1. INTRODUCTION

Cognitive radio has recently emerged as a useful technology to improve the efficiency of spectrum utilization [1]. Traditionally, the spectrum is assigned by the Federal Communications Commission (FCC) to specific users or applications, and each user can only utilize its pre-assigned bandwidth for communication. This discipline causes some bandwidth to be overcrowded while some other bandwidth may be seriously under utilized. The concept of cognitive radio aims at providing a flexible way of spectrum management, permitting secondary users to temporally access spectrum that is not currently used by legacy users. In this regard, the FCC has taken a number of steps towards allowing low-power devices to operate in the broadcast TV bands that are not being used by TV channels [2]. The TV bands include the following portions of the VHF and UHF radio spectrum: 54-72 MHz, 76-88 MHz, 174-216 MHz and 470-806 MHz. Each TV channel occupies a slot of 6 MHz bandwidth. If a TV frequency band is not used in a particular geographical region, it can be used by cognitive radios for transmission. To promote this development, IEEE has established the IEEE 802.22 Working Group to develop a standard for a cognitive radio-based device in TV bands.

A key challenge in the development of the IEEE 802.22 standard is that a cognitive radio should be able to reliably detect the presence of TV signals in a fading environment. Otherwise, the radio may use the frequency band that is occupied by a TV channel, and consequently cause serious interference to the TV receivers nearby. Many sensing or detection schemes have been recently reported in the IEEE 802.22 community, e.g., [3–6]. These schemes can be classified into two categories: single-user sensing and cooperative sensing. Due to the large variations in the received signal strength that are caused by path loss and fading, single-user sensing has proven to be unreliable, which consequently triggered the FCC to require geolocation-based methods for identifying unused frequency bands [7,8]. The geolocation approach is suitable for registered TV bands; however, its cost and operational overhead prevent its wide use in the opportunistic access to occasional “white spaces” in the spectrum. Cooperative sensing relies on multiple radios to detect the presence of primary users, and provides a reliable solution for cognitive radio networks.

The sensing performance is measured by two key factors: probability of detection errors and sensing time. A traditional way to design the sensing strategy is based on the Neyman-Pearson criterion, and the resulting test fixes the number of required samples, i.e., the sensing time. In the presence of multiple radios, the determination of the optimal threshold value for the test usually involves complex numerical computations. In this paper, we show that a sequential detector is more suitable for collaborative sensing. It is easy to implement and can significantly reduce the average sensing time. In the proposed scheme, each cognitive radio computes the Log-Likelihood ratio for its own signal-present events are defined as:

\[ \mathcal{H}_0 : \text{target signal is absent; } \mathcal{H}_1 : \text{target signal is present.} \]

The signal acquired by the \( m \)-th, \( m = 1, 2, \ldots, M \), cognitive radio device is represented by

\[
\mathcal{H}_0 : X_m[n] = W_m[n], \quad \mathcal{H}_1 : X_m[n] = S_m[n] + W_m[n], \quad n = 1, 2, \ldots,
\]
where $S_m[n]$ is the target signal and $W_m[n]$ is the additive noise. In the paper, we assume that the samples acquired by different radios are statistically independent, and that the samples acquired by the same radio are independent and identically distributed (i.i.d.). Under $H_0$ and $H_1$, the distribution of the acquired signal at the $m^{th}$, $m = 1, 2, \ldots, M$, radio is characterized by the probability density functions $p_{1,m}(X_m[n])$ and $p_{1,m}(X_m[n])$, respectively. The performance of detecting $H_0$ against $H_1$ is measured by the probability of false alarm and the probability of miss detection. False alarm refers to the error of accepting $H_1$ when $H_0$ is true, while miss detection refers to the error of accepting $H_0$ when $H_1$ is true. The probability of false alarm is represented by $P_{\text{FA}} = \Pr(H_1 | H_0)$ and the probability of miss detection is represented by $P_{\text{MISS}} = \Pr(H_0 | H_1)$, where $H$ represents the detector output.

