Theoretical Performance Analysis of 
Eigenvalue-based Detection

Federico Penna, Student Member, IEEE, and Roberto Garello, Senior Member, IEEE

Abstract

In this paper we develop a complete analytical framework based on Random Matrix Theory for the performance evaluation of Eigenvalue-based Detection. While, up to now, analysis was limited to false-alarm probability, we have obtained an analytical expression also for the probability of missed detection, by using the theory of spiked population models. A general scenario with multiple signals present at the same time is considered. The theoretical results of this paper allow to predict the error probabilities, and to set the decision threshold accordingly, by means of a few mathematical formulae. In this way the design of an eigenvalue-based detector is made conceptually identical to that of a traditional energy detector. As additional results, the paper discusses the conditions of signal identifiability for single and multiple sources. All the analytical results are validated through numerical simulations, covering also convergence, identifiability and non-Gaussian practical modulations.

Index Terms

Cognitive Radio, Spectrum Sensing, Random Matrix Theory, Spiked Population Models.

I. INTRODUCTION

Eigenvalue-based Detection (EBD) has been introduced [1], [2] as an efficient technique to perform spectrum sensing in Cognitive Radio (CR). Using the EDB approach, the secondary receiver is able to infer the presence or the absence of a primary signal based on the largest and the smallest eigenvalue of the received signal’s covariance matrix. This technique requires a cooperative detection setting, which may be accomplished by multiple antennas or cooperation among different users. In addition to the CR

F. Penna is with TRM Lab - Istituto Superiore Mario Boella (ISMB) and with the Department of Electrical Engineering (DELEN), Politecnico di Torino, Italy. R. Garello is with the Department of Electrical Engineering (DELEN), Politecnico di Torino, Italy. e-mail: {federico.penna, roberto.garello}@polito.it.
context, the detection of signal components in noisy covariance matrices is a very general problem, with a wide variety of applications in communications, statistics, genetics, mathematical finance, artificial learning.

The main advantage offered by EDB is its robustness to the problem of noise uncertainty, which affects all the previously proposed detection schemes including the widely adopted Energy Detection (ED). However, while for ED there exist comprehensive theoretical results that allow to express the error probabilities through analytical formulae, a corresponding theory for EBD has not been fully developed yet.

In general, a signal detection scheme can be characterized by defining two types of error probabilities: the probability of false alarm and the probability of missed detection (see Sec. II-A for a formal definition). These probabilities depend on the decision threshold (the value used by the algorithm to decide whether a signal is present or absent). If analytical formulae are available, it is possible to:

a) predict the error probabilities of the system as a function of the decision threshold;

b) set the decision threshold according to the required error constraints.

Such formulae are well-known in case of ED. For EBD, up to now, only approximated criteria were proposed for the estimation of the false-alarm probability [1], [2] and, to the best of our knowledge, no exact analytical results have been found for the missed-detection probability yet.

In this paper, by exploiting the spectral properties of the sample covariance matrix under the two complementary conditions of signal present/absent, we derive analytical expressions both for the false-alarm and the missed-detection probability. The result is a complete probabilistic framework that allows to evaluate the performance of EBD and to determine the proper decision threshold through analytical formulae.

Whereas most of the works on detection consider only the case of a single signal to be detected, our results also apply to the case of multiple primary signals. This generalization is of interest for the applications in CR, since a secondary user might be located in such a way as to hear different primary signals (each with a different channel). The analytical results derived in this paper show that the number of signals simultaneously present, as well as their powers and their channels, have an impact on the detection performance.

The paper is organized as follows: Sec. II introduces the signal model and the theoretical foundations of eigenvalue-based detection; Sec. III and IV derive analytical results for the probabilities of false alarm and missed detection, and for the signal identifiability condition; Sec. V discusses the problem of setting a proper decision threshold; Sec. VI validates the analysis through numerical results; Sec. VII concludes.
II. EIGENVALUE-BASED DETECTION

**Notational remark:** In the following, upper-case boldface letters indicate matrices, lower-case boldface letters indicate vectors, the symbols $^T$ and $^H$ indicate respectively the transpose and conjugate transpose (Hermitian) operators, $\text{tr}(\cdot)$ is the trace of a matrix, $\| \cdot \|$ is the Euclidean norm of a vector, $\text{diag}(x)$ indicates a square diagonal matrix whose main diagonal entries are taken from the vector $x$, $I_N$ is the identity matrix (of size $N$ if specified), $0_{M,N}$ is a $M \times N$ matrix of zeros; the symbol $\triangleq$ stands for “defined as”, the symbol $\sim$ for “distributed with law”, $\xrightarrow{\text{a.s.}}$ indicates the almost sure convergence, and $\xrightarrow{\text{D}}$ the convergence in distribution; $I_{\{\alpha\}}$ is the indicator function which takes value 1 where the condition $\alpha$ is true and 0 elsewhere.

**A. Signal model**

We consider a cooperative detection framework in which $K$ receivers (or antennas) collaborate to sense the spectrum. Denote with $y_k$ be the discrete baseband complex sample at receiver $k$, and define the $K \times 1$ vector $y = [y_1 \ldots y_K]^T$ containing the $K$ received signal samples.

The goal of the detector is to discriminate between two hypotheses:

- $\mathcal{H}_0$ (absence of primary signal). The samples contain only noise:
  \[
  y|_{\mathcal{H}_0} = v
  \]
  where $v \sim \mathcal{N}_C(0_{K,1}, \sigma_v^2 I_K)$ is a vector of circularly symmetric complex Gaussian (CSCG) noise samples;

- $\mathcal{H}_1$ (presence of primary signal). For sake of generality, we consider a model where $P$ primary signals may be simultaneously present:
  \[
  y|_{\mathcal{H}_1} = H s + v
  \]
  where $H$ is a $K \times P$ complex matrix, where each element $h_{kp}$ represents the channel between primary user $p$ and receiver $k$ (for simplicity, channels are assumed to be memoryless and constant for the sensing duration); $s$ is a $P \times 1$ vector containing the primary signal samples, each coming from one of the $P$ sources. The primary signals are assumed to be complex, zero-mean and mutually independent with covariance matrix
  \[
  \mathbb{E} ss^H \triangleq \Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_P^2)
  \]
  where $\sigma_p^2$ is the variance of the $p$-th primary signal.

\[\text{September 23, 2009 DRAFT}\]
Under $H_1$, we define the signal-to-noise ratio (SNR) as

$$\rho \equiv \frac{\mathbb{E} \|Hs\|^2}{\mathbb{E} \|v\|^2}$$  \hspace{1cm} (4)$$

This amounts to

$$\rho = \frac{\text{tr} \ H \Sigma H^H}{K\sigma_v^2} = \frac{\sum_{p=1}^{P} \sigma_p^2 \|h_p\|^2}{K\sigma_v^2}$$  \hspace{1cm} (5)$$

where $h_p$ is the $p$-th column of the matrix $H$, i.e., the channel vector referred to primary source $p$.

In the single-user case $(P = 1)$, we can drop the index $p$ and the expression of the SNR simplifies to

$$\rho|_{P=1} = \frac{\sigma_1^2 \|h\|^2}{K\sigma_v^2}$$  \hspace{1cm} (6)$$

Remark: All throughout this paper it is assumed that $P < K$. When this assumption is not verified, the covariance matrix lacks the necessary degrees of freedom to be able to distinguish the signal components from the noise. Notice that $P$ might be unknown, but to ensure a reliable detection $K$ (which is a receiver parameter) has to be chosen greater than the maximum possible number of primary signals.

B. Spectral properties of the statistical covariance matrix

Define the statistical covariance matrix of the received signal

$$R \equiv \mathbb{E} yy^H$$  \hspace{1cm} (7)$$

Under $H_0$ and $H_1$ it is equal to, respectively

$$R = \begin{cases} 
\sigma_v^2 I_K & (H_0) \\
H \Sigma H^H + \sigma_v^2 I_K & (H_1) 
\end{cases}$$  \hspace{1cm} (8)$$

Let $\lambda_1 \geq \ldots \geq \lambda_K$ be the eigenvalues of $R$ (without loss of generality, sorted in decreasing order).

