Abstract—Spectrum sensing is a fundamental component of cognitive radio. How to promptly sense the presence of primary users is a key issue to a cognitive radio network. The time requirement is critical in that violating it will cause harmful interference to the primary user, leading to a system-wide failure. The motivation of our work is to provide an effective spectrum sensing method to detect primary users as soon as possible. In the language of streaming based real-time data processing, short-time means small-sized data. In this paper, we propose a cumulative spectrum sensing method dealing with limited sized data. A novel method of covariance matrix estimation is utilized to approximate the true covariance matrix. The theoretical analysis is derived based on McDiarmid’s concentration inequalities and random matrix theory to support the claims of detection performance. Comparisons between the proposed method and other traditional matrix theory to support the claims of detection performance. The theoretical analysis is derived based on McDiarmid’s concentration inequalities and random matrix theory to support the claims of detection performance. Comparisons between the proposed method and other traditional

Index Terms—cognitive radio, spectrum sensing, covariance matrix estimation, quickest detection, concentration inequality

I. INTRODUCTION

A limited natural resource, wireless spectrum becomes increasingly scarce due to the evolution of various wireless technologies. However, it is not utilized efficiently; the current utilization of a licensed spectrum varies from 15% to 85% [1]. The number is even lower in rural areas. Cognitive radio (CR) is a key technology to mitigate the overcrowding of spectrum space based on its capability to perform dynamic spectrum access (DSA). When a primary user (PU) starts its transmission, the secondary user (SU) must vacate the frequency band as soon as possible. SUs that fail to sense the occupied spectrum and vacate the spectrum in time will cause unexpected harmful interference to PUs and even damage the whole cognitive radio network. With SUs frequently moving between regions with different densities of PUs, such as in vehicular applications, rapid PU detection is of great importance. Standards exist to address this detection requirement. For example, the IEEE 802.22 standard for unlicensed operation in the TV band regulates that PUs should be detected within 2 seconds of their appearance [2].

As we have seen, it is clear that one fundamental requirement of a CR system is for SUs to use spectrum sensing to find spectral holes. Each SU should be able to sense a PU’s existence accurately to avoid interference, even when the PU’s signal is weak. Looking at it in this light can lead one to see how spectrum sensing could be treated as a signal detection problem. There has been plenty of research on spectrum sensing using classical detection schemes, such as energy detection [3]–[5], matched filter detection [6]–[8], cyclostationary feature detection [9]–[11], and covariance matrix based detection [12]–[15]. Covariance matrix related spectrum sensing algorithms were extended by employing multiple antennas at the cognitive receiver [16]. A suboptimal multi-antenna detector under unknown noise has also been proposed [17]. Feature template matching (FTM) [18] extracts signal features as the leading eigenvector of signal’s covariance matrix. The feature is stable over time for non-white wide-sense stationary (WSS) signals while random for white noise [19]. Kernel feature template matching (KFTM) [20] extended the linear FTM to a nonlinear FTM by mapping data from a input space to a high dimensional feature space. The mapping is implemented by the so-called kernel trick. Applications of kernel-based learning in cognitive radio network have been proposed in literature [21], the algorithms after kernel mapping have gained significant performance improvements over their linear counterparts at the price of generally higher computational complexity. Generalized function of matrix detection (FMD) has been employed for spectrum sensing, through the use of the function of random matrix and matrix inequality [22]–[24]. A two-dimensional sensing framework has been proposed for spatial-temporal opportunity detection in cognitive radio [25], which exploits correlations in time and space simultaneously by fusing sensing results in a spatial-temporal sensing window. A selective-relay-based cooperative sensing scheme has been proposed for both the spectrum sensing and secondary transmissions to achieve a reliable and efficient cognitive radio system [26], in which a dedicated channel usually used for reporting initial detection results for fusion is not essentially needed.

The vast majority of the research on spectrum sensing required large-sized datasets for processing to make a final decision. It was difficult to solve the sensing problem with limited received signal samples under low SNR. To circumvent this difficulty, we try to explore and utilize the core idea of quickest detection. Quickest detection [27] tries to detect the change of two different random processes with the shortest delay. If the
change happens at the beginning of spectrum sensing, the goal of quickest detection is similar to that of sequential detection. The successive refinement algorithm was proposed which combined both the generalized likelihood ratio (GLR) and parallel cumulative sum (CUSUM) tests for quickest spectrum sensing [28]. This algorithm used only average run lengths to measure the detection delay performance without the results on detection probability performance. Collaborative quickest spectrum sensing via random broadcast was investigated by first deriving a necessary condition for the optimal broadcast probability versus SNR, then proposing a threshold broadcast scheme [29]. This algorithm used ROC curve of average detection delay and false alarm rate as the performance metric, but no ROC curve of detection probability and false alarm rate was provided. Besides, this reference did not consider the impact of the change of SNR. A sequential change detection framework for quickest detection has also been established for cognitive radio systems [30]. A hidden markov model (HMM) for quickest detection has been proposed for spectrum sensing, and the effectiveness has been verified by the experimental tests using industrial standard wireless communications signal [31]. However, this reference failed to show the superiority of the proposed algorithm over other algorithms. Linear-based CUSUM statistics for different cooperative sensing scenarios with unknown parameters of the distribution have also been investigated in [32]. A fast spectrum sensing algorithm based on the discrete wavelet packet transform has been proposed in [33]. However, this algorithm focused only on the coarse detection and needed a fine detection stage to complete the whole spectrum sensing.

In this paper, we propose a cumulative spectrum sensing method with small datasets. This method works as a real-time processing of streaming data; every incoming sample generates a metric value to compare with the threshold in real-time, thus minimizing the sensing latency. The input data are first collected by the system, after which this sequential data is transformed into a sample covariance matrix. Because the size of the datasets is too small for calculating an accurate sample covariance matrix, oracle-approximating shrinkage estimation is utilized to get an accurate estimate of the true covariance matrix. The cumulative average is then incorporated to smooth the detection metric. Two performance metrics, the number of sample data needed for detection versus SNR and detection probability versus SNR, are used to support the superiority of the cumulative spectrum sensing method.

The contributions of this paper are as follows:

1) A novel cumulative spectrum sensing method has been proposed for detection with small-sized datasets. This method is efficient and effective in practice in that it can blindly detect the PU’s signal without knowing any information about the noise or the PU’s signal. Meanwhile, this method can work in a relatively low SNR environment given that the sampling period is limited. In other words, the total sample data is fixed.

