD}$ is along the same lines. In this case particularly, since $\psi$ is bounded, one expects that $T_{n}^a$ will provide much better performance. This study explains the results observed in Section III-E.

V. ANALYSIS OF DISTRIBUTED ALGORITHM

Based on the simulation results in Section III and the theory in Section IV we now consider the distributed algorithm where each local node uses $M^2$-random walk and FC uses $M$-random walk. Thus, the observations at the local nodes and the fusion node after operation with the $\psi$ function are light-tailed, in fact bounded. Therefore, assumptions of Theorem 2 and 3 below, will be satisfied.

Let $\tilde{X}_k = \psi_1(X_k) - \frac{\mu_0 + \mu_1}{2}$, where $E_1[X_k] \geq \mu_1$, $E_0[X_k] \leq \mu_0$ and, $\mu_1 > \mu_0$ for $l = 1, 2, ..., L$.

We choose $\theta_0$ and $\theta_1$ such that

- $H_0 : E[\tilde{X}_k] \leq \theta_0$,
- $H_1 : E[\tilde{X}_k] \geq \theta_1$, $\theta_1 > \theta_0$.

Then, $T_{id} = \sum_{k=1}^{n} \left(\tilde{X}_k - \frac{\theta_0 + \theta_1}{2}\right)$ and $W_n = \psi_0\left(\tilde{Y}_k - \frac{\theta_0 + \theta_1}{2}\right)$, where $\tilde{Y}_k$ and $\tilde{X}_k$ are selected properly such that $\tilde{Y}_k \leq \tilde{X}_k \leq \tilde{X}_k$.

We use the following notation:

$\Delta_i = \text{mean drift of } W_k\text{ when all local nodes decide } H_i$,

$D_{i,tot} = \sum_{t=1}^{L} E_1[\tilde{X}_k - \theta_0 + \theta_1, 2]$,

$N = \inf\{k : F_k \geq \beta_1 \text{ or } F_k \leq -\beta_0\}$,

$\zeta^*_t = \psi_0(L b_1 + Z_t)$,

$R_i = \log \inf E_1[\psi_0(-t\tilde{X}_k)]$.

We choose $\psi_0$ such that $\Delta_0 < 0$ and $\Delta_1 > 0$.

Theorem 1: For any finite thresholds $\gamma_{lt}$, $\beta_l$, $P_{t}[N < \infty] = 1$.

Proof: Please see the appendix.

For Theorems 2-4, we will use the following thresholds:

- $\beta_0 = -|\log c|$, $\beta_1 = |\log c|$, $\gamma_0 = -\gamma_1 |\log c|$, $\gamma_1 = -\rho_1 |\log c|$, where,

$\gamma = \frac{E_0[\tilde{X}_k - \theta_0 + \theta_1]}{D_{i,tot}}$,

$\rho_i = \frac{E_0[\tilde{X}_k - \theta_0 + \theta_1]}{D_{i,tot}}$.

Theorem 2: Let $E_0[|\tilde{X}_k|^\alpha] < \infty$ for $l = 1, 2, ..., L$ and $E_0[|\tilde{Z}_k|^\alpha] < \infty$ for some $\alpha > 1$. Then under $H_i$,

$\limsup_{c \to 0} \frac{N}{|\log c|} \leq \frac{1}{D_{i,tot}} + M_i$ a.s.

and in $L_1$ where $M_i = \frac{\alpha_0 - c_0}{\alpha_1}, c_0 = -[1 + \frac{E_0[\zeta^*_t]}{D_{i,tot}}]$, $c_1 = [1 + \frac{E_0[\zeta^*_t]}{D_{i,tot}}]$.

Proof: Please see the appendix.