Under $H_0$, it is immediate to verify that

$$\lambda_i|_{H_0} = \sigma_v^2 \quad \forall i = 1, \ldots, K$$  \hspace{1cm} (9)$$

Under $H_1$, there are $(K - P)$ eigenvalues equal to $\sigma_v^2$ and $P$ greater, since $H \Sigma H^H$ is positive-semidefinite with rank $P$. The eigenvalues in this case can be written as

$$\lambda_i|_{H_1} = \begin{cases} 
s_i + \sigma_v^2 & (1 \leq i \leq P) \\
\sigma_v^2 & (P < i \leq K) 
\end{cases}$$  \hspace{1cm} (10)$$

where $s_1 \geq \ldots \geq s_P > 0$ denote the $P$ non-zero eigenvalues of the “signal covariance matrix” $H \Sigma H^H$, and are found by solving the characteristic equation

$$\det \left( H \Sigma H^H - sI_K \right) = 0$$  \hspace{1cm} \text{s.t. } s \neq 0 \hspace{1cm} (11)$$
Because of the assumption $P < K$, the rank of the signal covariance matrix is $P$. It is possible to reduce the degree of the characteristic polynomial down to $P$ by applying the generalized Matrix Determinant Lemma (MDL) [19] 

$$\det \left( \mathbf{H} \Sigma \mathbf{H}^H - s \mathbf{I}_K \right) =$$

$$= \det(\Sigma) \det(-s \mathbf{I}_K) \det \left( \mathbf{\Sigma}^{-1} - \frac{1}{s} \mathbf{H}^H \mathbf{H} \right) =$$

$$= \left( \prod_{p=1}^{P} \sigma_p^2 \right) (-s)^{K-P} \det \left( \mathbf{H}^H \mathbf{H} - s \mathbf{\Sigma}^{-1} \right) \quad (12)$$

We note that the left-hand factor in (12) is a constant with respect to $s$, the middle term gives rise to the $(K - P)$ trivial solutions $s = 0$, while the right-hand term determines the non-zero roots. The signal eigenvalues $s_1, \ldots, s_P$ may therefore be calculated from the simplified characteristic equation

$$\det \left( \mathbf{H}^H \mathbf{H} - s \mathbf{\Sigma}^{-1} \right) = 0 \quad (13)$$

which has degree $P$ instead of $K$. Since $\Sigma$ is diagonal, $\mathbf{\Sigma}^{-1} = \text{diag}(\sigma_1^{-2}, \ldots, \sigma_P^{-2})$.

In the case of single primary user ($P = 1$), there is one single signal eigenvalue and, from (13), it has a very simple expression:

$$s_1|_{P=1} = \|\mathbf{h}\|^2 \sigma^2 \quad (14)$$

where the index has been dropped like in (6).

The spectral properties of $\mathbf{R}$, summarized by (9) and (10), motivate the adoption of the ratio between the largest and the smallest eigenvalue of the covariance matrix as a test statistic to discriminate between the two hypotheses: under $\mathcal{H}_0$ the ratio is equal to 1, under $\mathcal{H}_1$ it is greater. This detection scheme was first proposed in [1], [2].

C. Sample covariance matrix

In practice, the statistical correlation matrix $\mathbf{R}$ is estimated through a sample covariance matrix. Introduce $N$ as the number of samples collected by each receiver during the sensing period. It is assumed that consecutive samples are uncorrelated and that all the random processes involved (signals and noise) remain stationary for the sensing duration. Then, let $s(n)$, $\mathbf{v}(n)$ and $\mathbf{y}(n)$ be, respectively, the transmitted signal vector, the noise vector and the received signal vector at time $n$; define the $P \times N$ matrix

$$\mathbf{S} \triangleq [s(1) \ldots s(N)] \quad (15)$$
and the $K \times N$ matrices
\[ V \triangleq [v(1) \ldots v(N)] \] (16)
\[ Y \triangleq [y(1) \ldots y(N)] = HS + V \] (17)

The $K \times K$ sample covariance matrix $R(N)$ is then defined as
\[ R(N) \triangleq \frac{1}{N} YY^H \] (18)

Denoting with $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_K$ its eigenvalues, the test statistic used for detection is
\[ T \triangleq \frac{\hat{\lambda}_1}{\hat{\lambda}_K} \] (19)

Although $R(N)$ converges to $R$ as $N$ tends to infinity, for finite $N$ its properties depart from those
of the statistical covariance matrix. In typical sensing applications $N$ is expected to be quite large (to
increase the detection reliability) but still not enormous (to reduce the sensing time). With such realistic
values of $N$, the eigenvalues have no longer a deterministic behavior as in (8), but are characterized
by a probability distribution. Therefore the discrimination criterion based on the eigenvalues is not as
sharp-cutting as in the ideal case and may be affected by two possible error events: false alarms and
missed detections. Denoting with $\gamma$ the decision threshold employed by the detector, such that
\[
\text{decision} = \begin{cases} H_0 & \text{if } T < \gamma \\ H_1 & \text{if } T \geq \gamma \end{cases},
\]
the probability of false alarm may be expressed as
\[ P_{fa} = \Pr(T \geq \gamma | H_0) \] (20)
and the probability of missed detection as
\[ P_{md} = \Pr(T < \gamma | H_1) \] (21)

These probabilities depend on the distribution of $T$ under the two hypotheses. The probability distribution
function (PDF) and the cumulative distribution function (CDF) of $T$ will be indicated as $f_{T|H_i}(t)$ and
$F_{T|H_i}(t)$, respectively, for $i \in \{0, 1\}$. Thus, (20) and (21) may be written as
\[ P_{fa} = 1 - F_{T|H_0}(\gamma) \] (22)
\[ P_{md} = F_{T|H_1}(\gamma) \] (23)

In the next sections the distribution of $T$ in both cases will be derived, using tools from Random Matrix
Theory (RMT) which allow to analyze the spectral properties of large-dimensional sample covariance
matrices. This makes it possible to evaluate the detection performance, given a decision threshold, as well as to express the threshold as a function of the required probabilities of false alarm or missed detection (by inverting (22) and (23)).

III. FALSE-ALARM PROBABILITY ANALYSIS

In this section, we first introduce some useful results from RMT that express the limiting distributions to which the largest and the smallest eigenvalues of $R(N)$ converge as $N$ and $K$ grow. Then, we exploit these theoretical results to find the limiting distribution of the test statistic $T$ and, through the relation (22), we derive the false-alarm probability.

Most of the results of this section also appear, in a slightly different form, in [21]. Here the results are stated in their entirety and are introduced by a more rigorous mathematical derivation. Also, a new notation is adopted to emphasize the link between the Wishart case ($\mathcal{H}_0$) and the spiked-population case ($\mathcal{H}_1$, discussed in Sec. IV).

A. Relevant results from Random Matrix Theory

Under $\mathcal{H}_0$, since the columns of $Y$ are zero-mean independent complex Gaussian vectors, the sample covariance matrix $R(N)$ is a complex Wishart matrix [4].

The fluctuations of the eigenvalues of Wishart matrices have been thoroughly investigated by RMT (see [3] and [6] for an overview). The most remarkable intuition of RMT is that in many cases the eigenvalues of matrices with random entries turn out to converge to some fixed distribution, when both the dimensions of the signal matrix tend to infinity with the same order. For Wishart matrices the limiting joint eigenvalue distribution has been known for many years [5]; then, more recently, also the marginal distributions of single ordered eigenvalues have been found.

By exploiting some of these results, we are able to express the asymptotical values of the largest and the smallest eigenvalue of $R(N)$ as well as their limiting distributions. We state the following theorem, which summarizes a number of relevant results.