2) The analysis based on McDiarmid’s concentration inequalities of statistics has been performed to demonstrate the superiority and effectiveness of the detection performance. The statistics on both hypotheses have been found to converge to distinct constant mean values. As a result, the signal and noise can be distinguished.

3) A threshold based on the false alarm probability has been proven to be stable without noise uncertainty problem. The probability distribution function of the test statistic under hypothesis $H_0$ has been derived, which is found approximately to be a Gaussian distribution.

The organization of this paper is as follows: in Section II a system model based on a binary hypothesis test and sample covariance matrix calculation is described; in Section III the cumulative spectrum sensing method with small datasets is presented with the corresponding mathematical foundations; Section IV presents a performance analysis including concentration inequalities of statistics and robust threshold; Section V provides the numerical results using a DTV signal, and comparisons are made with other popular detection methods.

**Notation:** In the following, we depict vectors in lowercase boldface letters and matrices in uppercase boldface letters. $(\cdot)^T$ means the transpose operator and $\text{Tr} (\cdot)$ means the trace operator. $\|\cdot\|_F$ represents the Frobenius norm and $\inf (\cdot)$ represents the infimum.

## II. System Model

### A. Binary Hypothesis Test

In a secondary network, we consider each SU with one receive antenna to detect one PU’s signal based on its own observation. Let $x(t)$ be the continuous-time received signal after unknown channel. Let $T_s$ be the sampling period. The received signal sample is

$$x[n] = x(nT_s)$$

(1)

There are two hypotheses to detect PU’s signal’s existence, $H_0$, only the noise (no PU’s signal) exists; and $H_1$, both the PU’s signal and the noise exist. The received signal samples under the two hypotheses are given respectively as follows:

$$H_0 : x[n] = w[n]$$

(2)

$$H_1 : x[n] = s[n] + w[n]$$

(3)

where $w[n]$ is the received white Gaussian noise, and each sample of $w[n]$ is assumed to be independent identical distribution (i.i.d.), with zero mean and variance $\sigma^2_w$. $s[n]$ is the received PU’s signal samples after unknown channel with unknown signal distribution. Though in practice, the noise $w[n]$ after analog-to-digital (ADC) is usually non-white, we can use pre-whitening techniques to whiten the noise samples. In the rest of this paper, the noise is considered white.

Two probabilities of interest are used to evaluate detection performance. One is detection probability $P_d$, that is, at hypothesis $H_1$, the probability having detected the PU’s signal. The other is false alarm probability $P_{fa}$, the probability having detected the PU’s signal at hypothesis $H_0$. Apparently, we want to obtain a high $P_d$ and a low $P_{fa}$. The requirements of $P_d$ and $P_{fa}$ depend on applications.
B. Sample Covariance Matrix

Assume spectrum sensing is performed based on the statistics of the \(i\)th sensing segment \(\Gamma_{x,i}\), which consists of \(N_{tot}\) total sample data. The sensing segment \(\Gamma_{x,i}\) can be formed as \(N\) sensing vectors with \(L\) (called “smoothing factor”) consecutive output samples in each vector:

\[
\Gamma_{x,i} = \{x_{(i-1)N+1}, x_{(i-1)N+2}, \ldots, x_{(i-1)N+N}\}
\]

\(x_i = [x[i], x[i+1], \ldots, x[i+L-1]]^T\)

where \(x_i \sim \mathcal{N}(\bar{x}, \Sigma_x)\) is the mean of \(x_i\) and \(\Sigma_x\) is covariance matrix of \(x_i\). Thus we have the equation

\[
N_{tot} = N + L - 1
\]

Here, to distinguish \(N_{tot}\) and \(N\), \(N_{tot}\) is named total data size, and \(N\) is named sample size (the number of columns of sensing segment matrix). \(\Gamma_{x,i}\) and \(s_i\) are defined in the same way as \(\Gamma_{x,i}\) and \(x_i\). The graphical illustration of sensing segment and sensing vector is provided in Fig. 1 to facilitate the understanding of how to form the data matrix.

In practice, the covariance matrix of the observed signals is unknown. Thus, the unstructured classical estimator of \(\Sigma_x\), the sample covariance matrix, is adopted and defined as

\[
\hat{\Sigma}_x = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \bar{x})^T
\]

Here, we assume the sample mean to be zero,

\[
\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i = 0
\]

Then, the sample covariance matrix is simplified as

\[
\hat{\Sigma}_x = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^T
\]

Large sample work in multivariate analysis has traditionally assumed that \(\frac{N}{L}\), the number of observations per variable, is large. Today, it is common for \(L\) to be large or even huge, and so \(\frac{N}{L}\) may be moderate to small and in extreme cases less than one [33]. In such case, an appropriate covariance matrix estimation is essentially needed.

III. SPECTRUM SENSING WITH SMALL-SIZED DATASETS

Detecting the presence of PU’s signal promptly is the basis of cognitive radio network. As soon as the PU is detected in the home channel, the SU has to vacate its home channel immediately. It is apparent the time requirement for cognitive radio system is of crucial importance to avoid interference. As a result, spectrum sensing should be performed in a short time. As the sampling rate is fixed for SU, sampling in a short time will only generate limited total size data \(N_{tot}\). In some sense, detection in short time is equivalent to detection using small-sized datasets. In the following, we will focus more on the data sample size than the detection time, since the sample size is directly involved into the algorithm design.

A. Algorithm Fundamentals

When \(N_{tot}\) is small, the sample size \(N\) would be comparable to matrix dimension \(L\), even \(N < L\). In such case, the sample covariance matrix \(\hat{\Sigma}_x\) is known to be a poor estimator of \(\Sigma_x\), which cannot describe the accurate statistical relationship within each sample. Many shrinkage estimators have been proposed under different performance measures by minimizing the mean-squared error (MSE) to approximate the true covariance matrix. In oracle-approximating shrinkage (OAS) estimation [35], the estimator \(\hat{\Sigma}\) is a trade-off between low bias and low variance, which is the solution to

\[
\min_{\rho} \mathbb{E} \left\{ \| \hat{\Sigma} - \Sigma_x \|_F^2 \right\}
\]

s.t. \(\hat{\Sigma} = (1 - \rho) \hat{\Sigma}_x + \rho \hat{F}\)

where \(\hat{\Sigma}_x\) is the sample covariance matrix defined in Eq. (7). The matrix \(\hat{F}\) is referred to as the shrinkage target, defined as

\[
\hat{F} = \frac{\mathrm{Tr} (\hat{\Sigma}_x)}{L} \mathbf{I}
\]

where \(\mathbf{I}\) is a \(L\) dimensional unitary matrix. Shrinkage coefficient \(\rho\), usually between 0 and 1, is aimed at minimizing the MSE. The solution is shown in Theorem 1 and the immediately followed iterations [36].

