Asymptotics of eigenbased collaborative sensing

Pascal Bianchi and Jamal Najim
Telecom Paristech
Paris, France
{jamal.najim,pascal.bianchi}@telecom-paristech.fr

G. Alfano and Mérouane Debbah
SUPELEC
Gif-sur-Yvette, France
{giusi.alfano, merouane.debbah}@supelec.fr

***Abstract*—In this contribution, we propose a new technique for collaborative sensing based on the analysis of the normalized (by the trace) largest eigenvalues of the sample covariance matrix. Assuming that several base stations are cooperating and without the knowledge of the noise variance, the test is able to determine the presence of mobile users in a network when only few samples are available. Unlike previous heuristic techniques, we show that the test has roots within the Generalized Likelihood Ratio Test and provide an asymptotic random matrix analysis enabling to determine adequate threshold detection values (probability of false alarm). Simulations sustain our theoretical claims.**

I. Introduction

Recently, cognitive networks [1] have been advocated as one of the major solutions to increase the spectral efficiency of wireless systems by making full use of the available spectrum. This requires networks to be able to exploit opportunistically spectrum left-overs, by means of knowledge of the environment and cognition capability, and to adapt their radio parameters accordingly. Unfortunately, usual cognitive radio techniques, based on sensing, require either a large number of received samples or the precise knowledge of the system parameters (such as the noise variance or the structure of the signal). The techniques range from energy detector [2], [3], [4], matched filter detector or ciclostationary methods.

In this contribution, we propose a collaborative sensing technique using very limited knowledge on the signal model (noise variance unknown) adapted to highly mobile environments where only a few number of samples can be acquired. We use the space dimension and suppose that various base stations in the network can cooperate (through a virtual MIMO system) to sense the received signal. The technique based on the analysis of the normalized (by the trace) of the maximum eigenvalue of the sample covariance matrix originates from the derivation of the Generalized Likelihood Ratio Test (GLRT). Interestingly, we are able to compute threshold detection values (probabilities of false alarm) by using recent results of asymptotic random matrix theory and show that the statistics of the test converge to a multivariate Tracy-Widom distribution. The results are valid for any number of users in the network. The paper is articulated as follows: next Section focuses on the problem formulation and contains the signal model, while in Section III the GLRT approach is exploited. Maximum likelihood estimates of the unknown parameters involved in the test are evaluated and their asymptotic behavior are discussed in Section IV. A brief numerical assessment follows in Section V, while Conclusions are given in Section VI.

II. Problem Formulation

A. Signal Model

Consider a secondary wireless network formed by K nodes, working in sensing mode. We assume that all K nodes are simultaneously sensing a given sub-band $\mathcal{B}$ of the spectrum. For each $k = 1, \dots, K$, we denote by $y_k(n)$ the complex envelope of the signal received by the kth sensor in band $\mathcal{B}$ after proper filtering and sampling. Denote by $\boldsymbol{y}(n) = [y_1(n), \dots, y_K(n)]^T$ the vector obtained when stacking all K sensors' observations at time n into a column vector. The aim is to detect the presence of one or several primary transmitters in band $\mathcal{B}$. We respectively denote by H_0 and H_1 the hypotheses corresponding to the case where "band $\mathcal{B}$ is free" and "one or several primary devices are already transmitting in band $\mathcal{B}$":

$$\boldsymbol{y}(n) = \begin{cases} \boldsymbol{w}(n): & H_0 \\ \mathbf{H}\, \boldsymbol{s}(n) + \boldsymbol{w}(n): & H_1 \end{cases}, \tag{1}$$

where $\boldsymbol{w}(n)$ represents a complex circular temporally-white Gaussian noise vector with zero mean and covariance matrix equal to $\sigma^2 \mathbf{I}_K$. In the H_1-case, $\boldsymbol{s}(n) = [s_1(n), \dots, s_P(n)]^T$ denotes the unknown P-dimensional process sent by the primary active devices. Integer P denotes the number of active transmitters in the band of interest. Sequence $\boldsymbol{s}(n)$ is assumed to be an independent identically distributed (i.i.d.) zero mean random sequence with independent entries. We assume without restriction that $s_p(n)$ has unit variance for each p. Matrix $\mathbf{H} \in \mathbb{C}^{K \times P}$ represents the complex-valued Multiple-Input Multiple-Output (MIMO) channel between the P transmitters and the K receiving nodes. In our context, most parameters are unknown. In particular:

- the noise variance σ^2 is unknown,
- the channel matrix $\mathbf{H}$ is unknown.