2. SEQUENTIAL DETECTION WITH KNOWN STATISTICS

Assume that the number $N_{\text{fix}}$ of samples (acquired by each cognitive radio) is fixed. To detect $H_0$ and $H_1$, the Likelihood ratio test (LRT) is performed at the base station according to

\[
\begin{align*}
\text{Accept } H_1 & \text{ if } \text{LLR} > \eta \\
\text{Accept } H_0 & \text{ if } \text{LLR} \leq \eta
\end{align*}
\]

where the Log-Likelihood ratio LLR is computed as

$$\text{LLR} = \ln \left( \prod_{n=1}^{N_{\text{fix}}} \frac{p_{1,m}(X_m[n])}{p_{0,m}(X_m[n])} \right).$$

The threshold value $\eta$ is selected such that the probability of false alarm and the probability of miss detection are equal to or less than some pre-assigned values $\alpha$ and $\beta$, respectively, i.e.,

$$P_{\text{FA}} \leq \alpha, \quad P_{\text{MISS}} \leq \beta. \quad (2)$$

The determination of $\eta$ usually involves complex numerical computations or simulations.

To reduce the number of required samples, instead of using a fixed sample size $N_{\text{fix}}$, we can implement the LRT for every acquired sample in a sequential manner [11]. That is, for $N = 1, 2, \ldots$, we perform the following test:

\[
\begin{align*}
\text{Accept } H_1 & \text{ and terminate if } \text{LLR}_N \geq A \\
\text{Accept } H_0 & \text{ and terminate if } \text{LLR}_N \leq B \\
\text{Take one more sample to repeat the test if } B < \text{LLR}_N < A
\end{align*}
\]

where

$$\text{LLR}_N = \ln \left( \prod_{n=1}^{N} \prod_{m=1}^{M} \frac{p_{1,m}(X_m[n])}{p_{0,m}(X_m[n])} \right)$$

$$= \sum_{n=1}^{N} \sum_{m=1}^{M} \ln \left( \frac{p_{1,m}(X_m[n])}{p_{0,m}(X_m[n])} \right)$$

and $A$, $B$ are predetermined constants according to the sensing objective (2). In the context of cooperative sensing, each radio computes the Log-Likelihood ratio for its every acquired sample, and the base station sequentially accumulates the Log-Likelihood statistics and performs the above test, as described in Algorithm 1.

**Algorithm 1** Cooperative Sequential Detection

\begin{algorithm}[H]
0: Set $N = 0.$
0: Set LLR$_0$ = 0 at the base station.
1: repeat
2: $N = N + 1.$
3: The $m^{th}$, $m = 1, 2, \ldots, M$, radio acquires sample $X_m[n]$ and computes $\ln (p_{1,m}(X_m[n])/p_{0,m}(X_m[n]))$.
4: Each radio sends its $\ln (p_{1,m}(X_m[n])/p_{0,m}(X_m[n]))$ to the base station.
5: The base station updates the sequential Log-Likelihood ratio LLR$_N$ according to

$$\text{LLR}_N = \text{LLR}_{N-1} + \sum_{m=1}^{M} \ln \left( \frac{p_{1,m}(X_m[n])}{p_{0,m}(X_m[n])} \right).$$

6: until LLR$_N \geq A$ or LLR$_N \leq B$.
7: If LLR$_N \geq A$, “$H_1$: Target Signal is Present” is claimed; if LLR$_N \leq B$, “$H_0$: Target Signal is Absent” is claimed.

To see how $A$ and $B$ are determined, we need to study how $P_{\text{FA}}$ and $P_{\text{MISS}}$ depend on $A$ and $B$. Assume that the detection procedure terminates at $N = N_{\text{stop}}$. At $N = N_{\text{stop}}$, we have

$$\text{LLR}_{N_{\text{stop}}} \approx A \quad \text{or} \quad \text{LLR}_{N_{\text{stop}}} \approx B,$$

provided that the change in LLR$_N$ at each step is relatively small when compared to the absolute values of $A$ and $B$. This holds when $\alpha$ and $\beta$ are sufficiently small.