Theorem 3.1: Convergence of the smallest and largest eigenvalues under $\mathcal{H}_0$. Let

$$c \triangleq \frac{K}{N}$$  \hspace{1cm} (24)

and assume that for $K, N \to \infty$

$$c \to \bar{c} \in (0, 1)$$  \hspace{1cm} (25)
Define:
\[
\mu_\pm(c) \triangleq \left(\frac{c^{1/2} \pm 1}{2}\right)^2 \\
\nu_\pm(c) \triangleq \left(\frac{c^{1/2} \pm 1}{2}\right) \left(\frac{c^{-1/2} \pm 1}{2}\right)^{1/3}
\]
(26, 27)

Then, as \(N, K \to \infty\), the following holds:

(i) **Almost sure convergence of the largest eigenvalue**
\[
\hat{\lambda}_1 \overset{a.s.}{\to} \sigma_v^2 \mu_+(c) 
\]
(28)

(ii) **Convergence in distribution of the largest eigenvalue**
\[
N^{2/3} \frac{\hat{\lambda}_1 - \sigma_v^2 \mu_+(c)}{\sigma_v^2 \nu_+(c)} \overset{D}{\to} W_2 
\]
(29)

(iii) **Almost sure convergence of the smallest eigenvalue**
\[
\hat{\lambda}_K \overset{a.s.}{\to} \sigma_v^2 \mu_-(c) 
\]
(30)

(iv) **Convergence in distribution of the smallest eigenvalue**
\[
N^{2/3} \frac{\hat{\lambda}_K - \sigma_v^2 \mu_-(c)}{\sigma_v^2 \nu_-(c)} \overset{D}{\to} W_2 
\]
(31)

where \(W_2\) is the Tracy-Widom law of order 2, defined in Appendix A.

**Proof:** The claims of this theorem follow from different results of RMT, up to some changes of variables and using a uniform notation. Proofs of the original theorems appear in the references listed below.

Claims (i) and (iii) descend from the work by Marchenko and Pastur [5], later extended by Silverstein, Bai, Yin, et al. [6].

Claim (ii) was proved, under the assumption of Gaussian entries, by Johansson [7], Johnstone [8] and Soshnikov [9], and generalized to the non-Gaussian case by Péché [10].

Claim (iv) derives from a very recent result by Feldheim and Sodin [11].

B. Derivation of \(F_{T|\mathcal{H}_0}\) and \(P_{fa}\)

The results of Theorem 3.1 allow, through some algebraic manipulations, to determine the limiting distribution of the test statistic \(T\) under the hypothesis \(\mathcal{H}_0\). Although the resulting distribution is obtained under the joint limit \(K, N \to \infty\), simulations show that it provides an accurate estimation of the false-alarm probability already for not-so-large values of \(K\) and \(N\). Numerical results investigating this issue are presented in Sec. VI.
In order to apply claims (ii) and (iv), we define:

\[ L_1 \triangleq N^{2/3} \frac{\hat{\lambda}_1 - \sigma^2 \mu_+ (c)}{\sigma^2 \nu_+ (c)} \]  

\[ L_K \triangleq N^{2/3} \frac{\hat{\lambda}_K - \sigma^2 \mu_- (c)}{\sigma^2 \nu_- (c)} \]  

For the above-mentioned theorem, both \( L_1 \) and \( L_K \) converge in distribution to the Tracy-Widom law \( \mathcal{W}_2 \):

\[ f_{L_1} (z), f_{L_K} (z) \to f_{\mathcal{W}_2} (z) \]  

where \( f_{\mathcal{W}_2} (\cdot) \) represents the PDF associated with the law \( \mathcal{W}_2 \), as defined in Appendix A.

Then, from (19), the test statistic \( T \) becomes

\[ T = \frac{\hat{\lambda}_1}{\hat{\lambda}_K} = \frac{N^{-2/3} \nu_+ (c) L_1 + \mu_+ (c)}{N^{-2/3} \nu_- (c) L_K + \mu_- (c)} \]  

Notice that the term \( \sigma^2 \) is canceled out in the ratio (this is the reason that makes the detection threshold “blind” with respect to the noise power). We denote with \( l_1 \) and \( l_K \), respectively, the numerator and the denominator of \( T \), and with \( \mathcal{F}_{l_1} (z) \) and \( \mathcal{F}_{l_K} (z) \) their limiting PDFs for \( N, K \to \infty \). These distributions are the same as those of \( L_1 \) and \( L_K \), up to a linear random variable transformation:

\[ \mathcal{F}_{l_1} (z) = \frac{N^{2/3}}{\nu_+ (c)} f_{\mathcal{W}_2} \left( \frac{N^{2/3}}{\nu_+ (c)} (z - \mu_+ (c)) \right) \]  

For the denominator, it must be observed that \( \nu_- (c) < 0 \) for the considered range \( c \in (0, 1) \). Thus

\[ \mathcal{F}_{l_K} (z) = \frac{N^{2/3}}{|\nu_- (c)|} f_{\mathcal{W}_2} \left( \frac{N^{2/3}}{|\nu_- (c)|} (\mu_- (c) - z) \right) = - \frac{N^{2/3}}{\nu_- (c)} f_{\mathcal{W}_2} \left( \frac{N^{2/3}}{\nu_- (c)} (z - \mu_- (c)) \right) \]  

To express the distribution of \( T \), we assume that \( f_{l_1} (l_1) \) and \( f_{l_K} (l_K) \) are asymptotically independent, as it is reasonable for the size of the covariance matrix tending to infinity (and confirmed by following numerical results):

\[ \mathcal{F}_{l_1, l_K} (l_1, l_K) \approx \mathcal{F}_{l_1} (l_1) \mathcal{F}_{l_K} (l_K) \]  

Then, using the formula for the quotient of random variables [20], the resulting ratio distribution writes:

\[ \mathcal{F}_{T | H_0} (t) = \left[ \int_{-\infty}^{+\infty} |x| \mathcal{F}_{l_1, l_K} (tx, x) dx \right] \cdot I_{\{t>1\}} = \left[ \int_{0}^{+\infty} x \mathcal{F}_{l_1} (tx) \mathcal{F}_{l_K} (x) dx \right] \cdot I_{\{t>1\}} \]  

September 23, 2009 DRAFT
where the lower integration limit has been changed to 0 instead of \(-\infty\), since the covariance matrix is positive-semidefinite therefore all the eigenvalues are non-negative; the condition \(t > 1\) is necessary to preserve the order of the eigenvalues, since the distributions are defined under the assumption \(l_1 > l_K\).

Finally, we denote with \(F_{T|\mathcal{H}_0}(\gamma)\) the CDF corresponding to (39). For \(N\) and \(K\) large enough, we can approximate \(F_{T|\mathcal{H}_0}(\gamma)\), which is needed to compute \(P_{fa}\) from (22), with the asymptotical distribution:

\[
F_{T|\mathcal{H}_0}(\gamma) \approx F_{T|\mathcal{H}_0}(\gamma)
\]

The expression of \(F_{T|\mathcal{H}_0}\) depends on \(N\) and \(c\), i.e., \(N\) and \(K\). Simulation results show that the approximation is accurate for practical values of \(N\) and \(K\), also quite far from the asymptotical region.

Clearly, the practical interest in the relation between \(P_{fa}\) and \(\gamma\) found here is that it allows to determine the decision threshold as a function of the required false-alarm probability; this application is discussed in more detail in Sec. V.

It is interesting to note that the distribution \(F_{T|\mathcal{H}_0}\) for finite \(N\) and \(K\) can also be expressed exactly, by following a completely different approach. This exact distribution and the corresponding detection threshold have been found in [22]. The drawback of the “exact” approach is its complexity, which makes implementation difficult when \(K\) and \(N\) are large.

IV. MISSED-DETECTION PROBABILITY ANALYSIS

In this section we use an approach based on RMT to derive the limiting distribution of \(T\) under \(\mathcal{H}_1\) and consequently \(P_{md}\). As a preliminary step, we show that under this hypothesis \(R(N)\) can be reduced to a so-called spiked population model, i.e., a model where the statistical covariance matrix is a finite-rank perturbation of the identity. Spiked population models were introduced by Johnstone [8] and have an important role in Principal Component Analysis (PCA), with many statistical applications ranging from genetics to mathematical finance. The fluctuations of the eigenvalues of sample covariance matrices constructed from spiked models are nowadays a hot topic in RMT.