The following result provides the rate of convergence of $P_{FA}$ and $P_{MD}$ to zero as the local and FC thresholds tend to $\infty$ (and $-\infty$) for the light-tailed $F$. Let $g$ be the moment generating function of $\zeta^*_t$ and $\Lambda(\alpha) = \sup_{\lambda}(\alpha \lambda - \log \gamma(\lambda))$, $\alpha^* = \sup \zeta^*_t$. Also, let

$s(\eta) = \left\{ \begin{array}{ll}
\frac{n}{\eta - \alpha} & \text{if } \eta \geq \Lambda(\alpha^*) \\
\frac{n}{\alpha(\alpha^*)} & \text{if } \eta \in (0, \Lambda(\alpha^*))
\end{array} \right.$
Theorem 3: Let $g(\lambda) < \infty$ in a neighbourhood of zero. Then,
(a) $\lim_{\lambda \to 0} P_{FA} = 0$ if for some $0 < \eta < R_0$, $s_0(\eta) > 1$.
(b) $\lim_{\lambda \to 0} P_{MD} = 0$ if for some $0 < \eta < R_1$, $s_1(\eta) > 1$.

Proof: Please see the appendix.

The following result is for heavy-tailed case. This is provided to show that if we do not robustify the observations at the local nodes and/or FC, the penalty for heavy-tailed EMI/outliers can be high. This holds for single node case also as demonstrated in Section IV. For the following theorem, we work with the random walk algorithm.

Theorem 4: If there is an $r_1 > 0$ and $r_2 > 0$ such that the distribution of $X_{1l} \in R(-r_1-1)$ for all $l = 1, 2, ..., L$ and $i = 0, 1$ and the distribution of $Z_l \in R(-r_2-1)$ then

$$P_{FA} \leq O(\log c^{-\min(r_1, r_2)})$$

$$P_{MD} \leq O(\log c^{-\min(r_1, r_2)})$$

Proof: Please see the appendix.

Thus we see that heavy-tailed noise/EMI affects the performance of $P_{FA}$ and $P_{MD}$.

![Energy detection for BPSK at −5dB (no fading)](image1)

![Energy detection for BPSK at −5dB in the presence of block shadowing–fast fading](image2)

Fig. 7: Energy detection in the presence of Gaussian noise. Top: Without fading. Bottom: Log $\mathcal{N}$ shadowing - Rayleigh fast fading.

VI. SIMULATION RESULTS FOR DISTRIBUTED ALGORITHM

We have considered $L = 5$ local nodes reporting their decisions to the FC. The distributions of fading, EMI and outliers at the local nodes and the FC are the same as in Section III. Also, $b_0 = 1, b_1 = 1$. The receiver noise at the local nodes is $\mathcal{N}(0,1)$ and at the FC is $\mathcal{N}(0,5)$. From Figures 7 and 8 we see that the distributed algorithm performs much better than the single node algorithm using $M^2$-random walk, especially in the low probability of error regime. Figure 7 shows the comparison when local nodes run $M^2$-random walk and FC runs $M$-random walk. The distributed algorithm performs better in the presence of EMI (along with shadowing-fading) at the local nodes and the FC (Figure 8). It also shows that the presence of 5% outliers (along with shadowing-fast fading) at the FC does not make a considerable difference in the performance of the distributed algorithm. We have considered the presence of outliers only when the signal is present (under $H_1$ at local nodes and under both the hypotheses at the FC). With all these impairments, the reporting channel becomes bad and the improvement over the single node case is seen for small probability of error only (Figure 8 bottom).

We see that the distributed algorithm performs much better than the single node $M^2$-random walk, especially at low probability of error. It is also quite robust to the effects of fading, EMI and outliers at the local nodes and the FC.

VII. CONCLUSION

We propose a distributed algorithm for spectrum sensing in Cognitive radio. Various impairments such as additive noise, EMI, shadowing and multipath effects and outliers were taken into account while designing the algorithm. The local nodes perform energy detection for lack of knowledge about the primary’s transmission parameters and the FC is signalled via BPSK. We find that robust versions of random walk algorithm developed recently, perform well in case of energy detection at the local nodes and binary signalling over the reporting MAC channel. We have performed simulations to demonstrate this and have theoretically validated the observations.