Depending on the context, the number of transmitters P may either be known or unknown. In case P is unknown, it is usually reasonable to assume that there exists a known integer P_{max} such that $P \le P_{max} < K$. In that case, it is always possible to test hypothesis H_0 versus H_1, where P is replaced with P_{max}. More involved order detection methods may as well be used, but such methods are out of the scope of this paper. In the sequel, we assume that P is known. Robustness issues related to model order mismatch will be investigated in an extended version of this paper.

B. Main Objective

In the sequel, we denote by N the number of samples observed by each sensor k. Consider the following $K \times N$ data matrix $\mathbf{Y}$:

$$\mathbf{Y} = [\boldsymbol{y}(0), \ldots, \boldsymbol{y}(N-1)] \,. \tag{2}$$

In order to test hypothesis H_0 versus H_1, the aim is to construct a relevant test function $\varphi : \mathbb{C}^{K\times N} \to \{0,1\}$ with the sense that one decides hypothesis H_0 (resp. H_1) whenever $\varphi(\mathbf{Y}) = 0$ (resp. $\varphi(\mathbf{Y}) = 1$). As usual, we restrict ourselves to the search for test functions such that the probability of false alarm does not exceed a predefined threeshold ϵ *i.e.*,

$$\mathbb{P}_{H_0}\left[\varphi(\mathbf{Y}) = 1\right] \leq \epsilon \,, \tag{3}$$

where $\mathbb{P}_{H_0}[\mathcal{E}]$ represents the probability of a given event $\mathcal{E}$ under hypothesis H_0.

III. Generalized Likelihood Ratio Test

In the present section, we investigate the case where input symbols $\boldsymbol{s}(n)$ are supposed to be Gaussian distributed: $\boldsymbol{s}(n) \sim \mathcal{CN}(0, \mathbf{I}_P)$ where $\mathbf{I}_P$ denotes the $P\times P$ identity matrix. In this case, the generalized likelihood ratio test can be expressed.

A. Likelihood Ratio

We respectively denote by $p_0(\mathbf{Y};\sigma^2)$ and $p_1(\mathbf{Y};\mathbf{H},\sigma^2)$ the likelihood functions of the observation matrix $\boldsymbol{y}$ indexed by the unknown parameters $\mathbf{H}$ and σ^2 under hypotheses H_0 and H_1 respectively:

$$p_0(\mathbf{Y};\sigma^2) = (\pi\sigma^2)^{-NK} \exp\left(-\frac{N}{\sigma^2}\mathrm{tr}\,\hat{\mathbf{R}}\right) \tag{4}$$

$$p_1(\mathbf{Y};\mathbf{H},\sigma^2) = (\pi^K \det \mathbf{R})^{-N} \exp\left(-N\mathrm{tr}\,(\hat{\mathbf{R}}\mathbf{R}^{-1})\right) \tag{5}$$

where $\mathbf{R} = \mathbf{R}(\mathbf{H},\sigma^2)$ is the true covariance matrix under H_1 defined by

$$\mathbf{R} = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_K$$

and where $\hat{\mathbf{R}}$ is the sampled covariance matrix:

$$\hat{\mathbf{R}} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H \,.$$

In the ideal case where parameters $\mathbf{H}$ and σ^2 are supposed to be available, it is well known that a uniformly most powerful test rejects the null hypothesis when ratio

$$L_N(\mathbf{Y}) = \frac{p_0(\mathbf{Y};\sigma^2)}{p_1(\mathbf{Y};\mathbf{H},\sigma^2)} \,. \tag{6}$$

lies below a certain threeshold which is selected so that (3) holds. Unfortunately, parameters $\mathbf{H}$ and σ^2 are unknown in our context so that a uniformly powerful test can no longer be defined. In this case, a suboptimal but classical approach consists in replacing the true likelihood ratio by the so-called generalized likelihood ratio (GLR) $\hat{L}_N(\mathbf{Y})$.