A. Reduction to the Spiked Population Model

Under \(\mathcal{H}_1\), the received signal matrix \(Y\) contains some Gaussian entries, like in the Wishart case, along with a certain number \((P)\) of signal components. In order to put into evidence the spiked structure of \(R(N)\), the received signal matrix \(Y\) (16) needs to be rewritten in the form

\[
Y = TZ
\]
where $T$ is a block matrix of size $K \times (P + K)$ defined as

$$T = \left[ \begin{array}{c|c} \frac{1}{\sigma_v} H \Sigma^{1/2} & I_K \end{array} \right]$$

(42)

and $Z$, of size $(P + K) \times N$, is defined as

$$Z = \left[ \begin{array}{c|c} \sigma_v \Sigma^{-1/2} S \end{array} \right]$$

(43)

This decomposition has been chosen such that all the entries $z_{ij}$ of $Z$ ($1 \leq i \leq P + K, \ 1 \leq j \leq N$) have the following properties:

$$E z_{ij} = 0$$

(44)

$$E |z_{ij}|^2 = \sigma_v^2$$

(45)

which are necessary conditions for Theorem 4.1 to hold. The covariance matrix becomes then

$$R(N) = \frac{1}{N} T Z Z^H T^H$$

(46)

which is exactly the model of [12], [13] and [14].

Finally, we denote with $t_1, \ldots, t_K$ the eigenvalues of $TT^H$. It follows from the structure of $T$ that $P$ eigenvalues are different from 1 (without loss of generality we put them in the first $P$ positions: $t_1 \geq \ldots \geq t_P$) and the remaining $K - P$ are ones. To express the $P$ “spike eigenvalues” (that represent the perturbation with respect to the pure-noise model), we notice that

$$TT^H = \frac{1}{\sigma_v^2} H \Sigma H^H + I_K$$

(47)

and the eigenvalues $t_1, \ldots, t_P$ result from the solution of

$$\det \left( H \Sigma H^H - \sigma_v^2 (t - 1) I_K \right) = 0$$

s.t. $t \neq 1$

(48)

The structure of the problem is identical to that of (11), with the change of variable $s = \sigma_v^2 (t - 1)$. We can therefore conclude that the “spike eigenvalues” $t_p$ are linked to the non-zero eigenvalues of the statistical covariance matrix, $s_p$, by the relation

$$t_p = \frac{s_p}{\sigma_v^2} + 1, \ 1 \leq p \leq P$$

(49)

In general, the values of $s_p$ are calculated using (12); in the case of single primary user ($P = 1$), there is the simplified expression (14) which leads to

$$t_1 \big|_{P=1} = \|h\|^2 \sigma_v^2 + 1$$

(50)
Relation between spike eigenvalues and SNR: The spike eigenvalues are related with the SNR; this fact turns out to be useful especially in the case of \( P = 1 \). From (49) we can write

\[
\sum_{p=1}^{P} t_p = \frac{1}{\sigma_v^2} \sum_{p=1}^{P} s_p + P
\]  

(51)

but, from the eigendecomposition of \( H \Sigma H^H \) and from (5) it follows that

\[
\sum_{p=1}^{P} s_p = \text{tr} \ H \Sigma H^H = \rho K \sigma_v^2
\]  

(52)

hence

\[
\sum_{p=1}^{P} t_p = K \rho + P
\]  

(53)

Therefore, in the case of one primary user (\( P = 1 \)), the (unique) spike eigenvalue may be expressed directly as a function of the SNR:

\[
t_1|_{P=1} = K \rho + 1
\]  

(54)

Note that, by exploiting the property (52), one could also obtain (14) without resorting to the characteristic equation.

In the case of multiple primary signals (\( P > 1 \)), the sum of the spike eigenvalues is related to the SNR, but not the single eigenvalues. Therefore, to compute the \( t_p \) (in particular \( t_1 \), which is needed to apply Theorem 4.1), it is necessary to know the channel matrix and the power of primary signals and use (13).

B. Relevant results from Random Matrix Theory

We are now ready to state the following theorem which provides a useful result on the convergence of the largest eigenvalue in spiked population models.

**Theorem 4.1: Convergence of the largest eigenvalue under \( H_1 \).** Again, assume that for \( K, N \rightarrow \infty \)

\[
c = \frac{K}{N} \rightarrow \tau \in (0, 1)
\]  

(55)

In addition, assume that for all \( i, j \) s.t. \( 1 \leq i \leq P + K, \ 1 \leq j \leq N \):

\[
(A_1) \ E \ z_{ij} = 0
\]

\[
(A_2) \ E \ (\Re z_{ij})^2 = E \ (\Im z_{ij})^2 = \frac{\sigma_v^2}{2}
\]

\[
(A_3) \ \forall k > 0, \ E \ |z_{ij}|^{2k} < \infty \text{ and } E \ (\Re z_{ij})^{2k+1} = E \ (\Im z_{ij})^{2k+1} = 0
\]

\[
(A_4) \ E \ (\Re z_{ij})^4 = E \ (\Im z_{ij})^4 = \frac{3}{4} \sigma_v^4
\]
Define:

\[ \mu_s(t_1, c) \triangleq t_1 \left(1 + \frac{c}{t_1-1}\right) \]  \hspace{1cm} (56) \\
\[ \nu_s(t_1, c) \triangleq t_1 \sqrt{1 - \frac{c}{(t_1-1)^2}} \]  \hspace{1cm} (57)

Then, as \( N, K \to \infty \), the following holds:

(i) **Almost sure convergence of the largest eigenvalue: phase transition phenomenon**

If \( t_1 > 1 + c^{1/2} \):

\[ \hat{\lambda}_1 \xrightarrow{a.s.} \sigma_v^2 \mu_s(t_1, c) \]  \hspace{1cm} (58)

If \( t_1 \leq 1 + c^{1/2} \):

\[ \hat{\lambda}_1 \xrightarrow{a.s.} \sigma_v^2 \mu_+(c) \]  \hspace{1cm} (59)

(ii) **Convergence in distribution of the largest eigenvalue**

Let \( m \) (with \( 1 \leq m \leq P \)) be the multiplicity of the first spike eigenvalue \( t_1 \).

If \( t_1 = \ldots = t_m > 1 + c^{1/2} \):

\[ N^{1/2} \frac{\hat{\lambda}_1 - \sigma_v^2 \mu_s(t_1, c)}{\sigma_v^2 \nu_s(t_1, c)} \xrightarrow{D} G_m \]  \hspace{1cm} (60)

If \( t_1 = \ldots = t_m = 1 + c^{1/2} \):

\[ N^{2/3} \frac{\hat{\lambda}_1 - \sigma_v^2 \mu_+(c)}{\sigma_v^2 \nu_+(c)} \xrightarrow{D} A_m \]  \hspace{1cm} (61)

If \( t_1 < 1 + c^{1/2} \):

\[ N^{2/3} \frac{\hat{\lambda}_1 - \sigma_v^2 \mu_+(c)}{\sigma_v^2 \nu_+(c)} \xrightarrow{D} \mathcal{W}_2 \]  \hspace{1cm} (62)

where \( A_m \) and \( G_m \) are distribution laws defined in Appendix B and C respectively.

**Proof:** The proof of claim (i) is due to Baik and Silverstein [12]; claim (ii) was found by Baik, Ben Arous and Péché [13] under the additional assumption of \( z_{ij} \) Gaussian with unit variance, and was generalized into this form by Féral, Péché [14] using results from Bai and Yao [15].

C. **Interpretation of the results**

1) **Validity of the assumptions:** All the assumptions \((A_1)-(A_4)\) are verified exactly for the noise part of \( Z \), whose entries are complex Gaussian random variables. For the signal part, the first two assumptions are guaranteed by construction of \( Z \): \((A_1)\) is given by (44) and \((A_2)\) is equivalent to (45) (provided that the variance of \( s \) is equally distributed between real and imaginary part, which is true for all types of complex signals used in communications). Assumption \((A_3)\) is also verified in practical cases.
Assumption \((A_4)\) is satisfied exactly by Gaussian signals, while there exist several types of signals (e.g. PSK, QAM) whose fourth moment is lower than that of a Gaussian random variable. However, since the type of primary signal is usually unknown to the secondary users, the Gaussian assumption is reasonable in general. In addition, since \(P < K\), most of matrix \(Z\) is represented by the noise part which does always satisfy \((A_4):\) therefore the theorem can be applied in almost all practical cases, even when this assumption does not hold exactly. The approximation introduced in this way is small and becomes negligible when the SNR of the primary signal is low, as shown in Sec. VI-E.

