APPLICATIONS OF LARGE EMPIRICAL SPATIO-TEMPORAL COVARIANCE MATRIX IN MULTIPATH CHANNELS DETECTION

G.T. Pham, P. Loubaton

Université Paris-Est Marne-la-Vallée
LIGM, UMR CNRS 8049
5 Boulevard Descartes
77454 Marne la Vallee Cedex 2, France

ABSTRACT

This paper addresses the detection of a single signal in a multipath propagation channel using a sensors array in the case where the number of sensors M and the number of observations N are large and of the same order of magnitude and where the number of paths P is much smaller than M and N. In contrast with the single path context, the GLRT test cannot be implemented, and we evaluate the behaviour of tests based on the largest eigenvalues of the empirical spatio-temporal covariance matrix. Using a technical result showing that the largest singular values of low rank deterministic pertubation of certain Gaussian block-Hankel large random matrices behave as if the entries of the latter random matrices were independent identically distributed, we obtain a clear understanding of the advantages of the use of the spatial-temporal covariance matrix.

1. INTRODUCTION

The multi-antenna detection of low rank non observable narrow band signals corrupted by an additive spatially and temporally white Gaussian noise is a fundamental problem that was studied extensively in the contexts of array processing (see e.g. [6], [4]) and more recently of spectrum sensing (see among others [17], [14], [7]). The most popular method to solve the above problem is the GLRT test (see e.g [13]), which, in the present case, can be expressed in closed form. In order to obtain some insights on the statistical performance of the GLRT test, it is standard to assume that the number of observations N converge towards ∞, and to characterize the asymptotic distribution of the GLRT statistics under the null and the alternative hypothesis. In practice, this approach provides reasonable results when N is much larger than the number of antennas M at the receiver side. When the antenna array is large, the assumption that $N >> M$ is often not justified and the standard asymptotic analysis does not provide reliable results (see e.g. [8] in the context of supervised detection). In this context, it is now standard to consider the large system regime $M \to +\infty$, $N \to +\infty$ in such a way that $\frac{M}{N} \to c$ where $c > 0$. We refer the reader to the papers [1], [16], [12], [5] in which this approach is developed.

In this paper, we assume that M and N are large and of the same order of magnitude. We address the detection of a single signal in a multipath propagation channel, i.e. its contribution to the observation coincides with the output of an unknown finite impulse response SIMO filter driven by an unobservable deterministic scalar sequence $s = (s_n)_{n \in \mathbb{Z}}$. The signal to be detected is thus a rank 1 wideband signal. We assume moreover that the number of paths P, or equivalently the number of coefficients of the SIMO filter is much smaller than M. [17] studied the GLRT test when s is an i.i.d. Gaussian sequence and the filter has an infinite impulse response, or equivalently when $P = +\infty$. Under certain assumptions, [17] proposed to evaluate the log likelihood ratio using the Witthle approximation, and obtained an expression based on integrals over the frequency domain. When P is finite, the GLRT test cannot be expressed in closed form because the maximization of the likelihood over sequence s and the filter coefficients $(\mathbf{h}_p)_{p=0,\dots,P-1}$ has no explicit solution.

As the GLRT test cannot be used, a pragmatic approach is to observe that the signal to be detected can be interpreted as a superposition of P narrow band deterministic signals. Therefore, it is possible to use the corresponding GLRT test which consists in comparing the sum of the P greatest eigenvalues of the empirical spatial covariance matrix of the observation to a threshold, at least if the noise variance is known. However, it is intuitively more appealing to consider the greatest eigenvalues of the empirical spatio-temporal covariance matrix in order to take benefit of the particular convolutive structure of the signal to be detected. We compare these 2 approaches in the asymptotic regime $M \to +\infty$, $N \to +\infty$ in such a way that $\frac{M}{N} \to c$ where $c > 0$. In this regime, the first order behaviour of the largest eigenvalues of the empirical spatial covariance matrix is well known, and this allows to evaluate the relevance of the "narrow band" test. In this paper, we use the recent result [15] in order to evaluate the behaviour of the greatest eigenvalues of the empirical spatio-temporal covariance matrix. This allows to have a clear understanding of the advantages of the use of the spatio-temporal covariance matrix.

This paper is organized as follows. In section 2, we introduce the signal models and the underlying assumptions. In section 3, we present our results concerning the behaviour of the greatest eigenvalues of the empirical spatio-temporal covariance matrix, and deduce from this in section 4 the first order behaviour of the detection test based on this matrix. Finally, section 5 presents numerical experiments sustaining our theoretical results.