B. ML Estimates

The GLR is simply obtained by replacing the unknown parameter values $\mathbf{H}$ and σ^2 by their maximum likelihood (ML) estimates:

$$\hat{L}_N(\mathbf{Y}) = \frac{p_0(\mathbf{Y};\hat{\sigma}_0^2)}{p_1(\mathbf{Y};\hat{\mathbf{H}}_1,\hat{\sigma}_1^2)} \,. \tag{7}$$

where $\hat{\mathbf{H}}_1$ is the ML estimate of $\mathbf{H}$ under hypothesis H_1 and where $\hat{\sigma}_0^2$ (resp. $\hat{\sigma}_1^2$) is the ML estimate of σ^2 under hypothesis H_0 (resp. H_1). Denote by $\lambda_1 > \lambda_2 \cdots > \lambda_K \geq 0$ the ordered eigenvalues of $\hat{\mathbf{R}}$ (all distincts with probability one). For each $k = 1 \ldots K$, denote by $\boldsymbol{e}_k$ the $K \times 1$ eigenvector associated with λ_k. We provide the expression of the ML estimates $\hat{\sigma}_0^2$, $\hat{\sigma}_1^2$ and $\hat{\mathbf{H}}_1$. Note that the likelihood function is unchanged by right-multiplication of $\mathbf{H}$ with a $P \times P$ unitary matrix, thus $\mathbf{H}$ is identifiable only up to a unitary matrix. ML estimates are given by:

$$\begin{aligned} \hat{\sigma}_0^2 &= \frac{1}{K}\sum_{k=1}^{K}\lambda_k, \qquad \hat{\sigma}_1^2 = \frac{1}{K-P}\sum_{k=P+1}^{K}\lambda_k \\ \hat{\mathbf{H}}_1 &= [\boldsymbol{e}_1,\ldots,\boldsymbol{e}_P]\,\mathrm{diag}\left(\sqrt{\lambda_1 - \hat{\sigma}_1^2},\ldots,\sqrt{\lambda_P - \hat{\sigma}_1^2}\right)\mathbf{U}_P \end{aligned}$$

where $\mathbf{U}_P$ is a $P\times P$ unitary matrix indeterminacy. The proof of the above lemma is omitted due to the lack of space. We may now evaluate the GLR by substituting the values σ^2 and $\mathbf{R}$ in equations (4)-(5) with the corresponding ML estimates $\hat{\sigma}_0^2$ and $\hat{\mathbf{H}}_1\hat{\mathbf{H}}_1^H + \hat{\sigma}_1^2\mathbf{I}_K$ respectively. For each p, we define:

$$\mu_p = \frac{\lambda_p}{\frac{1}{K}\,\mathrm{tr}\,\hat{\mathbf{R}}} \,. \tag{8}$$

C. Proposed Hypothesis Test

The following result is a direct consequence of Lemma 1.

Proposition 1. *The GLR writes* $\hat{L}_N(\mathbf{Y}) = C^N \exp N\mathcal{L}_N$ *where* $C = \left(1 - \frac{P}{K}\right)^{K-P}$ *is a constant and where* $\mathcal{L}_N = \mathcal{L}_N(\mu_1,\ldots,\mu_P)$ *is the statistic defined by*

$$\mathcal{L}_N = \sum_{p=1}^{P}\log\mu_p + (K-P)\log\left(1 - \frac{1}{K}\sum_{p=1}^{P}\mu_p\right) \,. \tag{9}$$

The above result implies that the "trace-normalized" P largest eigenvalues $\mu_1,\ldots,\mu_P$ of the sampled covariance matrix form in some sense a sufficient statistic for the generalized likelihood ratio test. For technical reasons which will become clear in the sequel, we rather focus on the following "centered and rescaled" generalized log-likelihood ratio:

$$\bar{\mathcal{L}}_N = N^{2/3}\beta_N\left(\mathcal{L}_N - \alpha_N\right) \,. \tag{10}$$