It can be shown that

$$P_{\text{FA}} = \Pr(\text{LLR}_{N_{\text{stop}}} \geq A | H_0) \approx \frac{1 - e^B}{e^A - e^B}, \quad (4)$$

$$P_{\text{MISS}} = \Pr(\text{LLR}_{N_{\text{stop}}} \leq B | H_1) \approx e^{-A} - \frac{1}{e^{-A} - e^{-B}}. \quad (5)$$

To find appropriate $A$ and $B$, we set (4) and (5) to be equal to $\alpha$ and $\beta$, respectively, and get

$$A = \ln \left( \frac{1 - \beta}{\alpha} \right), \quad B = \ln \left( \frac{\beta}{1 - \alpha} \right). \quad (6)$$

We can see that $A$ and $B$ do not depend on specific distributions and are convenient to compute.

In the sequential test, the number of samples required to reach a decision is a random variable. By using Wald’s equation, the average number of required samples under $H_0$ is given

\begin{footnotesize}
1. It can be shown that this would happen with probability one [11].

2. It is seen in (6) that $|A|$ and $|B|$ can be sufficiently large when $\alpha$ and $\beta$ are sufficiently small.
\end{footnotesize}
To summarize the results, we have the following theorem.

**Theorem 2.1** For the sequential sensing procedure defined in (3), if $\alpha$ and $\beta$ are sufficiently small and the thresholds $A$ and $B$ are set according to (6), then

$$P_{\text{FA}} \approx \alpha, \quad P_{\text{MISS}} \approx \beta$$

and the average number of samples required to reach a decision under $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by (7) and (8), respectively.

Expressions (7) and (8) show that the average required sensing time depends on the KL distance term provided by each sensing radio. Intuitively, the larger the KL distance term, the more the two hypotheses differ from each other, which thus requires less sensing time to detect. In order to save processing power and communication bandwidth, we may need to select a subset of available radios for spectrum sensing. Based on (7) and (8), we choose radios with larger KL distance in order to minimize

$$\max \left\{ \mathbf{E}_{\mathcal{H}_0} \{N_{\text{stop}}\}, \mathbf{E}_{\mathcal{H}_1} \{N_{\text{stop}}\} \right\}.$$  

### 3. SEQUENTIAL DETECTION WITH MODEL UNCERTAINTIES

In Section 2, we explained how to perform sequential sensing in a network when the distributions of the target signal and noise are exactly known. In practice, however, there usually exist uncertainties or errors in the assumed statistical models of signal and noise. To see how the sequential detector is implemented in a robust manner, we make the following assumptions:

1. Under $\mathcal{H}_0$, all possible distributions of $X_m[n]$ form a set denoted by $\Phi_{0,m}$, i.e., $\Phi_{0,m} = \{p_{0,m}(X_m[n])\}$.
2. Under $\mathcal{H}_1$, all possible distributions of $X_m[n]$ form a set denoted by $\Phi_{1,m}$, i.e., $\Phi_{1,m} = \{p_{1,m}(X_m[n])\}$.

For $m = 1, 2, \ldots, M$, we select a representative distribution for each of the two hypotheses, denoted by $p_{0,m}(X_m[n]) \in \Phi_{0,m}$ and $p_{1,m}(X_m[n]) \in \Phi_{1,m}$. The proposed robust test is obtained by replacing $p_{0,m}(X_m[n])$ with $\bar{p}_{0,m}(X_m[n])$ and $p_{1,m}(X_m[n])$ with $\bar{p}_{1,m}(X_m[n])$ in (3), i.e.,

- Accept $\mathcal{H}_1$ and terminate if $\text{LLR}_N \geq A$
- Accept $\mathcal{H}_0$ and terminate if $\text{LLR}_N \leq B$

Take one more sample to repeat the test if $B < \text{LLR}_N < A$ where

$$\text{LLR}_N = \sum_{n=1}^{N} \sum_{m=1}^{M} \ln \left( \frac{\bar{p}_{1,m}(X_m[n])}{\bar{p}_{0,m}(X_m[n])} \right)$$

and $A$, $B$ are given by (6). If the representative distributions satisfy the following conditions, we can still achieve the sensing objective (2):