In the following, $\mathcal{N}_c(\mathbf{x}, \mathbf{\Gamma})$ represents the M-variate complex Gaussian (i.e. circular) distribution with mean $\mathbf{x}$ and covariance matrix $\mathbf{\Gamma}$

2. PROBLEM FORMULATION.

In the following, we denote by $(\mathbf{y}_n)_{n=1,\dots,N}$ the M-dimensional signal received on the M-sensors array. Under hypothesis H_0, the observation is reduced to a spatially and temporally complex Gaussian noise, i.e.

$$\mathbf{y}_n = \mathbf{v}_n, \; n = 1, \dots, N \tag{1}$$

where $(\mathbf{v}_n)_{n=1,\dots,N}$ are i.i.d. $\mathcal{N}_c(0, \sigma^2 \mathbf{I})$ distributed random vectors. We assume from now on that σ^2 is known in order to simplify

Work supported by Project ANR-12-MONU-OOO3 DIONISOS

the exposition, but our results can be easily generalized if σ^2 is unknown (see below). Under hypothesis H_1, the observation is given by

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p} + \mathbf{v}_n,\ n = 1, \ldots, N \tag{2}$$

where $(s_n)_{n\in\mathbb{Z}}$ is a non observable deterministic scalar sequence and where the $M \times 1$ transfer function $\mathbf{h}(z) = \sum_{p=0}^{P-1} \mathbf{h}_p z^{-p}$ is unknown. We however assume that P is known, which, in practice, means that an upper bound of the number of paths is available. s is assumed deterministic in order to avoid formulating restrictive hypotheses,e.g. that signal $(s_n)_{n\in\mathbb{Z}}$ is an i.i.d. Gaussian sequence.

In order to test hypothesis H_0 versus H_1, the GLRT test cannot be implemented because, under H_1, the maximum likelihood estimator of filter $\mathbf{h}(z)$ and sequence $(s_n)_{n=-(P-1),\ldots,N}$ cannot be expressed in closed form (see e.g. [18], [11]). We note that when s is an i.i.d. Gaussian sequence, [17] derived an approximate GLRT test based on the Whittle approximation, but without assuming that filter $\mathbf{h}(z)$ is FIR. Moreover, the approach of [17] needs the observation of at least M independent realizations of the observation $(\mathbf{y}_n)_{n=1,\ldots,N}$, an hypothesis which is not formulated in the present paper. Finally, the approach of [17] cannot be adapted to the case of a deterministic signal $(s_n)_{n\in\mathbb{Z}}$.

As the GLRT test cannot be implemented, we study pragmatic alternative approaches. The most obvious solution is based on the observation that signal $[\mathbf{h}(z)]s(n) = \sum_{p=0}^{P-1} \mathbf{h}_p s_{n-p}$ can be interpreted as a superposition of P narrow band signals. It is thus possible to test the hypothesis H_0 against hypothesis $\mathrm{H}_1^{'}$ defined by

$$\mathbf{y}_n = \sum_{p=0}^{P-1} \mathbf{h}_p s_n^{(p)} + \mathbf{v}_n \tag{3}$$

where signals $(s^{(p)})_{p=0,\ldots,P-1}$ are non observable deterministic signals. Hypothesis $\mathrm{H}_1^{'}$ is of course not equivalent to H_1 because the particular structure of $s_n^{(p)} = s_{n-p}$ is ignored in the formulation of $\mathrm{H}_1^{'}$. We denote by $\mathbf{Y}$ the $M \times N$ matrix defined by $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$. Then, $\mathrm{H}_1^{'}$ is of course equivalent to

$$\mathbf{Y} = \mathbf{A} + \mathbf{V} \tag{4}$$

where $\mathbf{A}$ is a rank P deterministic matrix, and where $\mathbf{V}$ is defined as $\mathbf{Y}$. The corresponding GLRT test is easy to derive, and consists in comparing the statistics

$$\eta_N = \sum_{i=1}^{P} \lambda_i \left(\frac{\mathbf{Y}\mathbf{Y}^*}{N} \right) \tag{5}$$

to a threshold. Here, $\left(\lambda_i\left(\frac{\mathbf{Y}\mathbf{Y}^*}{N}\right)\right)_{i=1,\ldots,M}$ represent the eigenvalues of $\frac{\mathbf{Y}\mathbf{Y}^*}{N}$ arranged in decreasing order.

Matrix $\frac{\mathbf{Y}\mathbf{Y}^*}{N}$ coincides with the empirical spatial covariance matrix of the observations. In order to take benefit of the particular convolutive structure of signal $[\mathbf{h}(z)]s_n$, it seems however more appropriate to consider a statistics based on the largest eigenvalues of empirical spatio-temporal covariance matrices. If L is an integer, we denote by $\mathbf{y}_n^{(L)}$ the ML–dimensional vector defined by

$$\mathbf{y}_n^{(L)} = \left(\mathbf{y}_{1,n}, \ldots, \mathbf{y}_{1,n+L-1}, \ldots, \mathbf{y}_{M,n}, \ldots, \mathbf{y}_{M,n+L-1}\right)^T$$

and by $\mathbf{Y}^{(L)}$ the $ML \times N$ block-Hankel matrix defined by $\mathbf{Y}^{(L)} = (\mathbf{y}_1^{(L)}, \ldots, \mathbf{y}_N^{(L)})$. Thus $\frac{\mathbf{Y}^{(L)}(\mathbf{Y}^{(L)})^*}{N}$ represents the spatio-temporal covariance matrix. We note that $\mathbf{Y}^{(L)}$ depends on $(\mathbf{y}_n)_{n=1,\ldots,N+L-1}$ while, in principle, the observation $\mathbf{y}_n$ is available until $n = N$. As we consider in the following asymptotic regimes in which $N \to +\infty$ while L remains fixed, the above mentioned end effect has no consequence on our results.