Here, we defined the centering constant α_N by

$$\begin{aligned} \alpha_N &= 2P\log\left(1+\sqrt{c}\right) \\ &\quad - P\left(1+\sqrt{c}\right)^2\left(1 - \frac{P}{cN}\right) - \frac{P^2(1+\sqrt{c})^4}{2Nc} \end{aligned} \tag{11}$$

and the normalisation constant β_N by

$$\beta_N = \frac{-\left(1+\sqrt{c}\right)^{2/3}}{(2+\sqrt{c})\,c^{1/3}} \tag{12}$$

where we defined $c = K/N$. Clearly, the GLRT which consists in comparing (7) with a predefined threeshold is equivalent (in terms of false alarm and miss probabilities) to the following test:

$$\bar{\mathfrak{L}}_N \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_N \tag{13}$$

where γ_N is a suitable threshold.

Comments

• Clearly, the test statistic $\bar{\mathfrak{L}}_N$ is equivalent to the generalized likelihood ratio $\hat{L}_N(\mathbf{Y})$ in the sense that there exists a one-to-one mapping from one to other. The reason for using $\bar{\mathfrak{L}}_N$ instead of $\hat{L}_N(\mathbf{Y})$ or $\mathfrak{L}_N$ is that, when the dimensions N and K are large enough, the threeshold value γ_N does not depend on the system characteristics N, K or K/N. Universal threeshold tables are likely to be provided irrespective to the particular cognitive system of interest.

• In order to complete the definition of the test, we must determine the threshold value γ_N. As usual, γ_N is fixed so that the probability of false alarm associated with the test does not exceed a certain predefined value ϵ, as required by constraint (3). In order to maximize the power of our test while keeping the latter constraint satisfied, we select the threshold γ_N such that $F_{N,P}(\gamma_N) = \epsilon$, where

$$F_{N,P}(x) = \mathbb{P}_{H_0}\left[\bar{\mathfrak{L}}_N \leq x\right] \tag{14}$$

is the distribution function of random variable $\bar{\mathfrak{L}}_N$ under H_0. Otherwise stated, γ_N can simply be defined as $\gamma_N = F_{N,P}^{-1}(1-\epsilon)$, where $F_{N,P}^{-1}$ denotes the inverse of $F_{N,P}$ with respect to composition. Unfortunately, the expression of function $F_{N,P}$ is hardly tractable and its inversion would require involved numerical methods. In the next section, we propose to study the asymptotic behaviour of $F_{N,P}(x)$ in order to simplify the computation of the threshold value.

• The proposed test benefits from the prior knowledge of the number P of primary transmitters (or at least an upper bound on this number). In case such an information is not available, or in case $P \geq K$, the GLRT would reduce to the standard sphericity test based on the statistic $\det \hat{\mathbf{R}} / \left(\mathrm{tr}\, \hat{\mathbf{R}}\right)^K$. The proposed test can be interpreted as an extension of the above sphericity test to the case where the dimension P of the "signal-subspace" is known to be strictly less than K.

IV. Asymptotic Analysis

In the present section, we provide a simple procedure allowing to determine the threeshold value, based on the asymptotic analysis of the test statistic $\bar{\mathfrak{L}}_N$ under H_0. Our analysis is relevant in contexts where the number K of sensors is assumed to be large (*i.e.*, significantly larger than the number P of sources). Due to cognitive radio constraints, the secondary system must be able to decide the presence/absence of primary transmitters in a moderate amount of time. Therefore, we focus on the context where the number K of sensors and the number N of samples have the same order of magnitude. Otherwise stated, we consider the following asymptotic regime:

$$N \to \infty,\ K \to \infty,\ \ K/N \to c,\ \ P \text{ is fixed}, \tag{15}$$

where $0 < c < 1$ is a constant. It is worth stressing that under H_0, the distribution of $\mu_p = \lambda_p / (\frac{1}{K}\,\mathrm{tr}\,\hat{\mathbf{R}})$ does not depend on σ^2. Therefore, the distribution of $\bar{\mathfrak{L}}_N$ does not depend on σ^2. As a consequence, there is no restriction in assuming that

$$\sigma^2 = 1$$

in the present section, for the sake of analysis.