**Theorem 1:** Let \(\hat{\Sigma}_x\) be the sample covariance of a set of \(L\)-dimensional vectors \(\{x_i\}_{i=1}^{N}\). If \(\{x_i\}_{i=1}^{N}\) are i.i.d. Gaussian vectors with covariance \(\Sigma_x\), then the solution to (10) is

\[
\rho_0 = \frac{(1 - \frac{2}{L}) \mathrm{Tr} (\hat{\Sigma}_x) + \mathrm{Tr}^2 (\hat{\Sigma}_x)}{(N + 1 - \frac{2}{L}) \mathrm{Tr} (\hat{R}_x) + (1 - \frac{2}{L}) \mathrm{Tr}^2 (\hat{R}_x)}
\]

Since \(\hat{R}_x\) is hard to obtain, the OAS estimator is trying to approximate the solution in Eq. (12) via an iterative procedure. It initializes iterations with an initial guess of \(\hat{R}_x\) and iteratively refine it. The iteration procedure is continued until convergence, which is

\[
\hat{\rho}_{t+1} = \frac{(1 - \frac{2}{L}) \mathrm{Tr} (\hat{\Sigma}_{t} \hat{R}_x) + \mathrm{Tr}^2 (\hat{\Sigma}_{t})}{(N + 1 - \frac{2}{L}) \mathrm{Tr} (\hat{\Sigma}_{t} \hat{R}_x) + (1 - \frac{2}{L}) \mathrm{Tr}^2 (\hat{\Sigma}_{t})}
\]
The initial guess $\hat{\Sigma}_0$ could be the sample covariance matrix $\hat{\mathbf{R}}_x$. The initial guess $\hat{\rho}_0$ could be any value between 0 and 1. Here, $\text{Tr}(\hat{\Sigma}_j^2)$ is replaced by $\text{Tr}(\hat{\Sigma}_j \hat{\mathbf{R}}_x)$ since $\text{Tr}(\hat{\Sigma}_j^2)$ would always force $\hat{\rho}_j$ to converge to 1 while $\text{Tr}(\hat{\Sigma}_j \hat{\mathbf{R}}_x)$ is not.

When the above iteration converges, we can get the following estimation:

$$\hat{\rho}_{\text{OAS}} = \min \left( \frac{(1 - \frac{2}{N}) \text{Tr} (\hat{\mathbf{R}}_x^2) + \text{Tr}^2 (\hat{\mathbf{R}}_x)}{(N + 1 - \frac{2}{N}) \text{Tr} (\hat{\mathbf{R}}_x^2) - \frac{\text{Tr}^2 (\hat{\mathbf{R}}_x)}{L}}, 1 \right)$$  

(15)

In addition, $0 < \hat{\rho}_{\text{OAS}} < 1$.

After using $\hat{\rho}_{\text{OAS}}$ to substitute $\rho$ in (10), we can get the estimated covariance matrix as

$$\hat{\Sigma}_{\text{OAS}} = (1 - \hat{\rho}_{\text{OAS}}) \hat{\mathbf{R}}_x + \hat{\rho}_{\text{OAS}} \hat{\mathbf{F}}$$  

(16)

Eigenvalues of sample covariance matrix are widely used in detection. The maximum-minimum eigenvalue (MME) algorithm [12] uses the ratio of maximum and minimum eigenvalue, obtained from sample covariance matrix, as the detection metric. However, if the data size is not huge enough, the detection performance of MME will be compromised. It is natural to replace the eigenvalues from the sample covariance matrix with the eigenvalues from the estimated covariance matrix. The eigenvalues decomposed from $\hat{\Sigma}_{\text{OAS}}$ are denoted as: $\lambda_1 \geq \cdots \geq \lambda_L$. The maximum eigenvalue $\lambda_1$ and minimum eigenvalue $\lambda_L$ are defined as

$$\lambda_1 = \max_{\mathbf{u}^H \mathbf{u} = 1} \mathbf{u}^H \hat{\Sigma}_{\text{OAS}} \mathbf{u}$$  

(17)

$$\lambda_L = \max_{\dim(\psi) = L} \min_{\mathbf{u}^H \mathbf{u} = 1} \mathbf{u}^H \hat{\Sigma}_{\text{OAS}} \mathbf{u}$$  

(18)

where $\text{dim}(\psi)$ denotes the dimension of the subspace $\psi$.

The ratio of maximum and minimum eigenvalue can be calculated as

$$T = \frac{\lambda_1}{\lambda_L}$$  

(19)

The CUSUM test [37] is the optimal solution for minimizing delay and the central algorithm of non-Bayesian quickest detection, which requires the perfect knowledge of the distribution. In this paper, the distribution parameters of $H_1$ are unknown. However, the fundamental idea of CUSUM test can be simplified and utilized here.

The stopping time for detecting the change is defined by

$$t^{\text{stop}} = \inf \{ t \mid Q_t \geq \gamma \}$$  

(20)

where the $\gamma$ is the detection threshold and $Q_t$ is the detection statistic at time slot $t$. As we have discretized the time serials data into digital samples in Eq. (1), similar to the above formula, the stopping total data samples for detecting the presence of PU’s signal is given by

$$N_{\text{stop}}^{\text{stop}} = \inf \{ N_{\text{stop}} \mid Q_{N_{\text{stop}}} \geq \gamma \} + L - 1$$  

(21)

Here, $Q_{N_{\text{stop}}}$ is the metric for detection when $N_{\text{stop}}$ sample size are involved and defined as

$$Q_{N_{\text{stop}}} = \frac{q_{N_{\text{stop}}}}{N_{\text{stop}}}$$  

(22)

In some cases, the environment is so harsh that all the received data will be processed for the spectrum sensing decision, so $N_{\text{stop}}$ simply equals to $N$, which leads to

$$Q_N = \frac{q_N}{N}$$  

(23)

$q_r$ can be computed recursively:

$$q_r = \max (q_{r-1} + T_r, 0), 1 \leq r \leq N$$  

(24)

where $T_r$ is defined by (19) when $r$ sample size is used for covariance matrix calculation. The initial value of $q_0$ is 0.