2) Phase transition phenomenon: The first important result implied by the theorem is the existence of a critical value of \(t_1\) that determines whether a signal component is identifiable or not. This behavior is called phase transition phenomenon. In fact, when \(t_1 \leq 1 + c^{1/2}\), the largest eigenvalue of the covariance matrix converges to the same value as in the pure-noise model, whereas for \(t_1 > 1 + c^{1/2}\), it converges to a larger value: \(\mu_s(t_1, c) > \mu_+(c)\). This property makes it possible to detect the presence of signals.

In case of \(P = 1\), the critical value can be expressed directly in terms of the SNR using (54):

\[
\rho > \frac{1}{\sqrt{KN}}
\]  

(63)

This relation also allows to determine the minimum number of samples for the detector to be able to identify signals with a given SNR.

3) Limiting distributions: The second claim of the theorem clarifies how the largest eigenvalue converges to the asymptotical limit. For non-identifiable components, the limiting distribution is the same as in the case of no signal. For components with eigenvalues placed exactly on the critical point, the limiting distribution is a generalization of the one encountered in the previous case: in fact, for \(m = 0\), \(A_0\) reduces to the Tracy-Widom law (Appendix B). For components above the critical value, we find the distributions \(G_m\): for \(m = 1\), which is the most common case in practical applications, \(G_1\) is simply the normal distribution; for \(m = 2\), we have derived a simple expression of the CDF of \(G_2\) in terms of the Gaussian error function (see Appendix C).

Finally, notice that both the events of eigenvalues exactly equal to the critical point and of eigenvalues with multiplicity larger than one are asymptotically events with zero probability. The results concerning these cases are mentioned for completeness, but are not important for practical applications. Therefore, the case (60) with \(G_1\) is by far the most important result of this theorem and allows to express \(P_{md}\). Furthermore, \(G_1\) does not even involve complicated calculations because it reduces to the Gaussian distribution.
D. Derivation of $F_{T|\mathcal{H}_1}$ and $P_{md}$

Thanks to the results of Theorem 4.1, we are now able to express the limiting probability distribution of the test statistic $T$ under the hypothesis $\mathcal{H}_1$ and, consequently, to derive an analytical expression for the probability of missed detection. From now on, we refer to the case of identifiable signals, i.e., we assume the $P$ signal components produce spiked eigenvalues above the critical limit $1 + c^{1/2}$.

The approach that we adopt is the same as in the case of $\mathcal{H}_0$: we define again

$$L_1 \triangleq N^{1/2} \frac{\hat{\lambda}_1 - \sigma_v^2 \mu_s(t_1, c)}{\sigma_v^2 \nu_s(t_1, c)} \quad (64)$$

which, for claim (ii), has a limiting PDF

$$f_{L_1}(z) \rightarrow f_{G_m}(z) \quad (65)$$

where $f_{G_m}(\cdot)$ represents the PDF associated with $G_m$ ($m$ is the multiplicity of $t_1$), as defined in Appendix C.

As for the distribution of smallest eigenvalue, we introduce the following theorem.

**Theorem 4.2:** Distribution of the $K - P$ smallest eigenvalues under $\mathcal{H}_1$. Assume that for $K, N \to \infty$

$$c = \frac{K}{N} \to \bar{c} \in (0, 1) \quad (66)$$

and that $t_p > 1 + c^{1/2}$ for $1 \leq p \leq P$, the eigenvalues $\hat{\lambda}_{P+1}, \ldots, \hat{\lambda}_K$ of $R(N)$ have asymptotically the same limiting distribution as those of a $(K - P) \times (K - P)$ Wishart matrix.

**Proof:** The result follows from the proof of Lemma 2 in [16].

Therefore, the distribution of the smallest eigenvalue is not affected by the presence of “spikes” and claims (iii) and (iv) of Theorem 3.1 can be applied also in this case with the only difference that, instead of $c$ (24), now

$$c' = \frac{K - P}{N} \quad (67)$$

Thus, we define

$$L_K \triangleq N^{2/3} \frac{\hat{\lambda}_K - \sigma_v^2 \mu_-(c')}{\sigma_v^2 \nu_-(c')} \quad (68)$$

which still converges in distribution to the Tracy-Widom law

$$f_{L_K}(z) \rightarrow f_{W_2}(z) \quad (69)$$

Then the test statistic $T$ becomes

$$T = \frac{\hat{\lambda}_1}{\hat{\lambda}_K} = \frac{N^{-1/2} \nu_s(t_1, c)L_1 + \mu_s(t_1, c)}{N^{-2/3} \nu_-(c')L_K + \mu_-(c')} \quad (70)$$
Also in this case the noise variance $\sigma^2_v$ is canceled out in the ratio. However, an implicit dependence on $\sigma^2_v$ remains in the term $t_1$, except for the case of single primary user ($P = 1$) where $t_1$ is a function of the SNR only (54).

We denote with $l_1$ and $l_K$, respectively, the numerator and the denominator of $T$ and with $\overline{f}_{l_1}(z)$ and $\overline{f}_{l_K}(z)$ their limiting PDFs for $N,K \to \infty$. Through a random variable transformation, they may be expressed as

$$\overline{f}_{l_1}(z) = \frac{N^{1/2}}{\nu_s(t_1,c)} f_{\mathcal{G}_m} \left( \frac{N^{1/2}}{\nu_s(t_1,c)} (z - \mu_s(t_1,c)) \right)$$

(71)

$$\overline{f}_{l_K}(z) = \frac{N^{2/3}}{2^3 c'} f_{\mathcal{W}_2} \left( \frac{N^{2/3}}{2^3 c'} (\mu - (c') - z) \right)$$

(72)

Notice that, as a consequence of the observations in IV-B, $\mathcal{G}_m$ is with probability one a Gaussian distribution and thus it can be written in a more practical form as

$$\overline{f}_{l_1}(z) = \left( \frac{N}{2\pi} \right)^{1/2} \exp \left[ -\frac{N}{2 \nu_s^2(t_1,c)} (z - \mu_s(t_1,c))^2 \right]$$

(73)

Also in this case, we assume $f_{l_1}(l_1)$ and $f_{l_K}(l_K)$ as asymptotically independent. The resulting limiting ratio distributions is

$$\overline{f}_{T|\mathcal{H}_1}(t) = \left[ \int_{-\infty}^{+\infty} x \overline{f}_{l_1,t_K}(tx,x) dx \right] \cdot I_{\{t>1\}}$$

$$= \left[ \int_{0}^{+\infty} x \overline{f}_{l_1}(tx) \overline{f}_{l_K}(x) dx \right] \cdot I_{\{t>1\}}$$

(74)

where, like in the previous case, the domain of integration has been restricted to non-negative values, and the condition $t > 1$ is necessary to ensure that $l_1 > l_K$.

Finally, denoting with $\overline{F}_{T|\mathcal{H}_1}(\gamma)$ the CDF corresponding to the PDF in (74), we can take the approximation

$$\overline{F}_{T|\mathcal{H}_1}(\gamma) \approx \overline{F}_{T|\mathcal{H}_1}(\gamma)$$

(75)

that, in the asymptotical limit for $N$ and $K$, is the expression of the missed detection probability as it is given by (23). Numerical results show that the approximation is quite accurate for all cases of practical interest.

The relation between $P_{md}$ and $\gamma$ allows to predict the missed-detection probability of the detector with a given threshold, or to express the decision threshold as a function of the required probability of missed detection. The problem of setting the threshold is discussed in more detail in the next section.
V. SETTING THE DECISION THRESHOLD

The results presented in the previous sections express $P_{fa}$ and $P_{md}$ as a function of $\gamma$; therefore, by inverting the relations (22) and (23), the threshold can be expressed as a function of the error probabilities.