VIII. APPENDIX

The proofs of Theorems 1 – 4 are provided in this appendix. We will use the following notation:
\[ N_0^l = \inf \{ n : T_{nl} \leq -\gamma_0 l \} \]
\[ N_1^l = \inf \{ n : T_{nl} \geq \gamma_1 l \} \]
\[ \tau(-\gamma_0 l) = \text{the last time random walk } T_{nl} \text{ is above } -\gamma_0 l. \]
\[ \tau = \max_l \tau(-\gamma_0 l) \]
\[ \nu(a) = \text{Starting from } 0, \text{ the first time } W_k \text{ crosses } a \]
\[ \text{when all local nodes are transmitting } -b_0. \]
\[ \tau(c) = \max_l \tau(-\gamma_l | \log c|). \]

**Proof of Theorem 1.** We show \( P_0[N < \infty] = 1. \) Similarly we can show for \( P_l[N < \infty] = 1. \)

Under \( H_0, \) \( T_{nl} \) is a random walk with finite negative mean, for each \( l. \) Thus, \( T_{nl} \to -\infty \text{ a.s. and hence } \tau(-\gamma_0 l) < \infty \text{ a.s.} \)
for any finite \( \gamma_0 l. \) Therefore \( \tau < \infty \text{ a.s.} \) After \( \tau, \) all local nodes transmit \(-b_0 \) and hence increments of \( W_k \) have a negative mean \((= \Delta_0). \) Therefore,
\[
N < \tau + \nu(-W_{\tau+1} - \beta_0). \tag{16}
\]
Since \( \tau < \infty \text{ a.s. and } E[Y_k] < \infty, \) \( |W_{\tau+1}| < \infty \text{ a.s. and hence } \nu(-W_{\tau+1} - \beta_0) < \infty \text{ a.s.} \) Therefore, \( P_0[N < \infty] = 1. \)

**Proof of Theorem 2.** We prove for \( H_0. \) From (16),
\[
\frac{N}{|\log c|} \leq \frac{\tau(c)}{|\log c|} + \frac{\nu(-W_{\tau(c)+1} - |\log c|)}{|\log c|}.
\]
Also, from (32)
\[
\lim_{c \to 0} \frac{\tau(c)}{\log c} = \frac{1}{D_{\text{tot}}} \text{ a.s.}
\]
Furthermore,
\[
\frac{\nu(-|\log c| - W_{\tau(c)+1})}{|\log c|} \leq \frac{\nu(-|\log c|)}{|\log c|} + \frac{\nu(-W_{\tau(c)+1}) W_{\tau(c)+1}}{W_{\tau(c)+1}} \frac{\tau(c) + 1}{\tau(c) + 1} \frac{1}{|\log c|}
\]
and from (32),
\[
\lim_{c \to 0} \frac{\nu(-|\log c|)}{|\log c|} = -\frac{1}{\Delta_0} \text{ a.s.}
\]
and as in the proof of Theorem 1 in (11)
\[
\lim_{c \to 0} \frac{\nu(-W_{\tau(c)}) W_{\tau(c)}}{W_{\tau(c)} |\log c|} \frac{\tau(c)}{|\log c|} \leq \frac{(-1)}{\Delta_0} \frac{\mathbb{E}[\zeta_1]}{D_{\text{tot}}} \text{ a.s.}
\]