We noticed that $T_r$ is positive under both hypotheses, thus (24) is equivalent to

$$q_r = \sum_{k=1}^{r} T_k$$  

(25)

and (23) can also be written as

$$Q_N = \frac{1}{N} \sum_{k=1}^{N} T_k$$  

(26)
B. Proposed Algorithm and System Architecture

Base on the above analysis, we propose the cumulative spectrum sensing approach with small-sized datasets, which is summarized as Algorithm 1. Meanwhile, the overall processing module architecture and the data flow diagram are shown in Fig. 2. A sequence of data are streamed into the system, every new coming sample will generate a new detection metric value, until the decision is made. The data streaming and detection are working simultaneously. The minimal required number of total data size is \( L \), so that at least one multivariate vector is contained in the sensing segment.

Algorithm 1 Cumulative Spectrum Sensing With Small-Sized Datasets

1: \( \gamma \leftarrow \) Set the detection threshold
2: \( \Gamma_{x,t} \leftarrow \) Sensing segment
3: Initial:
4: \( k = 1 \), \( q_0 = 0 \)
5: while True do
6: \( \hat{R}_{x,k} \leftarrow \) Calculate sample covariance matrix using Eq. \ref{eq:1} where \( N \) is set to be \( k \)
7: \( \hat{F}_k, \hat{\rho}_{OAS_k} \leftarrow \hat{R}_{x,k} \)
8: \( \Sigma_{OAS,k} \leftarrow \hat{R}_{x,k}, \hat{F}_k, \hat{\rho}_{OAS_k} \)
9: \( \lambda_{k}, \lambda_{L,k} \leftarrow \Sigma_{OAS,k} \) // Eigen-decomposition of \( \Sigma_{OAS,k} \)
10: \( T_k \leftarrow \lambda_{k}/\lambda_{L,k} \)
11: \( q_k \leftarrow q_{k-1}, T_k \)
12: \( \hat{q}_k \leftarrow q_k/k \)
13: if \( \hat{q}_k > \gamma \) then
14: \( \) PU exists and SU vacates the channel
15: \( \) return
16: else if \( k \geq N \) then
17: \( \) PU does not exist
18: \( \) return
19: else \( k = k + 1 \)
20: end if
21: end while

The advantage of Algorithm 1 is twofold. First, in a specific situation where the environment is not so harsh, the number of stopping total data sample \( N_{tot} \) is sufficient to detect the PU promptly once the threshold is reached. The goal is to detect as quickly as possible with minimal delay, which is particularly useful for detection in vehicular applications. Second, if we use all the received data for the detection decision, the proposed algorithm is able to work in a relatively low SNR environment. Three essential properties of this algorithm are worth mentioning here:

1) For detection under low SNR environment, more data brings better performance.
2) The threshold is robust which is not related to the noise power.
3) The algorithm is blind without any knowledge of the signal or the noise.

Both covariance matrix estimation and cumulative iteration contribute to the proposed algorithm. In order to show how much each part contributes in this Algorithm 1, we propose an additional Algorithm 2 which merely involves covariance matrix estimation. The initial \( k \) is set to be \( N \); \( T_N \) then can be obtained as before, after that \( T_N \) is directly compared with the threshold \( \gamma \) to make a decision. The performance of Algorithm 2 will be provided in Section IV as well as a reference.

IV. PERFORMANCE ANALYSIS

A. Concentration Inequalities of Statistics

The concentration inequalities of statistics will be analyzed in this section for the proposed algorithm. The detailed proof will be given based on the following (simplified version of the) theorem by McDiarmid [38], [39].

Theorem 2: Let \( x_1, \ldots, x_N \) be independent random variables taking values in a set \( A \), and let \( f : \mathbb{R}^N \rightarrow \mathbb{R} \) be a measurable function such that these is a constant \( c \) with

\[
|f(x_1, \ldots, x_m, \ldots, x_N) - f(x_1, \ldots, x_m', \ldots, x_N)| \leq c
\]

for all \( 1 \leq m \leq N \), \( x_1, \ldots, x_m, x_m', \ldots, x_N \in A \), and the sequence \( x_1, \ldots, x_m, \ldots, x_N \) and \( x_1, \ldots, x_m', \ldots, x_N \) differ only in the \( m \)th co-ordinate. Then for all \( t > 0 \),

\[
P(|f(x_1, \ldots, x_N) - E(f(x_1, \ldots, x_N))| \geq t) \leq 2 \exp \left(-\frac{2t^2}{Nc^2}\right)
\]

Let sample covariance matrix based on sequence \( x_1, \ldots, x_m, \ldots, x_N \) be defined as

\[
\hat{R}_{x,m} = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^T + x_m x_m^T
\]

where \( m = 1, \ldots, N \).

Accordingly, the OAS estimated covariance matrix is obtained as

\[
\hat{\Sigma}_{OAS} = (1 - \hat{\rho}_{OAS}) \hat{R}_{x,m} + \hat{\rho}_{OAS} \hat{F}'
\]

The above equation can be treated as a matrix perturbation to the original OAS estimated covariance matrix, written as

\[
\hat{\Sigma}'_{OAS} = \hat{\Sigma}_{OAS} + \hat{E}
\]

where \( \hat{E} \), which is Hermitian, is a perturbation matrix to \( \hat{\Sigma}_{OAS} \).

The eigenvalues obtained from \( \hat{\Sigma}'_{OAS} \) and \( \hat{E} \) are \( \lambda_1' \geq \cdots \geq \lambda_L' \) and \( \lambda_1(\hat{E}) \geq \cdots \geq \lambda_L(\hat{E}) \), respectively.