(i) for all $p_{0,m}(X_m[n]) \in \Phi_{0,m}$, the solution $\theta_0^*$ to the equation

$$\mathbf{E}_{\mathcal{H}_0} \left( \frac{\prod_{m=1}^{M} \bar{p}_{1,m}(X_m[n])}{\prod_{m=1}^{M} p_{0,m}(X_m[n])} \right)^{\theta_0^*} = 1$$

is greater than or equal to 1, where

$$\mathbf{E}_{\mathcal{H}_0} \left( \frac{\prod_{m=1}^{M} \bar{p}_{1,m}(X_m[n])}{\prod_{m=1}^{M} p_{0,m}(X_m[n])} \right)^{\theta_0} \propto e^{-A}$$

$$\mathbf{E}_{\mathcal{H}_1} \left( \frac{\prod_{m=1}^{M} \bar{p}_{1,m}(X_m[n])}{\prod_{m=1}^{M} p_{0,m}(X_m[n])} \right)^{\theta_1} = 1$$

is less than or equal to 1, where

$$\mathbf{E}_{\mathcal{H}_1} \left( \frac{\prod_{m=1}^{M} \bar{p}_{1,m}(X_m[n])}{\prod_{m=1}^{M} p_{0,m}(X_m[n])} \right)^{\theta_1} \propto e^{-B}$$

For some $\Phi_{0,m}$ and $\Phi_{1,m}$, we are not able to find proper $p_{0,m}(X_m[n])$ and $p_{1,m}(X_m[n])$ satisfying the above conditions, e.g., when $\Phi_{0,m}$ and $\Phi_{1,m}$ have overlap. Nevertheless, for many practical cases, we can distinguish $\Phi_{0,m}$ and $\Phi_{1,m}$ such that proper $\bar{p}_{0,m}(X_m[n])$ and $\bar{p}_{1,m}(X_m[n])$ can be found. Under the existence of such $\bar{p}_{0,m}(X_m[n])$ and $\bar{p}_{1,m}(X_m[n])$, we have the following theorem.

**Theorem 3.1** Given $\bar{p}_{0,m}(X_m[n]) \in \Phi_{0,m}$, $\bar{p}_{1,m}(X_m[n]) \in \Phi_{1,m}$, and the above conditions (i), (ii), if $\alpha$ and $\beta$ are sufficiently small and the thresholds $A$ and $B$ are set according to (6), the sequential sensing procedure defined in (9) approximately achieves

$$P_{\text{FA}} \approx \frac{1 - e^{\theta_0^* B}}{e^{\theta_0^* A} - e^{\theta_0^* B}} \leq \alpha,$$

$$P_{\text{MISS}} \approx \frac{1 - e^{\theta_1^* A}}{e^{\theta_1^* B} - e^{\theta_1^* A}} \leq \beta,$$

and the average number of samples required to reach a decision under $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by (12) and (13) (see the equations at the top of next page), respectively.
4. EXAMPLE

In this section, we illustrate the proposed sequential sensing technique by an example. In the signal model given by (1), the target signal $S_m[n]$ and the noise $W_m[n]$ are i.i.d. Gaussian with mean zero and variances $\sigma^2_{S,m}$ and $\sigma^2_{W,m}$, respectively. Hence, $X_m[n]$ is i.i.d. Gaussian with mean zero under both $\mathcal{H}_0$ and $\mathcal{H}_1$. For convenience of notation, we denote the variances of $X_m[n]$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ by $\sigma^2_{0,m}$ and $\sigma^2_{1,m}$ ($\sigma^2_{1,m} > \sigma^2_{0,m}$), respectively, i.e.,

$$\sigma^2_{0,m} = \sigma^2_{W,m}, \quad \sigma^2_{1,m} = \sigma^2_{S,m} + \sigma^2_{W,m}.$$