**Proof of Theorem 3.** We prove the result for \( P_{F,A}. \) It holds for \( P_{F,M}. \) in the same way.

We have,
\[
P_{F,A} = P_0[\text{Declare } H_1 \text{ up to } \tau(c)] + P_0[\text{Declare } H_1 \text{ after } \tau(c)].
\]
Consider the first term on the RHS. Let \( W^*_k \) be the random walk formed by the i.i.d. sequence \( \{\zeta_k^l, k \geq 0\} \) where \( \zeta_k^l \equiv \zeta^l_k. \) Then \( W^*_k \) stochastically dominates \( W_k \) and can be generated such that \( W^*_k \geq W_k \) a.s. Thus,
\[
P_0[\text{Declare } H_1 \text{ up to } \tau(c)] \leq P_0[\sup_{0 \leq k \leq \tau(c)} W^*_k \geq |\log c|]
\]
\[
= P(\sum_{k=1}^{\tau(c)+1} |\zeta_k^l| \geq |\log c|).
\]
Then, from (36), for \( 0 < \eta < R_0 \), \( \mathbb{E}[e^{\eta \tau(c)}] < \infty \) for each \( l. \) Since \( \tau(c) \leq \sum_{i=1}^{L} \tau_i(c) \) and \( \tau_i(c) \) are independent, this implies that \( \mathbb{E}[e^{\eta \tau(c)}] < \infty \) for \( 0 < \eta < \min_i R_0 = R_0. \) Therefore, by Markov inequality,
\[
P(\tau(c) > l) \leq k_1 e^{-\eta l}.
\]
Also, then from (37)
\[
P(\sum_{k=1}^{\tau(c)} |\zeta_k^l| \geq |\log c|) \leq k_2 e^{-s_0(\eta) |\log c|}
\]
for any \( 0 < \eta < R_0 \) where \( k_2 \) is a constant. Thus, from (18),
\[
P_0[\text{Declare } H_1 \text{ up to } \tau(c)] \leq k_2 e^{-s_0(\eta) / c} \to 0,
\]
if \( s_0(\eta) > 1 \) for some \( \eta. \) Furthermore, as in (18). Theorem 2,
\[
\lim_{c \to 0} P_0[\text{Declare } H_1 \text{ after } \tau(c)] = 0.
\]

**Proof of Theorem 4.** Define \( A_l = \{ \text{local node } l \text{ makes a wrong decision some time}\}. \)
Then from (11),
\[
P_0[\text{relative } A] \leq P_0[\sup_{n \geq 0} T_{nl} \geq \rho_1 |\log c|]
\]
\[
\sim f_1(\rho_1 |\log c| / \rho_1 |\log c|)^{-r_1}
\]
where \( f_1 \) is a slowly varying function. Thus,
\[
P_0[\text{Declare } H_1] = P_0[\cap l A_l] P_0[\text{Declare } H_1 | \cap l A_l]
\]
\[
+ \left( \sum_{l} f_1(\rho_1 |\log c| / \rho_1 |\log c|)^{-r_1} \right) \frac{1}{|\log c|^{r_1}}
\]
\[
\text{Consider}
\]
\[
P_0[\text{Declare } H_1 | \cap l A_l] = P_0[\text{Declare } H_1 \text{ up to } \tau(c) | \cap l A_l]
\]
\[
+ P_0[\text{Declare } H_1 \text{ after } \tau(c) | \cap l A_l]
\]
\[
\leq P_0[\sup_{0 \leq k \leq \tau(c)} \sum_{k=1}^{n} Z_k > |\log c|]
\]
\[
+ P_0[\text{Declare } H_1 \text{ after } \tau(c) | \cap l A_l]
\]
\[
\text{Also, from (35), since } Z_k \text{ distribution } \in R(-r_2 - 1),
\]
\[
P_0[\sup_{1 \leq k \leq \tau(c)} \sum_{k=1}^{n} Z_n > |\log c|],
\]
\[
\sim |\log c|^{r_2-1} - |\log c|^{-r_2} g_1(|\log c|) \text{ a.s.}
\]
where \( g_1 \) is a slowly varying function.