Theorem 3: [40, p. 34] Let \( A, B \) be Hermitian matrices with eigenvalues \( \lambda_1(A) \geq \cdots \geq \lambda_L(A) \) and \( \lambda_1(B) \geq \cdots \geq \lambda_L(B) \), respectively. Then,

\[
\lambda_j(A) + \lambda_L(B) \leq \lambda_j(A + B) \leq \lambda_j(A) + \lambda_L(B)
\]

Theorem 4: [40, p. 34] Let \( A, B \) be Hermitian matrices with eigenvalues \( \lambda_1(A) \geq \cdots \geq \lambda_L(A) \) and \( \lambda_1(B) \geq \cdots \geq \lambda_L(B) \), respectively. Then,

\[
\max_j |\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|
\]

where \( \|A - B\| = \max \{|\lambda_1(A - B)|, |\lambda_L(A - B)|\} \)
Since $\Sigma_{QAS}$ and $E$ are both Hermitian, substituting $j$ with 1 and $L$ into the inequality of Theorem 3 leads to the following results:

\[
\lambda_1 + \lambda_L (E) \leq \lambda_1' \leq \lambda_1 + \lambda_1 (E) \tag{33}
\]
\[
\lambda_L + \lambda_L (E) \leq \lambda_L' \leq \lambda_L + \lambda_1 (E) \tag{34}
\]

Remember now that

\[
\|E\| \leq \lambda_{L,k} \tag{ii} \leq 1
\]

The result is also based on Theorem 3 similar to Case 1.

Consequently,

\[
|f(x_1, \ldots, x_m, \ldots, x_N) - f(x_1, \ldots, x_m', \ldots, x_N)| \leq \max_k \left\{ \frac{|E|}{\lambda_{L,k}} \max \left( \frac{\lambda_{1,k}}{\lambda_{L,k}}, 1 \right) \right\} \tag{42}
\]

Based on Theorem 2, we finally obtain that for all $t > 0$

\[
P (|f(x_1, \ldots, x_N) - E(f(x_1, \ldots, x_N))| \geq t) \leq 2 \exp \left( \frac{-2t^2}{Nc^2} \right) \tag{43}
\]

where $c = \max_k \left\{ \frac{|E|}{\lambda_{L,k}} \max \left( \frac{\lambda_{1,k}}{\lambda_{L,k}}, 1 \right) \right\}$.

This concentration analysis shows that for a small value of $t$, with high probability

\[
E(Q_N) - t \leq Q_N \leq E(Q_N) + t \tag{44}
\]

which indicates $Q_N$ highly concentrates at its mean value $E(Q_N)$. Therefore, under null hypothesis and alternative hypothesis

\[
E(Q_{N,0}) - t \leq Q_N \leq E(Q_{N,0}) + t \tag{45}
\]
\[
E(Q_{N,1}) - t \leq Q_N \leq E(Q_{N,1}) + t \tag{46}
\]

where $E(Q_{N,0})$ and $E(Q_{N,1})$ are the mean values under $H_0$ and $H_1$. The detection statistic is able to discriminate between these two hypotheses when $E(Q_{N,1}) > E(Q_{N,0})$, denoted as

\[
H_0: Q_N \leq E(Q_{N,0}) + t \tag{47}
\]
\[
H_1: Q_N \geq E(Q_{N,1}) - t \tag{48}
\]

Two thousand Monte-Carlo simulations for statistics under both hypotheses are shown in Fig. 3 with SNR = -5 dB and $N = 300$. Fig. 3 shows that the statistics concentrate at two different values under null and alternative hypotheses, meanwhile $E(Q_{N,1})$ is greater than $E(Q_{N,0})$. 
B. Robust Threshold

Generally, we have no information on the signal, it is difficult to set the threshold based on the $P_d$. Thus, we choose the threshold based on $P_{fa}$. In our proposed algorithm, the sample covariance matrix is obtained with very limited samples which is inaccurate to true covariance matrix. After the covariance matrix estimation, the effect of OAS estimated covariance matrix $\Sigma_{OAS}$ is equivalent to sample covariance matrix $\mathbf{R}_x(N_{eq})$ with large number of samples, denoted as $N_{eq}$, though we don’t know exactly how large $N_{eq}$ is. Hence, we can use the distribution of eigenvalues obtained from $\mathbf{R}_x(N_{eq})$ to approximate the distribution of eigenvalues obtained from $\Sigma_{OAS}$.

**Theorem 5:** [13], [41] Assume that noise is real. Let $A(N_{eq}) = \frac{N_{eq}}{\sigma^2} \mathbf{R}_x(N_{eq})$, $\mu = \left(\sqrt{N_{eq} - 1 + \sqrt{L}}\right)$ and $\nu = \left(\sqrt{N_{eq} - 1 + \sqrt{L}}\right) \left(\frac{1}{\sqrt{N_{eq} - 1}} + \frac{1}{\sqrt{L}}\right)^{\frac{1}{2}}$. Then, $\frac{\lambda_1(A(N_{eq}) - \mu)}{\sqrt{\lambda_1(A(N_{eq}) - \mu)}}$ converges to the Tracy-Widom distribution of order 1 ($W_1$).

The mean and variance of Tracy-Widom distribution of order 1 can be found [42] to be $\mu_{tw} = -1.20653$ and $\sigma^2_{tw} = 1.60778$. It’s easy to obtain the mean and variance of $\lambda_1(A(N_{eq}))$ as $\mu + \nu \mu_{tw}$ and $\nu^2 \sigma^2_{tw}$, respectively. And hence, the the mean and variance of $\frac{\lambda_1(\mathbf{R}_x(N_{eq}))}{\sqrt{N_{eq} - 1}}$ are written as

$$\frac{\sigma^2}{N_{eq}} \text{ and } \frac{\sigma^4 \nu^2}{N_{eq}}$$ respectively.

**Theorem 6:** [13], [43] Assume that $\lim_{N_{eq} \to \infty} \frac{L}{N_{eq}} = y (0 < y < 1)$. Then $\lim_{N_{eq} \to \infty} \lambda_1(\mathbf{R}_x(N_{eq})) = \sigma^2 (1 - \sqrt{y})^2$ (with probability one).

Based on the Theorem 6 the smallest eigenvalue of $\mathbf{R}_x(N_{eq})$ tend to be deterministic value $\frac{\sigma^2}{N_{eq}} \left(\sqrt{N_{eq} - 1} - \sqrt{L}\right)^2$ when $N_{eq}$ is large. In such case, $\frac{\lambda_1(\mathbf{R}_x(N_{eq}))}{\lambda_3(\mathbf{R}_x(N_{eq}))}$ can be viewed as a new random variable $T$ obtained from random variable $\lambda_1(\mathbf{R}_x(N_{eq}))$ with a coefficient $\frac{1}{\lambda_3(\mathbf{R}_x(N_{eq}))}$.