A. Threshold as a function of $P_{fa}$

The first relation
\[
\gamma(P_{fa}) = F_{T|\mathcal{H}_0}^{-1}(1 - P_{fa})
\]  
(76)
allows to set the decision threshold accurately even if the noise power ($\sigma^2_v$) is unknown, since $F_{T|\mathcal{H}_0}$ depends only on the number of receivers ($K$) and of samples ($N$). The threshold set in this way, as a function of a target $P_{fa}$, is therefore a “blind” decision scheme as it is insensitive both to the noise and to the signal power.

In a previous work, Zeng and Liang [2] proposed a similar approach to set the decision threshold as a function of the probability of false alarm. Their detection algorithm was based on an approximated distribution of $T$, calculated taking into account only the limiting distribution of the largest eigenvalue (Theorem 3.1(ii)), and therefore provides non-optimal detection performance. In [1] another eigenvalue-based detection scheme was proposed, based only on the asymptotical values of $\hat{\lambda}_1$ and $\hat{\lambda}_K$ (Theorem 3.1(i)(iii)). For this reason, it does not allow to adjust the threshold as a function of $P_{fa}$ and is strongly sub-optimal with respect to our scheme unless $N$ and $K$ are extremely large.

A detailed performance comparison between the threshold based on the limiting distribution $F_{T|\mathcal{H}_0}$ and these two previous approaches was provided in [21].

B. Threshold as a function of $P_{md}$

The second relation is
\[
\gamma(P_{md}) = F_{T|\mathcal{H}_1}^{-1}(1 - P_{md})
\]  
(77)
Whereas $\gamma(P_{fa})$ has been found to depend only on $K$ and $N$, the expression of $\gamma(P_{md})$ depends also on the characteristics of the signal to be detected. In particular, two cases have to be considered separately:

- when $P = 1$, the only additional parameter needed to compute $P_{md}$ is the SNR $\rho$. In this case, the detector may still be defined “blind” since it does not need to know explicitly the noise power nor the signal power. (Clearly, the detection performance has to be related, at least, with the SNR.
For instance, in the case of Energy Detection, the SNR and the noise power are needed to compute \( P_{md} \).

- when \( P > 1 \), the knowledge of additional parameters is needed, namely the noise power (\( \sigma_v^2 \)), the number of primary users (\( P \)), their powers (\( \sigma_1^2, \ldots, \sigma_P^2 \)), and the channel (\( H \)). These dependences arise from the nonlinear expression of \( t_1 \) (48).

In general, all these parameters (even the SNR and the potential number of primary users) might be unknown. Therefore, the relation between \( \gamma \) and \( P_{md} \) should better be used in the forward way, to predict the \( P_{md} \) achieved using a given threshold under the possible primary signal scenarios, rather than to set the decision threshold according to a target \( P_{md} \). Nevertheless, if the system imposes a certain requirement on \( P_{md} \) to keep the interference caused by the secondary network below a maximum level, the formula is useful to determine \( \gamma \) based on the worst-case scenario (i.e., the one with the highest missed-detection probability) so as to guarantee in all cases the required protection to the primary network.

C. Complexity and practical implementation

As shown in [2] and [21], eigenvalue-based detection schemes offer a substantial performance improvement compared to ED (and a complete protection to noise uncertainty) at the price of an increased complexity. Most of the computational complexity of these algorithms derives from the computation of the covariance matrix and of its eigenvalues: in [2] it is estimated that such operations lead to a complexity that grows as \( K^3 \), whereas in the case of ED it grows linearly with \( K \). This increased computational cost is not dramatic, since the number of receivers is never enormous. On the other hand, in terms of the sample number (which is, actually, very large) the complexity remains linear with \( N \) for both EBD and ED.

However, it is important to remark that the computational complexity is not influenced by the computation of the threshold. Even if the formulae found in this paper to express the threshold are very complex, they are always implemented off-line, and what the detector uses is simply a look-up table (LUT) containing several values of \( \gamma \) as a function of \( N, K, P_{fa} \), and/or \( P_{md} \) and SNR. The use of LUTs also allows to change the decision threshold “on the fly”, in case of modifications of the system requirements.

Finally, for the computation of the distribution functions defined in this paper, routines are available on the web (e.g., [18] for the Tracy-Widom distributions) or can be implemented directly from the definitions given in the Appendices.
VI. NUMERICAL RESULTS

In this section, the results derived analytically in the previous sections are validated by comparing them with empirical results, obtained from Matlab Monte-Carlo simulations. The parameters used in the simulations are described in each sub-section; when referring to the SNR, it is defined according to (4).

A. Distribution of $T$ under $\mathcal{H}_0$

Figure 1 represents the probability of false alarm, i.e., the complementary CDF of $T$ under $\mathcal{H}_0$, for $N = 1000$ and different values of $K$ (i.e., of $c$). The value of $\sigma_v^2$ has no effect, as it gets canceled out in the test statistic.

The curve predicted using the analytical expression turns out to be consistent with the empirical data in all the considered cases. Comparing the three curves obtained with different values of $K$, one may observe that for a given $\gamma$ the probability of false alarm increases with $K$. However, this does not mean that the detector performance worsens for larger $K$, because also the curve of $P_{md}$ shifts rightwards, and consequently the decision threshold. The overall effect is indeed an improvement of performance when $K$ gets larger, as expected intuitively.

B. Distribution of $T$ under $\mathcal{H}_1$

Figures 2, 3 and 4 show the probability of missed detection, i.e., the CDF of $T$ under $\mathcal{H}_1$, for the same values of $N$ and $K$ as in the previous case.

The entries of $H$ are taken as zero-mean complex Gaussian random coefficients (Rayleigh fading), with a variance normalized so as to obtain the desired SNR. In the first figure the SNR is $-10$ dB with $P = 1$ primary signal; in the second one, the SNR is $-20$ dB again with $P = 1$; in the third one, $P = 2$ with a global SNR of $-10$ dB (from (5) with: $\rho_1 = \frac{\sigma^2_1 \|h_1\|^2}{K \sigma_c^2} \approx -12.2$ dB; $\rho_2 = \frac{\sigma^2_2 \|h_2\|^2}{K \sigma_c^2} \approx -14.0$ dB; $\sigma_v^2 = 1$). Notice that in the last case ($P > 1$) the largest spike eigenvalue $t_1$, which determines $P_{md}$, depends on all the entries of $H$ and not only on the SNR. In our simulations $t_1 = \{2.25, 4.04, 7.60\}$, respectively for $K = \{20, 50, 100\}$.

Also in this case, the analytical curves fit the empirical data well in all the considered cases. We have considered low values of SNR, since the low-SNR region is the most important both from the theoretical point of view ($t_1$ close to the critical value of identifiability) and from the practical point of view (the challenge for cognitive radios is to detect signals also in presence of fading or shadowing).
As previously mentioned, the curves of $P_{md}$ shift rightwards as $K$ increases, i.e., the missed-detection probability gets lower for a given $\gamma$. This fact compensates the increase of $P_{fa}$ resulting in a larger separation between $1 - F_{T|H_0}$ and $F_{T|H_1}$ for larger $K$.

C. Convergence

Figures 5 and 6 show the convergence of the empirical CDFs to the analytical CDFs, which are calculated under asymptotical assumptions for $N$ and $K$. Four different couples of $\{N, K\}$ have been considered while keeping their ratio $c$ fixed at 0.1. Remarkably, even though the CDFs are asymptotical they provide an accurate approximation of the empirical CDFs also for low $K$ and $N$.

In the case $H_0$, as $N$ and $K$ increase the CDF tends to a step function, because the largest and the smallest eigenvalues converge (almost surely) to the values $\mu_+(c)$ and $\mu_-(c)$, respectively; the variance instead depends also on $N$ (it gets smaller for larger $N$).