The second term on RHS of (20),
\[
P_0[\text{Declare } H_1 \text{ after } \tau(c) | \cap l A_l]
\]
\[
\leq P[\text{random walk at FC with mean } \Delta_0]
\]
\[
\text{and initial condition } \sum_{k=1}^{\tau(c)+1} Z_k \text{ crosses } |\log c|,
\]
\[
\leq P[\text{FC random walk with mean } \Delta_0]
\]
\[
\text{and initial condition } \delta \text{ crosses } |\log c|,
\]
\[
+ P[\sum_{k=1}^{\tau(c)+1} Z_k > (1 - \delta) |\log c|].
\]
Also, for slowly varying functions \( g_2, g_3 \),
\[
P[ \sum_{k=1}^{\infty} Z_k > (1 - \delta) | \log c |] \\
\leq P[ \sup_{0 \leq \xi \leq \tau(c) + 1} \sum_{j=0}^{k} Z_j > (1 - \delta) | \log c |] \\
\sim g_2((1 - \delta) | \log c |) \xi | \tau(c) + 1 | | \log c |^{-\gamma - 1} \\
\sim g_2((1 - \delta) | \log c |) | \log c |^{-\gamma}
\]
(22)

and
\[
P[\text{random walk with mean } -\Delta_0 \text{ and initial condition } \delta | \log c | \text{ crosses } | \log c |] \\
\leq g_3((1 - \delta) | \log c |) | \log c |^{-\gamma^*}. 
\]
(23)

From (20), (21), (22) and (23),
\[
P_{FA} \lesssim \left( \sum \frac{f_l(\rho_l | \log c |)}{\rho_l^2} \right) \frac{1}{(\log c)^\gamma} \\
+ g_1((1 | \log c |) | \log c |^{-\gamma}) \\
+ g_2((1 - \delta) | \log c |) | \log c |^{-\gamma} \\
+ g_3((1 - \delta) | \log c |) | \log c |^{-\gamma^*}. 
\]

\[\blacksquare\]

REFERENCES

[1] J. Mitola and J. Maguire, "Cognitive radio: making software radios more personal," IEEE Personal Communication, vol. 6, no. 4, pp. 13–18, Aug 1999.
[2] I. F. Akyildiz, B. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," Physical Communication, vol. 4, no. 1, pp. 40–62, 2011.
[3] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," Signal Processing Magazine, IEEE, vol. 29, no. 3, pp. 101–116, May 2012.
[4] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in 42nd Allerton Conference on Communication, Control, and Computing, 2004.
[5] Z. Govindarajulu, Sequential statistics. World Scientific Publishing Co., 2004.
[6] S. Atapattu, C. Tellambura, and H. Jiang, Energy Detection for Spectrum Sensing in Cognitive Radio. DOI 10.1007/978-1-4439-0494-5, Springer-Briefs in Computer Science, 2014.
[7] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," Proceedings of the IEEE, vol. 97, no. 2, pp. 205–230, 2009.
[8] J. Park, G. Shevlyakov, and K. Kim, "Maximun distributed detection in the presence of impulsive alpha-stable noise," IEEE Transactions on Wireless Communications, pp. 1687 – 1691, June 2011.
[9] P. J. Huber and E. M. Ronchetti, Robust Statistics. 2nd edition, A. John Wiley and sons Inc. publication, 2009.
[10] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge University press, 2005.
[11] J. D. Parsons, The mobile radio propagation channel. 2nd edition, John Wiley and Sons, 2000.
[12] A. Ghasemi and E. Sousa, “Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs,” IEEE Communications Magazine, vol. 46, no. 4, pp. 32–39, 2008.
[13] Y. Zeng, Y. C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," EURASIP Journal on Advances in Signal Processing, no. Article id 381465, 2010.
[14] 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, July 2009.
[15] J. Chamberland and V. Veeravalli, "Wireless sensors in distributed detection applications," IEEE Signal Processing Magazine, vol. 24, pp. 16–25, 2007.
[16] V. Veeravalli, "Sequential decision fusion: theory and applications," Journal of the Franklin Institute, vol. 336, pp. 301–322, 1999.
[17] T. Banerjee, V. Sharma, V. Kavitha, and A. Jayaprakasam, "Generalized analysis of a distributed energy efficient algorithm for change detection," IEEE Transactions on Wireless Communication, vol. 10, pp. 91–101, 2011.