As a result, the mean and variance of $T$ are written as

$$\mu_T = \frac{\left(\sqrt{N_{eq} - 1 + \sqrt{L}}\right)^2}{\left(\sqrt{N_{eq} - 1} - \sqrt{L}\right)^2} \frac{\left(\frac{1}{\sqrt{N_{eq} - 1}} + \frac{1}{\sqrt{L}}\right)^{\frac{1}{2}} \mu_{tw}}{(\sqrt{N_{eq} - 1} - \sqrt{L})^{\frac{1}{2}}}$$

(49)

$$\sigma^2_T = \frac{\left(\sqrt{N_{eq} - 1 + \sqrt{L}}\right)^2 \left(\frac{1}{\sqrt{N_{eq} - 1}} + \frac{1}{\sqrt{L}}\right)^{\frac{1}{2}} \sigma^2_{tw}}{(\sqrt{N_{eq} - 1} - \sqrt{L})^{\frac{1}{4}}}$$

(50)

Fig. 4. Histogram and pdf of statistic under null hypothesis when $N = 300$, SNR = -10 dB

By taking advantage of covariance matrix estimation, multiple (i.e., $N$) random variables $T$ are generated provided limited total samples (i.e., $N + L - 1$). Because of the cumulative average, the variance of the random variable can be further reduced by a factor of $N$ to reach (52). The histogram and the estimated probability distribution function (pdf) of the statistic under $H_0$ are shown in Fig.4. We can see the pdf approximates a Gaussian distribution.

The false alarm probability can be transformed into standard Gaussian distribution form as

$$P_{fa} = P\left(\frac{Q_N - \mu_{Q_N}}{\sigma_{Q_N}} > \frac{\gamma - \mu_{Q_N}}{\sigma_{Q_N}}\right)$$

(53)

where

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-\frac{x^2}{2}} dx$$

(54)

then the threshold $\gamma$ can be calculated as

$$\gamma = \mu_{Q_N} + Q^{-1}(P_{fa}) \sigma_{Q_N}$$

(55)

Here the threshold is not affected by noise power or SNR, which is stable against environment changes.
V. NUMERICAL RESULTS

In this section, we will give some simulation results using a digital TV (DTV) signal, which was captured (field measurements) in Washington D.C. [44]. The data rate of the vestigial sideband (VSB) DTV signal is 10.762 MSamples/sec. The recorded DTV signal was sampled at receiver at 21.524476 MSamples/sec and down converted to a low central IF frequency of 5.381119 MHz. The SNR of the received signal is unknown. In order to use the signals for simulating the algorithms at low SNR, we need to add white Gaussian noise to obtain various SNR levels [45]. The smoothing factor \( L \) is chosen to be 32. False alarm probability is fixed with \( P_{fa} = 1\% \), and all the thresholds are determined by this 1% false alarm probability. Two thousand simulations are performed on different sample sizes or different SNR levels.

A. Simulations on Proposed Algorithms

In Fig. 5, SNR is fixed at -5 dB while \( N \) varies. For 100% detection probability, Algorithm 1 only needs a sample size of 100 to achieve that. It is equivalent to 131 total data, corresponding to 6.086 micro seconds. While Algorithm 2 needs about a sample size of 500, that is 531 total data, corresponding to 24.670 micro seconds. The original sample covariance MME requires more than a sample size of 500 to achieve the same detection probability. The performance of detection with SNR -10 dB is shown in Fig. 6. With lower SNR, Algorithm 1 needs more data to achieve required detection probability. Observing from the figure, it’s about a sample size of 320 to reach 100% detection probability, which needs 16.307 micro seconds. The detection probabilities of both Algorithm 2 and sample covariance MME are increasing slowly as the \( N \) grows. They need far more data to obtain a satisfied detection performance when SNR is low, yet Algorithm 2 still performs better than sample covariance MME.

Given the number of total data size, the proposed algorithms are able to work in a relatively low SNR environment. In Fig. 7, when \( N \) equals 30, Algorithm 1 and Algorithm 2 can work at SNR 0 dB and 8 dB to achieve 100% detection probability, respectively. However, sample covariance MME is ineffective at any SNR level. When \( N \) increases to 100, shown in Fig. 8, all the detection performance are improved. Algorithm 1 can work at -4 dB, which is a 8 dB gain compared with sample covariance MME working at 4 dB. Fig. 7 and Fig. 8 suggested that providing more data will help detect the signal in a lower SNR environment. The SNR and the sample size \( N \) needed to achieve 100% detection probability exhibited a linear relationship between them, as shown in Table I. When the SNR (in dB) decreases by 3 dB, which equals SNR (not in dB) reduces by a factor of 2, the \( N \) will accordingly increase by a factor around 2.

One of the properties of our proposed algorithm is that the

| SNR (in dB) | 3 dB | 0 dB | -3 dB | -6 dB | -9 dB | -12 dB |
|-------------|------|------|-------|-------|-------|--------|
| SNR (not in dB) | 2 | 0.5 | 0.25 | 0.125 | 0.0625 |
| \( N \) | 15 | 32 | 61 | 123 | 252 | 590 |
| \( N \) increases by | / | 2.13 | 1.91 | 2.02 | 2.05 | 2.34 |
threshold is robust. As shown in Fig. 9, the threshold is almost a constant between 1.25 and 1.3 no matter what SNR is.

B. Comparison with Other Algorithms

In the following, we will discuss some other spectrum sensing algorithms for comparison purposes. Arithmetic-to-geometric mean (AGM) [14] which derives from GLRT, is able to sense the spectrum without prior knowledge. Feature template matching (FTM) [18] utilizes feature, which can be learned blindly, as prior knowledge for detection. Estimator-correlator (EC) [46] requires both the original PU’s signal covariance matrix and the noise variance. Energy detection (ED) is easy to be implemented, but usually suffers the noise uncertainty problem.