A. Some Insights

The goal of the present section is to characterize the asymptotic behaviour of $\bar{\mathfrak{L}}_N$ as $N, K \to \infty$, under hypothesis H_0. In order to have some insights on this behaviour, assume for the sake of illustration that $P = 1$ (at most one source is likely to be active).

Case $P = 1$. From Proposition 1, the (rescaled) generalized log-likelihood ratio $\bar{\mathfrak{L}}_N$ is a function of the ratio $\mu_1 = \lambda_1 / (\frac{1}{K}\mathrm{tr}\hat{\mathbf{R}})$. The asymptotic analysis of $\bar{\mathfrak{L}}_N$ thus reduces to the separate study of λ_1 and $\frac{1}{K}\mathrm{tr}\,\hat{\mathbf{R}}$. First consider the largest eigenvalue λ_1 of $\hat{\mathbf{R}}$. Under hypothesis H_0, $\hat{\mathbf{R}}$ belongs to the Laguerre Unitary Ensemble (LUE). It is well known that λ_1 converges a.s. to the right edge of the Marchenko-Pastur distribution: $\lambda_1 \xrightarrow{a.s.} (1+\sqrt{c})^2$. A further result due to Johnstone(2001) [5] states that convergence holds at speed $1/N^{2/3}$ and, more precisely, that the centered and rescaled quantity

$$\ell_1 = N^{2/3}\left(\frac{\lambda_1 - (1+\sqrt{c})^2}{(1+\sqrt{c})\left(\frac{1}{\sqrt{c}}+1\right)^{1/3}}\right) \tag{16}$$

converges in distribution toward a standard Tracy-Widom distribution function F_1 which can be defined in the following way:

$$F_1(s) = \exp\left(-\int_s^\infty (x-s)q^2(x)\,dx\right)\ , \tag{17}$$

where q solves the Painlevé II differential equation:

$$q''(x) = xq(x) + 2q^3(x)\ ,$$
$$q(x) \sim \mathrm{Ai}(x) \quad \text{as} \quad x \to \infty$$

and $\mathrm{Ai}(x)$ denotes the Airy function. This result provides the asymptotic behaviour of the numerator λ_1 of μ_1.

Now consider the denominator $\frac{1}{K}\mathrm{tr}\,\hat{\mathbf{R}}$ of μ_1. By the law of large numbers, $\frac{1}{K}\mathrm{tr}\,\hat{\mathbf{R}}$ converges a.s. to $\sigma^2 = 1$. Furthermore, convergence holds at speed $1/N$ in the sense that $\frac{1}{K}\mathrm{tr}\,\hat{\mathbf{R}} = 1 + O_P(1/N)$ (where $O_P(1/N)$ stands for a term which is bounded in probability by C/N for a certain constant C). It is therefore straightforward to prove that ratio $\mu_1 = \lambda_1/(\frac{1}{K}\mathrm{tr}\,\hat{\mathbf{R}})$ has the same asymptotic behaviour as λ_1. As a consequence, the asymptotic behaviour of $\bar{\mathfrak{L}}_N$ can be expressed in terms of the Tracy-Widom law (17).

Case $P > 1$. When the number P of sources is larger than one, a similar behaviour occurs. In that case, the test statistics

$\bar{\mathcal{L}}_N$ is a continuous function of $(\mu_1, \ldots, \mu_P)$ where for each $p = 1 \ldots P$, $\mu_p = \lambda_p / (\frac{1}{K}\mathrm{tr}\,\hat{\mathbf{R}})$. Due to the same arguments, $(\mu_1, \ldots, \mu_P)$ has essentially the same asymptotic behaviour as $(\lambda_1, \ldots, \lambda_P)$. Therefore, the asymptotic distribution of $\bar{\mathcal{L}}_N$ can be expressed in terms of the asymptotic joint distribution of the P largest eigenvalues $(\lambda_1, \ldots, \lambda_P)$ in the LUE. For each p, we define ℓ_p as the r.h.s. of equation (16) when λ_1 is replaced with λ_p. Random variables $(\ell_1, \ldots, \ell_P)$ are the properly centered and rescaled largest eigenvalues of $\hat{\mathbf{R}}$. The following result can be found for instance in [6]. Denote by $\mathbf{1}_A$ the indicator of set A.