For the test with a fixed sample size $N_{\text{fix}}$, the LRT turns out to be an energy detector, i.e.,

$$\text{Accept } \mathcal{H}_1 \text{ if } \sum_{n=1}^{N_{\text{fix}}} \sum_{m=1}^{M} \frac{1}{2} \left( \frac{1}{\sigma^2_{0,m}} - \frac{1}{\sigma^2_{1,m}} \right) (X_m[n])^2 > \eta',$$

$$\text{Accept } \mathcal{H}_0 \text{ if } \sum_{n=1}^{N_{\text{fix}}} \sum_{m=1}^{M} \frac{1}{2} \left( \frac{1}{\sigma^2_{0,m}} - \frac{1}{\sigma^2_{1,m}} \right) (X_m[n])^2 \leq \eta'.$$

To find the value of $\eta'$, we treat the summation $\sum_{n=1}^{N_{\text{fix}}} (X_m[n])^2$ approximately as Gaussian distributed by the central limit theorem and obtain the expression for the minimum $N_{\text{fix}}$ (see (14) at the top of next page) to achieve (2), where $Q(\cdot)$ is the Q-function of the standard normal distribution. The associated $\eta'$ is given by

$$\eta' \approx \frac{N_{\text{fix}}}{2} \sum_{m=1}^{M} \left( 1 - \frac{\sigma^2_{0,m}}{\sigma^2_{1,m}} \right) + \sqrt{\frac{N_{\text{fix}}}{2} \sum_{m=1}^{M} \left( 1 - \frac{\sigma^2_{0,m}}{\sigma^2_{1,m}} \right)^2} \cdot Q^{-1}(P_{\text{FA}}).$$

For the sequential test given in (3), we have

$$D(p_{0,m}(X_m[n])/p_{1,m}(X_m[n])) = \frac{1}{2\lambda_m} - \frac{1}{2\sigma^2_{0,m}} + \frac{1}{2\sigma^2_{1,m}},$$

$$D(p_{1,m}(X_m[n])/p_{0,m}(X_m[n])) = \frac{1}{2\lambda_m} - \frac{1}{2\sigma^2_{0,m}} - \frac{1}{2\sigma^2_{1,m}}.$$

where $\lambda_m = \sigma^2_{1,m}/\sigma^2_{0,m} > 1$. Both the two KL distance terms are monotonically increasing functions of $\lambda_m$. So we prefer sensing radios with large $\lambda_m$, which is consistent with common sense.

The modeling uncertainty arises when $\sigma^2_{0,m}$ and $\sigma^2_{1,m}$ are not exactly known. To account for this uncertainty, we specify a range for the signal and noise variances. Since $\sigma^2_{1,m} > \sigma^2_{0,m}$, we assume that $\sigma^2_{0,m} \leq \bar{\sigma}^2_{0,m}$ and $\sigma^2_{1,m} \geq \bar{\sigma}^2_{1,m}$ for some known $\bar{\sigma}^2_{0,m}$ and $\bar{\sigma}^2_{1,m}$ ($\bar{\sigma}^2_{1,m} > \bar{\sigma}^2_{0,m}$), where $\bar{\sigma}^2_{0,m}$ is the upper bound for the noise variances and $\bar{\sigma}^2_{1,m}$ is the lower bound for the signal variances. To apply the robust sequential test (9), we choose

$$\bar{\sigma}^2_{0,m}(X_m[n]) \sim \mathcal{N}(0, \bar{\sigma}^2_{0,m}), \quad \bar{\sigma}^2_{1,m}(X_m[n]) \sim \mathcal{N}(0, \bar{\sigma}^2_{1,m})$$

according to the required conditions and hence

$$\text{LLR}'_N = \sum_{n=1}^{N} \sum_{m=1}^{M} \left[ \left( \frac{1}{2\bar{\sigma}^2_{0,m}} - \frac{1}{2\bar{\sigma}^2_{1,m}} \right) (X_m[n])^2 + \frac{1}{2\ln \left( \frac{\bar{\sigma}^2_{0,m}}{\bar{\sigma}^2_{1,m}} \right) \cdot Q^{-1}(P_{\text{FA}})} \right].$$