1) AGM: AGM finds an unstructured estimate of \( R_s \) to be \( R_s + \sigma^2_n I \) under \( H_1 \) and \( \sigma^2_n I \) under \( H_0 \). \( \lambda_m \) represents all eigenvalues of the sample covariance matrix and \( M \) is the dimension of the sample covariance matrix. Since the arithmetic mean is larger than geometric mean, the resulting detector computes the arithmetic-to-geometric mean of the eigenvalues of sample covariance matrix and compares with a threshold \[ T_{AGM} = \frac{1}{M} \sum_{m=1}^{M} \lambda_m > \gamma_{AGM} \] (56)

2) FTM: FTM extracts leading eigenvector as the feature, which is stable for signals and random for noise. FTM involves two steps. First, it learns the feature blindly by comparing the similarity of two consecutive sensing segments, namely feature learning algorithm (FLA) [18].

\[ T_{FLA} = \max_{l=1,2,\cdots,L-k+1} \left| \sum_{k=1}^{L} \eta_i[k] \eta_{i+1}[k+l] \right| > \gamma_{FTM} \] (57)

If \( T_{FLA} > \gamma_c \) feature is learned as \( \phi_{s,1} = \eta_{i+1} \). \( \gamma_c \) is the threshold determined empirically. Then, with the learned signal feature \( \phi_{s,1} \) as prior knowledge, this algorithm compares just the feature \( \phi_{x,1} \) from new sensing segment and signal feature \( \phi_{s,1} \) to determine if the signal exists.

\[ T_{FTM} = \max_{l=1,2,\cdots,L-k+1} \left| \sum_{k=1}^{L} \phi_{s,1}[k] \phi_{x,1}[k+l] \right| > \gamma_{FTM} \] (58)

3) EC: EC method assumes the signal follows zero mean Gaussian distribution with covariance matrix \( R_s \), and noise follows zero mean Gaussian distribution with covariance matrix \( \sigma^2_n I \),

\[ s \sim \mathcal{N}(0, R_s) \] (59)
\[ w \sim \mathcal{N}(0, \sigma^2_n I) \] (60)

Both \( R_s \) and \( \sigma^2_n \) are given priorly. The hypothesis \( H_1 \) is true if

\[ T_{EC} = \sum_{j=1}^{N} x_j^T R_s^{-1} (R_s + \sigma^2_n I)^{-1} x_j > \gamma_{EC} \] (61)

4) Comparison: The detection probabilities varied by SNR for Algorithm 1 compared with EC, FTM, AGM, MME and ED are shown in the following, where “ED x dB” means the energy detection with x dB noise uncertainty. In Fig. 10 when \( N = 100 \), if the noise variance is exactly known (ED 0 dB), the energy detection is pretty good. Speaking of more than 80% detection probability, Algorithm 1 exhibits almost the same performance with EC and ED 0.5 dB, even a little better. The rest of the algorithms including FTM, AGM, MME, ED 1 dB, all require higher SNR to achieve the same detection probability. In Fig. 11 when \( N = 300 \) and Fig. 12 when \( N = 500 \), Algorithm 1 is only worse than ED 0 dB when considering a more than 60% detection probability. In a word, the proposed algorithm demonstrates superior advantage, almost approximates the optimal EC and outperforms the performance of the rest of the algorithms mentioned above.

5) Discussion: The SNR change brought some impact to calculate \( T_{FLA} \) with each sample size \( \tau \), \( \tau = 1, \cdots, N \). Because of the cumulative addition, this impact from SNR was accumulated and amplified with N times in forming \( Q_N \) in alternative hypothesis, while \( Q_N \) in null hypothesis remained unchanged. That is the reason why our algorithm is sensitive to the SNR, as shown in Fig. 10 that the curve of Algorithm...
Fig. 10. Detection probabilities comparison among different algorithms at N = 100 with SNR varies

Fig. 11. Detection probabilities comparison among different algorithms at N = 300 with SNR varies

Fig. 12. Detection probabilities comparison among different algorithms at N = 500 with SNR varies

N increases, as shown in Fig. 11 and Fig. 12.

VI. CONCLUSION

We have considered the spectrum sensing for single PU with single antenna. A blind cumulative spectrum sensing method has been proposed focusing on small-sized datasets. If the total sample size is given, this method works in a lower SNR environment compared with some other algorithms. Concentration inequalities of statistics have been adopted to demonstrate the effectiveness and validity of the proposed method. Meanwhile, the threshold has been proven to be stable. All the results were verified by the simulations using a captured DTV signal.

This proposed method can also be extended to be a general detection framework. The MME detector in this method can be replaced by other covariance matrix based spectrum sensing algorithms [47] depend on different detection scenarios, the detection performance may be further improved consequently. In addition, this method can be applied to Smart Grid as well because real-time response of system changes is also a fundamental requirement in Smart Grid [48].

ACKNOWLEDGMENT

The authors would like to thank Dr. Zhen Hu, for the helpful discussions on this paper.