For the case $H_1$, we considered a scenario with $P = 1$ and, to make the comparison more evident, we kept $t_1$ fixed instead of the SNR ($\rho$ and $t_1$ are linked by a factor $K$, so they can not remain both constant with different $K$). In particular we chose the value $t_1 = 2$, which is above the critical value that is $1 + \sqrt{c} = 1.3162$ for all the considered couples of $\{N, K\}$. Similarly as in the previous case, the CDFs turn out to converge to a step function corresponding to the almost sure asymptotical limits of the eigenvalues.

D. Identifiability

As a result of the phase transition phenomenon of Theorem 4.1 signals below a certain power level are not identifiable. A detection limit as a function of the SNR is expressed by the relation (63), valid for $P = 1$. Figure 7 represents graphically the critical SNR for detection as a function of the number of samples $N$ and of receivers $K$. The relation may be used to determine the minimum sensing duration (i.e., the minimum number of samples) needed to detect signals for a required detector sensitivity.

A relation between identifiability threshold and SNR is valid only for $P = 1$. For multiple signals, the expression of $t_1$ is more complex and does not depend only on the SNR. However, it turns out that also for $P > 1$ the value of $t_1$ is determined essentially by the power of the largest signal, i.e., by the SNR as if the first component was alone. Therefore, we may define an approximated expression of the SNR, similar as (6), depending only on the power of the dominant signal component:

$$\rho \approx \frac{\max_p \left( \sigma_p^2 \| h_p \|^2 \right)}{K \sigma_v^2}$$

(78)
This expression can be used in (63) to determine, approximately, the parameters $N$ and $K$ of the detector. As an example, in figure 8 we consider the case $P = 2$ with $\rho_1$ fixed at $0.1 = -10 \text{ dB}$ and $\rho_2$ varying from 0 and $\rho_1$. The graph shows $t_1$ as a function of $\rho_2$, comparing the case when $t_1$ is calculated from the exact formula for $P = 2$ (11) with the case when it is calculated taking into account the largest component only (78) and with the case of a single component, but with double power (SNR = $2\rho_1$). It turns out that the actual value of $t_1$ is very close to the approximated one, even when the sum of $\rho_1$ and $\rho_2$ is close to $2\rho_1$. Furthermore, the approximated $t_1$ tends to underestimate the actual $t_1$, resulting in a conservative choice of $N$ and $K$.

E. Non-Gaussian signals

As pointed out in Sec. IV-B, the last assumption of Thereom 4.1 is often not satisfied in practice, since realistic signals have typically a fourth moment lower than that of a Gaussian random variable. Figures 9 and 10 show how the theoretical results, which rely on that assumption, fit empirical data obtained using more realistic types of primary signal. We considered four different types of signals, all with the same variance as in the Gaussian case, but with different fourth moments. The first curve refers to a 4-PSK modulated primary signal, with ideal rectangular pulse-shape filter and assuming a coherent reception; in the second curve, the signal is the same but passed through a square root raised cosine (SRRC) filter with roll-off $\alpha = 0.5$; the third curve is a PSK signal with non-coherent reception (i.e., each sample has a random phase); the last curve refers to a random complex signal whose real and imaginary parts are uniformly distributed.

In the first figure, when the SNR is very low ($-20 \text{ dB}$), the theoretical distribution fits the empirical data perfectly in spite of the fourth moment of the signals. When the SNR increases ($-10 \text{ dB}$), some difference between the theoretical and the empirical curve can be observed, especially for PSK signals. It is interesting to notice that the Gaussian approximation on the fourth moment affects the variance of the resulting distribution, but not the mean. The result is that the analytical formula overestimates the probability of missed detection (the interesting part of the curve is for $P_{md} < 0.5$, i.e., the left tail).

To obtain a more accurate estimation of the missed-detection probability in case of non-Gaussian signals, for high SNR, one should add a “correction coefficient” to the theoretical variance $\nu_s(t_1,c)$. Such coefficients would depend on the fourth moment of the signals, $\sigma_4^P$, and would be therefore specific of the modulation used. It might be possible to determine by simulation the correction coefficients for a particular signal as a function of the SNR, whereas determining them analytically is a more challenging task since the matrix $Z$ is composed of heterogeneous entries. However, the Gaussian assumption is valid.
asymptotically for $\rho \to -\infty$ (the signal part in $Z$ becomes negligible) and is accurate enough in the low-SNR region as shown by figure 9.

F. Receiver operating characteristics (ROC)

Figures 11 and 12 represent the performance of the eigenvalue-based detector in the form of complementary ROC (receiver operating characteristics), i.e., $P_{md}$ as a function of the target $P_{fa}$. The curves are plotted by setting the threshold as a function of the false-alarm probability and deriving the corresponding missed-detection probability for that threshold. The graphs compare the curves obtained from the empirical distributions with those obtained using the analytical expressions of this paper: (76) to set $\gamma(P_{fa})$, then (23) to compute $P_{md}(\gamma)$.

The first ROC graph refers to the same scenario as figures 1 (for $P_{fa}$) and 3 (for $P_{md}$), with $N = 1000$ and $K = 50$; the second one refers to the scenario of figure 4 (for $P_{md}$) with the same values of $N$ and $K$.

The overall detector performance expressed by the ROC improves as the separation between the $P_{fa}$ curve (monotonically decreasing) and the $P_{md}$ curve (monotonically increasing) gets larger, thus letting both $P_{fa}$ and $P_{md}$ be nearly zero for a wide range of $\gamma$. Such distance increases with $K$, $N$ and with the SNR. For this reason, in the second ROC the performance is almost ideal (zero $P_{md}$ for all the $P_{fa}$). In the first ROC on the contrary there are finite missed-detection probabilities for the considered range of $P_{fa}$; the analytical result also in this case turns out to be consistent with the empirical data.

VII. CONCLUSION

In this paper, analytical formulae have been found for the limiting distribution of the ratio between the largest and the smallest eigenvalue in sample covariance matrices, either constructed from pure-noise (Wishart) models or signal-and-noise (spiked population) models. These results have been applied to the problem of signal detection (in particular, in the context of Cognitive Radio), where eigenvalue-based detection has proved to be an efficient technique.

Among the main results of the paper, there are the analytical formulation of the missed detection probability as a function of the threshold, and the derivation and discussion of signal identifiability conditions. All the results have been validated via numerical simulations covering false-alarm and missed-detection vs. threshold, convergence behavior, identifiability for single and multiple primary users as a function of the SNR, validity of the approach for realistic modulated signals, ROC curves.
A. Tracy-Widom distribution

The Tracy-Widom distributions $W_2$ were introduced in [17], to express the distribution of the largest eigenvalue in a Gaussian Unitary Ensemble (GUE). Define the complex Airy function,

$$Ai(u) = \frac{1}{2\pi} \int_{\infty e^{i\pi/6}}^{\infty e^{i\pi/6}} e^{jua + j\frac{3}{2}a^3} da$$

(79)

the Airy kernel,

$$A(u, v) = \frac{Ai(u)Ai'(v) - Ai'(u)Ai(v)}{u - v}$$

(80)

and let the $A_x$ be the operator acting on $L^2((x, +\infty))$ with kernel $A(u, v)$. Then, the second-order Tracy-Widom CDF, $F_{W_2}(x)$, is defined in terms of the Fredholm determinant

$$F_{W_2}(x) = \det(1 - A_x)$$

(81)

It also admits an alternative expression. Let $q(u)$ be the solution of the Painlevé II differential equation

$$q''(u) = uq(u) + 2q^3(u)$$

(82)

satisfying

$$q(u) \sim -Ai(u), \quad u \to +\infty$$

(83)

Then

$$F_{W_2}(x) = \exp \left(- \int_x^{+\infty} (u - x)q^2(u) du \right)$$

(84)

Notice that this definition, and the index 2, are referred to the case of complex Gaussian variables. In the case of real signals, one should use the corresponding first-order Tracy-Widom distribution [17].