5. SIMULATIONS AND CONCLUSIONS

The simulated network has 4 cognitive radios for spectrum sensing. The signal and noise are Gaussian distributed, as in the example of Section 4. The received signal variances under $\mathcal{H}_0$ and $\mathcal{H}_1$ are listed in Table 1. Fig. 2 compares the average number of required samples for the detection method with a fixed sample size and the sequential method, assuming that the detector has exact information about $\sigma^2_{0,m}$ and $\sigma^2_{1,m}$. Figs. 3 and 4 show the performance of the robust sequential sensing method in the presence of model uncertainties. Table 1 also lists the upper bound of the noise variances and the lower bound of the signal variances. In Case 1, the noise and signal variances are exactly equal to $\bar{\sigma}^2_{0,m}$ and $\bar{\sigma}^2_{1,m}$, respectively; in Case 2, the noise and signal variances are given by $\sigma^2_{0,m}$ and $\sigma^2_{1,m}$. It can be seen that the sequential algorithm performs as well as predicted by theory. For the most critical case that $\sigma^2_{0,m} < \bar{\sigma}^2_{0,m}$ and $\sigma^2_{1,m} > \bar{\sigma}^2_{1,m}$, the achieved $P_{\text{FA}}$ and $P_{\text{MSS}}$ are closely equal to the target values $\alpha$ and $\beta$. When $\sigma^2_{0,m} < \bar{\sigma}^2_{0,m}$ and $\sigma^2_{1,m} > \bar{\sigma}^2_{1,m}$, the achieved $P_{\text{FA}}$ and $P_{\text{MSS}}$ are less than the target values $\alpha$ and $\beta$, and the average number of required samples is less than that for the most critical case. Compared to the method that fixes the sample size, the sequential method needs fewer samples on average.

6. REFERENCES

\[ N_{\text{fix}} \approx 2 \left( \frac{\sqrt{\sum_{m=1}^{M} \left( 1 - \frac{\sigma^2_{0,m}}{\sigma^2_{1,m}} \right)^2} Q^{-1}(P_{\text{FA}}) + \sqrt{\sum_{m=1}^{M} \left( \frac{\sigma^2_{1,m}}{\sigma^2_{1,m}} - 1 \right)^2} Q^{-1}(P_{\text{MISS}})} {\sum_{m=1}^{M} \sigma^2_{0,m} + \sum_{m=1}^{M} \sigma^2_{1,m} - 2M} \right)^2 \]  \tag{14}

<table>
<thead>
<tr>
<th>( m = 1 )</th>
<th>( m = 2 )</th>
<th>( m = 3 )</th>
<th>( m = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma^2_{0,m} )</td>
<td>1.0</td>
<td>0.3</td>
<td>0.8</td>
</tr>
<tr>
<td>( \sigma^2_{1,m} )</td>
<td>1.05</td>
<td>0.33</td>
<td>0.92</td>
</tr>
<tr>
<td>( \sigma^2_{0,m} )</td>
<td>1.0125</td>
<td>0.3075</td>
<td>0.83</td>
</tr>
<tr>
<td>( \sigma^2_{1,m} )</td>
<td>1.0375</td>
<td>0.3225</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Table 1. Simulated signal and noise variances.

Fig. 2. Average number of required samples when the noise and signal variances are exactly known, i.e., \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1.

Fig. 3. Average number of required samples when only the bounds of the noise and signal variances are known, i.e., \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1. In Case 1, the actual noise and signal variances are \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1; in Case 2, the actual noise and signal variances are \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1.

Fig. 4. Achievable \( P_{\text{FA}} \) and \( P_{\text{MISS}} \) when only the bounds of the noise and signal variances are known, i.e., \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1. In Case 1, the actual noise and signal variances are \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1; in Case 2, the actual noise and signal variances are \( \sigma^2_{0,m} \) and \( \sigma^2_{1,m} \) in Table 1.