Theorem 1. *[6] For each $x_1 \geq x_2 \cdots \geq x_P$,*

$$\lim_{N,K\to\infty} \mathbb{P}_{H_0}\left[\ell_1 \leq x_1, \ldots, \ell_P \leq x_P\right] = \mathcal{F}_P(x_1, \ldots, x_P)$$

where $\mathcal{F}_P(x_1, \ldots, x_P)$ is the P-variate Tracy-Widom distribution defined by

$$\mathcal{F}_P(x_1, \ldots, x_P) = \sum_{(i_1 \ldots i_P)\in I} \frac{1}{i_1! \ldots i_P!} \frac{\partial^{i_1+\cdots+i_P}}{\partial z_1^{i_1} \ldots \partial z_P^{i_P}} Det(1+K\chi_{z_1\ldots z_P})\big|_{z_1=\cdots=z_P=1}$$

where I consists of all sets of P nonnegative integers $i_1, \ldots, i_P$ such that $i_1 = 0$ and $i_1 + \cdots + i_{j+1} \leq j$ for each $j = 1 \ldots P-1$. Here, $Det(1+K\chi_{z_1\ldots z_P})$ represents the Fredholm determinant with kernel $K(x,y)\sum_{p=1}^P (z_p - 1)\mathbf{1}_{R_p}(y)$ where

$$K(x,y) = \frac{Ai(x)Ai'(y) - Ai(y)Ai'(x)}{x-y}$$

is the Airy kernel and where $R_1 \ldots R_p$ are the intervals (x_1, ∞), $(x_2, x_1]$,..., $(x_P, x_{P-1}]$ respectively.

Note that numerically efficient methods to compute Fredholm determinants and Tracy-Widom distributions have been studied, see for instance the works of [7]. Algorithms for constructing tables for $\mathcal{F}_P$ are however far beyond the scope of this paper. We may now express the main result of the present section.

B. Main Result

Let $\xrightarrow{\mathcal{D}}$ denote the convergence in distribution in the asymptotic regime (15). A sketch of the proof of the following result is provided in subsection IV-C.

Theorem 2. *Under hypothesis H_0,*

$$\bar{\mathcal{L}}_N \xrightarrow{\mathcal{D}} \sum_{p=1}^{P} X_p\,, \tag{18}$$

where $(X_1, \ldots, X_P)$ follows a standard P-variate Tracy-Widom distribution.

Denote by $F_P(x)$ the distribution function of $\sum_p X_p$ and by F_P^{-1} the inverse of F_P w.r.t. composition.

Corollary 1. *Any threeshold γ_N such that*

$$\gamma_N > F_P^{-1}(1-\epsilon) \tag{19}$$

ensures that the probability of false alarm $\mathbb{P}_{H_0}[\bar{\mathcal{L}}_N < \gamma_N]$ is no larger than ϵ for N large enough.

We now make the following comments.

• The above results provide a simple way to set the threshold γ_N or to compute p-values associated with the proposed test. It prevents from using tedious algorithms for approximating the distribution of $\bar{\mathcal{L}}_N$. Instead, it only relies on *pre-determined* tables of the distribution F_P. Such tables are well known in case $P = 1$. The case $P > 1$ has been subject to much less investigations at the present time. We note however that the main developments in the field of multivariate Tracy-Widom distributions are most recent (see for instance [6], [7] and references therein).

• Note that F_P does not depend on the technical parameters K, N or K/N. This observation is one of the main arguments for using test statistic $\bar{\mathcal{L}}_N$: the threshold selection procedure only depends on the desired probability of false alarm and on the maximum number of sources P likely to be present in a given band. It does not depend on the number of available snapshots. More importantly, it does not depend on the number of secondary users in the system. Such a flexibility of the test represents a particularly important feature for cognitive radio systems.