REFERENCES

[1] FCC, “Spectrum policy task force report,” ET Docket No. 02-155, Tech. Rep., Nov. 2002.
[2] H. Kim, C. Cordeiro, K. Challapali, and K. Shin, “An experimental approach to spectrum sensing in cognitive radio networks with off-the-shelf ieee 802.11 devices,” in Proc. Fourth IEEE Consumer Comm. and Networking Conf. Workshop Cognitive Radio Networks, 2007, pp. 1154–1158.
[3] A. Sonnenchein and P. Fishman, “Radiometric detection of spread-spectrum signals in noise of uncertain power,” Aerospace and Electronic Systems, IEEE Transactions on, vol. 28, no. 3, pp. 654 –660, jul 1992.
[4] V. Kostylev, “Energy detection of a signal with random amplitude,” in Communications, 2002. ICC 2002. IEEE International Conference on, vol. 3, 2002, pp. 1606 –1610.
[5] F. Digham, M.-S. Alouini, and M. Simon, “On the energy detection of unknown signals over fading channels,” in Communications, 2003. ICC ’03. IEEE International Conference on, vol. 5, may 2003, pp. 3575 –3579.
[6] S. Haykin, D. Thomson, and J. Reed, “Spectrum sensing for cognitive radio,” Proceedings of the IEEE, vol. 97, no. 5, pp. 849 –877, may 2009.
[7] J. Ma, G. Li, and B. H. Juang, “Signal processing in cognitive radio,” Proceedings of the IEEE, vol. 97, no. 5, pp. 805 –823, may 2009.
[8] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, “A review on spectrum sensing for cognitive radio: challenges and solutions,” EURASIP J. Adv. Signal Process, vol. 2010, pp. 2:2–2:2, January 2010.
[9] A. Dandawate and G. Giannakis, “Statistical tests for presence of cyclostationarity,” IEEE Transactions on Signal Processing, vol. 42, no. 9, pp. 2355–2369, 1994.
[10] P. Sutton, K. Nolan, and L. Doyle, “Cyclostationary signatures in practical cognitive radio applications,” Selected Areas in Communications, IEEE Journal on, vol. 26, no. 1, pp. 13 –24, jan. 2008.
[11] W. Gardner, “Signal interception: a unifying theoretical framework for feature detection,” IEEE Transactions on Communications, vol. 36, no. 8, pp. 897 –906, aug 1988.
[12] Y. Zeng and Y.-C. Liang, “Maximum-Minimum Eigenvalue Detection for Cognitive Radio,” 2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, Sep. 2007.
[13] Y. Zeng and Y.-C. Liang, “Eigenvalue-based spectrum sensing algorithms for cognitive radio,” IEEE Transactions on Communications, vol. 57, no. 6, pp. 1784 –1793, June 2009.
[14] T. J. Lim, R. Zhang, Y. Liang, and Y. Zeng, “Glt-based spectrum sensing for cognitive radio,” in Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008., Nov. 2008., pp. 1–5.

[15] Y. Zeng and Y.-C. Liang, “Covariance based signal detections for cognitive radio,” in New Frontiers in Dynamic Spectrum Access Networks, 2007, DySPAN 2007. 2nd IEEE International Symposium on, April 2007., pp. 202–207.

[16] P. Wang, J. Fang, N. Han, and H. Li, “Multiantenna-Assisted Spectrum Sensing for Cognitive Radio,” IEEE Transactions on Vehicular Technology, vol. 59, no. 4, pp. 1791–1800, May 2010.

[17] R. López-Valcarce, G. Vazquez-Vilar, and J. Sala, “Multiantenna spectrum sensing for cognitive radio: overcoming noise uncertainty,” in Cognitive Information Processing (CIP), 2010 2nd International Workshop on, IEEE, 2010., pp. 310–315.

[18] P. Zhang, R. Qiu, and N. Guo, “Demonstration of Spectrum Sensing with Blindly Learned Feature,” IEEE Communications Letters, vol. 15, no. 5, pp. 548–550, May 2011.

[19] P. Zhang and R. Qiu, “Glt-based spectrum sensing with blindly learned feature under rank-1 assumption,” Communications, IEEE Transactions on, vol. 61, no. 1, pp. 87–96, 2013.

[20] S. Hou and R. C. Qiu, “Kernel Feature Template Matching for Spectrumsensing,” to appear in IEEE Transactions on Vehicular Technology, 2013. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=2290013

[21] G. Ding, Q. Wu, Y.-D. Yao, J. Wang, and Y. Chen, “Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions,” IEEE Signal Processing Magazine, vol. 30, pp. 126–136, 2013.

[22] F. Lin, R. C. Qiu, Z. Hu, H. Li, P. Browning, and M. C. Wicks, “Generalized FMD Detection for Spectrum Sensing Under Low Signal-to-Noise Ratio,” IEEE Communications Letters, vol. 16, no. 5, pp. 604–607, May 2012.

[23] F. Lin, R. C. Qiu, Z. Hu, S. Hou, L. Li, J. P. Browning, and M. C. Wicks, “Cognitive Radio Network as Sensors: Low Signal-to-Noise Ratio Collaborative Spectrum Sensing,” in Proceeding of IEEE International Waveform Diversity & Design Conference, January 2012.

[24] R. Qiu and M. Wicks, Cognitive Networked Sensing and Big Data. Springer, 2013.

[25] Q. Wu, G. Ding, J. Wang, and Y.-D. Yao, “Spatial-temporal opportunity detection for spectrum-heterogeneous cognitive radio networks: Two-dimensional sensing,” Wireless Communications, IEEE Transactions on, vol. 12, no. 2, pp. 516–526, February 2013.

[26] Y. Zou, Y.-D. Yao, and B. Zheng, “Cooperative relay techniques for cognitive radio systems: spectrum sensing and secondary user transmissions,” Communications Magazine, IEEE, vol. 50, no. 4, pp. 98–103, 2012.

[27] H. V. Poor and O. Hadjiilikadi, Quickest Detection. Cambridge University Press, 2009.

[28] H. Li, C. Li, and H. Dai, “Quickest spectrum sensing in cognitive radio,” in Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on., IEEE, 2008., pp. 203–208.

[29] H. Li, H. Dai, and C. Li, “Collaborative quickest spectrum sensing via random broadcast in cognitive radio systems,” Wireless Communications, IEEE Transactions on, vol. 9, no. 7, pp. 2338–2348, 2010.

[30] L. Lai, Y. Fan, and H. Poor, “Quickest detection in cognitive radio: A sequential change detection framework,” in Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE., IEEE, 2008., pp. 1–5.

[31] Z. Chen, Z. Hu, and R. Qiu, “Quickest spectrum detection using hidden markov model for cognitive radio,” in Military Communications Conference, 2009. MILCOM 2009. IEEE., IEEE, 2009., pp. 1–7.

[32] S. Zarrin and T. Lim, “Cooperative quickest spectrum sensing in cognitive radios with unknown parameters,” in Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE., IEEE, 2009., pp. 1–6.

[33] Y. Youn, H. Jeon, J. H. Choi, and H. Lee, “Fast spectrum sensing algorithm for 802.22 wran systems,” in Communications and Information Technologies, 2006. ISCI’06. International Symposium on., IEEE, 2006., pp. 960–964.

[34] R. C. Qiu, C. Zhang, Z. Hu, and M. C. Wicks, “Towards A Large-Scale Cognitive Radio Network Testbed: Spectrum Sensing, System Architecture, and Distributed Sensing,” 2012, accepted by Journal of Communications.

[35] Y. Chen, A. Wiesel, and A. O. Hero, “Robust Shrinkage Estimation of High-dimensional Covariance Matrices,” September 2010. [Online]. Available: http://arxiv.org/pdf/1009.5331.pdf