B. Airy-type distributions

These distributions are defined in [13] as an extension of the Tracy-Widom (GUE) distribution. Let

$$s^{(m)}(u) = \frac{1}{2\pi} \int_{\infty e^{i\pi/6}}^{\infty e^{i\pi/6}} e^{jua + j\frac{3}{2}a^3} \frac{1}{(ja)^m} da$$

(85)

$$t^{(m)}(u) = \frac{1}{2\pi} \int_{\infty e^{i\pi/6}}^{\infty e^{i\pi/6}} e^{jua + j\frac{3}{2}a^3} (ja)^{m-1} da$$

(86)

Then, for $k \geq 1$, the CDFs of $A_k$ are defined as

$$F_{A_k}(x) = \det(1 - A_x) \cdot \det \left( \delta_{mn} - \frac{1}{1 - A_x} s^{(m)}, t^{(n)} \right)_{1 \leq m,n \leq k}$$

(87)
where $\langle \cdot , \cdot \rangle$ is the real inner product of functions in $L^2((x, +\infty))$.

For $k = 0$, this distribution reduces to the GUE distribution:

$$F_{A_0}(x) = F_{W_2}(x) \quad (88)$$

For $k = 1$, it can be written in the Painlevé form

$$F_{A_1}(x) = F_{W_2}(x) \exp \left( \int_x^{+\infty} q(u) du \right) \quad (89)$$

C. Finite GUE distributions

The distributions $G_k$ are defined in [13] as the distribution of the largest eigenvalue in a $k \times k$ GUE. Their CDF is

$$F_{G_k}(x) = (2\pi)^{-k/2} \left( \prod_{m=1}^{k} m! \right)^{-1} \cdot$$

$$\cdot \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} \prod_{1 \leq m < n \leq k} |\xi_m - \xi_n|^2 \cdot \prod_{m=1}^{k} e^{-\frac{1}{2} \xi_m^2} d\xi_1 \cdots d\xi_k \quad (90)$$

In the case $k = 1$, it is simply a zero-mean, unit-variance Gaussian distribution:

$$F_{G_1}(x) = \frac{1}{2\pi} \int_{-\infty}^{x} e^{-\frac{1}{2} \xi^2} d\xi \triangleq \mathcal{E}(x) \quad (91)$$

We also introduce here a compact expression for CDF and PDF in the case $k = 2$, in terms of the Gaussian error function:

$$F_{G_2}(x) = \mathcal{E}^2(x) - \frac{1}{\sqrt{2\pi}} xe^{-\frac{x^2}{2}} \mathcal{E}(x) - \frac{1}{2\pi} e^{-x^2} \quad (92)$$

$$f_{G_2}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left( 1 + x^2 \right) \mathcal{E}(x) + \frac{1}{2\pi} xe^{-x^2} \quad (93)$$

These expressions do not appear in [13].

REFERENCES

[1] L.S Cardoso, M. Debbah, P. Bianchi, J. Najim, “Cooperative spectrum sensing using random matrix theory”, 3rd International Symposium on Wireless Pervasive Computing (ISWPC) 2008, pp.334-338, 7-9 May 2008.

[2] Y. Zeng, Y.-C. Liang, “Maximum-Minimum Eigenvalue Detection for Cognitive Radio”, 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communication (PIMRC) ’07, 2007.

[3] A. M. Tulino and S. Verdú, “Random matrix theory and wireless communications”, in Foundations and Trends in Communications and Information Theory, vol. 1, (Hanover, MA, USA), Now Publishers Inc., 2004.

[4] J. Wishart “The generalized product moment distribution in samples from a normal multivariate population”, Biometrika, vol. 20A, pp. 32-52, 1928.
[5] V. A. Marchenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices”, *Math USSR-Sbornik*, vol.1, pp.457-483, 1967

[6] Z. D. Bai, “Methodologies in spectral analysis of large-dimensional random matrices, a review”, *Statistica Sinica*, vol.9, pp.611-677, 1999.

[7] K. Johansson, “Shape fluctuations and random matrices”, *Comm. Math. Phys.*, 209:437-476, 2000.

[8] I. Johnstone, “On the distribution of the largest eigenvalue in principal component analysis”, *Ann. Statist.*, 29:295-327, 2001.

[9] A. Soshnikov, “A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices”, *J. Statist. Phys.*, 108, no. 5-6, 10331056, 2002.

[10] S. Péché, “Universality results for largest eigenvalues of some sample covariance matrix ensembles”, preprint: [http://arxiv.org/abs/0705.1701](http://arxiv.org/abs/0705.1701) 2007

[11] O. N. Feldheim, S. Sodin, “A universality result for the smallest eigenvalues of certain sample covariance matrices”, preprint: [http://arxiv.org/abs/0812.1961](http://arxiv.org/abs/0812.1961) 2008

[12] J. Baik and J. Silverstein, “Eigenvalues of large sample covariance matrices of spiked population models”, *Journ. of Mult. Anal.*, 97, 1382-1408 (2006).

[13] J. Baik, G. Ben Arous, G. and S. Péché, “Phase transition of the largest eigenvalue for non-null complex sample covariance matrices”, *Ann. Probab.*, 33 no. 5, 1643-1697, 2005.

[14] D. Féral, S. Péché, “The largest eigenvalues of sample covariance matrices for a spiked population: diagonal case”, preprint: [http://arxiv.org/abs/0812.2320](http://arxiv.org/abs/0812.2320) 2008.

[15] Z. Bai, J. Yao, J, “Central limit theorems for eigenvalues in a spiked population model”, *Ann. Inst. H. Poincaré*, vol. 44, no. 3, 447474, 2008.

[16] S. Kritchman and B. Nadler, “Determining the number of components in a factor model from limited noisy data”, *Chemometrics and Intelligent Laboratory Systems*, 94:19-32, 2008.

[17] C. Tracy and H. Widom, “On orthogonal and symplectic matrix ensembles”, *Comm. Math. Phys.*, vol.177, pp.727-754, 1996.

[18] M. Dieng, *RMLab ver. 0.02: [http://math.arizona.edu/~momar/](http://math.arizona.edu/~momar/)* 2006

[19] M. Brookes, “The Matrix Reference Manual”, online: [http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html](http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html) 2005.

[20] J. H. Curtiss, “On the Distribution of the Quotient of Two Chance Variables”, *The Annals of Mathematical Statistics* 12(4):409-421, 1941.

[21] F. Penna, R. Garello, M. A. Spirito, “Cooperative Spectrum Sensing based on the Limiting Eigenvalue Ratio Distribution in Wishart Matrices”, accepted for publication in *IEEE Comm. Letters* (preprint: arxiv.org/abs/0902.1947), 2009

[22] F. Penna, R. Garello, D. Figlioli, M. A. Spirito, “Exact Non-asymptotic Threshold for Eigenvalue-based Spectrum Sensing”, accepted for publication in *Proc. 4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Hannover, Germany, Jun 2009
Fig. 1. False-alarm probability: empirical vs. analytical.

Fig. 2. Missed-detection probability: empirical vs. analytical. $P = 1$, $\rho = -10$dB.
Fig. 3. Missed-detection probability: empirical vs. analytical. \( P = 1, \rho = -20 \text{dB} \).

Fig. 4. Missed-detection probability: empirical vs. analytical. \( P = 2, \rho = -10 \text{ dB} \) \((\rho_1 = 0.06 \approx -12.2 \text{ dB}, \rho_2 = 0.04 \approx -14.0 \text{dB})\).
Fig. 5. False-alarm probability: convergence, for a fixed $c = K/N = 0.1$.

Fig. 6. Missed-detection probability: convergence, for a fixed $c = K/N = 0.1$. $P = 1$, $t_1 = 2$. 
Fig. 7. Phase transition phenomenon: minimum identifiable SNR vs. $N$ and $K$.

Fig. 8. Impact of the approximated SNR formula ($\rho \approx \rho_1$) on the calculation of the largest spike eigenvalue ($t_1$).
Fig. 9. Missed-detection probability: empirical vs. analytical with non-Gaussian signals.. $P = 1$, $\rho = -20\text{dB}$.

Fig. 10. Missed-detection probability: empirical vs. analytical with non-Gaussian signals.. $P = 1$, $\rho = -10\text{dB}$. 
Fig. 11. Complementary ROC: analytical vs. empirical. $P = 1, \rho = -20$ dB.

Fig. 12. Complementary ROC: analytical vs. empirical. $P = 2, \rho = -10$ dB.