• The selection of γ_N as in (19) ensures that the PFA is below ϵ at least from a certain value of N. However, Theorem 2 provides no information on *how large* should be N in order that the PFA stays below ϵ for finite values of N, K. Such a characterization would require to study more accurately the convergence speed in (18), and is out of the scope of this paper. However, some answers are provided in the simulations.

C. Sketch of the Proof of Theorem 2

Assume that hypothesis H_0 holds. We study the asymptotic behavior of $N^{2/3}(\mathcal{L}_N - \alpha_N)$, where $\mathcal{L}_N$ is defined by (9) and α_N is defined by (11). In the asymptotic regime (15), α_N converges to the constant $\alpha = 2P\log(1+\sqrt{c}) - P(1+\sqrt{c})^2$. The asymptotic study of $N^{2/3}(\mathcal{L}_N - \alpha_N)$ is thus equivalent to the study of $N^{2/3}(\mathcal{L}_N - \alpha)$. Define the following quantities:

$$Z_p = \frac{\mu_p}{(1+\sqrt{c})^2} - 1\,, \tag{20}$$

$$A_N = \frac{\log\left(1 - \frac{1}{K}\sum_p \mu_p\right)}{-\frac{1}{K}\sum_p \mu_p}\left(1 - \frac{P}{K}\right)\,.$$

Finally, define $B_N = (1+\sqrt{c})^2 P N^{2/3}(1-A_N)$.

Lemma 1. *The following equality holds true:*

$$N^{2/3}(\mathcal{L}_N - P\,\alpha_c) = \sum_{p=1}^{P} N^{2/3} Z_p\left(\frac{\log(1+Z_p)}{Z_p} - A_N(1+\sqrt{c})^2\right) + B_N\,. \tag{21}$$

The proof of Lemma 1 is obtained by straightforward expansion of the r.h.s. of (21). As μ_p converges a.s. to $(1+\sqrt{c})^2$, Z_p tends to zero as $N \to \infty$. This implies that the ratio $\frac{\log(1+Z_p)}{Z_p}$ converges a.s. to one. Similarly, $\frac{1}{K}\sum_p \mu_p$ tends to zero as $N, K \to \infty$. Thus, A_N converges a.s. to one. Moreover, it can be shown that A_N converges to one at speed $1/N$ *i.e.*, $N(A_N - 1)$ is bounded in probability as $N \to \infty$ (the proof is omitted due to the lack of space). As a consequence, $B_N \propto N^{2/3}(1 - A_N)$ converges in probability to zero. On the otherhand, the facts that $Z_p \to 0$ and $A_N \to 1$ a.s. imply that

$$\forall p, \quad \frac{\log(1+Z_p)}{Z_p} - A_N(1+\sqrt{c})^2 \xrightarrow{a.s.} 1 - (1+\sqrt{c})^2 .$$

By Theorem 1, $N^{2/3} Z_p$ is bounded in probability for each p. Therefore, using $1-(1+\sqrt{c})^2 = -\sqrt{c}(2+\sqrt{c})$, equation (21) leads to:

$$N^{2/3}\left(\mathfrak{L}_N - P\,\alpha_c\right) = -\sqrt{c}(2+\sqrt{c})\sum_{p=1}^{P} N^{2/3} Z_p + o_P(1)\,,$$

$$= -\frac{c^{1/3}(2+\sqrt{c})}{(1+\sqrt{c})^{2/3}}\sum_{p=1}^{P} N^{2/3}\left[\frac{\mu_p - (1+\sqrt{c})^2}{(1+\sqrt{c})\left(\frac{1}{\sqrt{c}}+1\right)^{1/3}}\right] + o_P(1)$$

where $o_P(1)$ stands for a term which tends to zero in probability. As explained in subsection IV-A, the asymptotic distribution of the trace-normalized eigenvalues $\mu_1, \ldots, \mu_P$ is equivalent to the asymptotic distribution of the non-normalized eigenvalues $\lambda_1, \ldots, \lambda_P$, and is therefore given by Theorem 1. We finally obtain:

$$N^{2/3}\left(\mathfrak{L}_N - P\,\alpha_c\right) \xrightarrow{\mathcal{D}} -\frac{c^{1/3}(2+\sqrt{c})}{(1+\sqrt{c})^{2/3}}\sum_{p=1}^{P} X_p$$

where $(X_1, \ldots, X_P)$ follows a standard P-variate Tracy-Widom distribution. This proves Theorem 2.

D. A remark on the convergence speed

One may question the reason for introducing the term

$$P\left(1+\sqrt{c}\right)^2 \frac{P}{cN} - \frac{P^2(1+\sqrt{c})^4}{2Nc} \tag{22}$$

into the definition of α_N in equation (11). Indeed, since this term is negligible in the asymptotic regime, Theorem 2 would still hold when simply replacing α_N with its limit $\alpha = 2P\log(1+\sqrt{c}) - P(1+\sqrt{c})^2$. Unfortunately, in this case, the convergence speed of $\bar{\mathfrak{L}}_N$ (in distribution) would hold at rate $1/N^{1/3}$, which is a rather slow convergence rate. This unfortunate behaviour is due to the presence of the term B_N in (21), which tends to zero only at rate $1/N^{1/3}$. The correction term (22) allows to circumvent this issue, and to provide a higher convergence rate of $\bar{\mathfrak{L}}_N$. This term has no impact on the results, but allows the asymptotic regime to be reached even for moderate values of N, K, as illustrated by the following simulations.

V. Simulations

Figure 1 represents the empirical c.d.f. of the test statistic $\bar{\mathfrak{L}}_N$ for different values of the number of sources P. 1000 realizations of matrix $\mathbf{Y}$ are used to evaluate each empirical

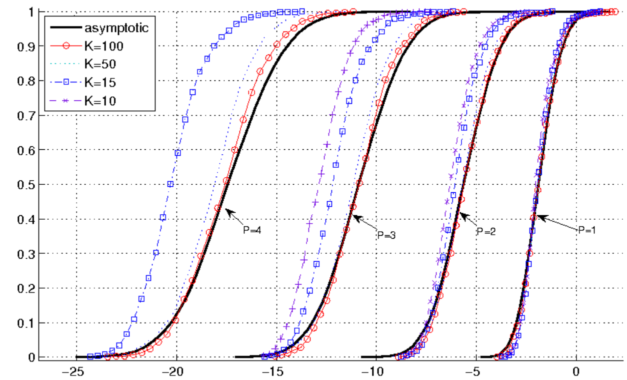

Figure 1. Cumulative distribution function of $\bar{\mathfrak{L}}_N$ – $P = 1, 2, 3, 4$ – $c = 0.4$

c.d.f. An arbitrary value of c ($c = 0.4$) is chosen. Solid bold lines represent the asymptotic distribution of $\bar{\mathfrak{L}}_N$ for each P, given by Theorem 2, which is independent of c. Figure 1 shows that, even for very moderate values of K, N, the empirical distribution function fits the asymptotic one. Typically, $K = 10$ is sufficient to achieve the asymptotic regime when $P = 1$ (although, as expected, when P increases, largest values of K are required to reach the asymptotic regime). This proves that our asymptotic analysis can reliably be used for the selection of the threeshold γ_N or for the computation of p-values, even for moderate values of N, K.

References

[1] J. Mitola. *Cognitive Radio An Integrated Agent Architecture for Software Defined Radio*. PhD thesis, Royal Institute of Technology (KTH), May 2000.

[2] F. F. Digham, M. S. Alouini, and M. K. Simon. On the energy detection of unknown signals over fading channels. 2003.

[3] V. I. Kostylev. Energy detection of a signal with random amplitude. 2002.

[4] H. Urkowitz. Energy detection of unknown deterministic signals. *Proceedings of the IEEE*, 55:523–531, 1967.

[5] I. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Annals of Stat.*, pages 295–327, 2001.

[6] Alexei Onatski. The tracy–widom limit for the largest eigenvalues of singular complex wishart matrices. *ANNALS OF APPLIED PROBABILITY*, 18:470, 2008.

[7] F. Bornemann. On the numerical evaluation of distributions in random matrix theory. *submitted to Markov Process. Related Fields*, (E-print arXiv:0904.1581), April 2009.