Under hypothesis H_0, matrix $\mathbf{Y}^{(L)}$ is reduced to $\mathbf{V}^{(L)}$, and under H_1, it holds that

$$\mathbf{Y}^{(L)} = \mathbf{H}^{(L)}\mathbf{S}^{(L)} + \mathbf{V}^{(L)} \tag{6}$$

where matrix $\mathbf{S}^{(L)}$ is the $(P + L - 1) \times N$ Hankel matrix defined by $(\mathbf{S}^{(L)})_{i,n} = s_{n+i-P}$, and where $\mathbf{H}^{(L)}$ is defined by $\mathbf{H}^{(L)} = (\mathbf{H}_1^{(L)T}, \ldots, \mathbf{H}_M^{(L)T})^T$ with $\mathbf{H}_m^{(L)}$ representing the $L \times (P+L-1)$ Toeplitz matrix corresponding to the convolution of signal $(s_n)_{n\in\mathbb{Z}}$ with sequence $(\mathbf{h}_{m,p})_{p=0,\ldots,P-1}$. We remark that matrices $\mathbf{Y}^{(1)}$ and $\mathbf{V}^{(1)}$ coincide with $\mathbf{Y}$ and $\mathbf{V}$. We also denote matrices $\mathbf{S}^{(1)}$ and $\mathbf{H}^{(1)}$ by $\mathbf{S}$ and $\mathbf{H}$. Instead of using η_N defined by (5), we propose to consider the statistics $\eta_N^{(L)}$ given by

$$\eta_N^{(L)} = \sum_{i=1}^{P+L-1} \lambda_i \left(\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}}{N} \right) \tag{7}$$

for a suitable value of integer L. We note that if $L = 1$, then $\mathbf{Y}^{(1)}$ and $\eta_N^{(1)}$ coincide with $\mathbf{Y}$ and η_N.

In order to obtain some insights on the merits of statistics $\eta_N^{(L)}$ in the case where M and N are large and of the same order of magnitude, we evaluate under both hypotheses the first order behaviour of $\eta_N^{(L)}$ in the asymptotic regime $M \to +\infty$, $N \to +\infty$ in such a way that $c_N = \frac{M}{N} \to c$ where $c > 0$. We also assume that P and L do not scale with M, N. In the following, $N \to +\infty$ should be understood as the above asymptotic regime. The study of $\eta_N^{(L)}$ when $N \to +\infty$ is equivalent to the study of the largest eigenvalues of matrix $\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}}{N}$. As we shall see below, the case $L = 1$ is well known, and followed from existing results (see e.g. [3]) concerning the behaviour of the largest singluar values of finite rank perturbation of the i.i.d. large random matrix $\mathbf{V}$. If $L > 1$, matrix $\mathbf{V}^{(L)}$ is block-Hankel, and its entries are of course not i.i.d. Our main result shows that the largest eigenvalues of $\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}}{N}$ behave as if the MLN entries of $\mathbf{V}^{(L)}$ were i.i.d. This behaviour appears as a consequence of the results of [15].

Remark 1: **The case σ^2 unknown.** When σ^2 is unknown, the GLRT test corresponding to hypotheses H_0 and $\mathrm{H}_1^{'}$ given by (4) consists in comparing statistics $\frac{\eta_N}{\frac{1}{M}\mathrm{Tr}(\mathbf{Y}\mathbf{Y}^*/N)}$ to a threshold. Therefore, it is relevant to replace statistics $\eta_N^{(L)}$ by $\frac{\eta_N^{(L)}}{\frac{1}{M}\mathrm{Tr}(\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}/N)}$. It is easily seen that $\frac{1}{M}\mathrm{Tr}(\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}/N)$ converges almost surely towards σ^2 in the absence and in the presence of signal. Therefore, the characterization of first order asymptotic behaviours of $\eta_N^{(L)}$ and of its normalized version are equivalent.

3. LARGEST EIGENVALUES OF THE EMPIRICAL SPATIO-TEMPORAL COVARIANCE MATRIX.

In order to simplify the notations, we denote by $\mathbf{W}_N^{(L)}$, $\mathbf{B}_N^{(L)}$ and $\mathbf{X}_N^{(L)}$ the matrices defined by $\mathbf{W}_N^{(L)} = \mathbf{V}_N^{(L)}/\sqrt{N}$, $\mathbf{B}_N^{(L)} = \frac{1}{\sqrt{N}}\mathbf{H}^{(L)}\mathbf{S}^{(L)}$ and $\mathbf{X}_N^{(L)} = \mathbf{B}_N^{(L)} + \mathbf{W}_N^{(L)}$. We notice that $\mathbf{Y}^{(L)}/\sqrt{N}$ coincides with $\mathbf{W}_N^{(L)}$ under H_0 and with $\mathbf{X}_N^{(L)}$ under H_1. This paper is based on a technical result which establishes that, in a certain sense, the eigenvalues of matrix $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ behave as if the entries of $\mathbf{W}_N^{(L)}$ were i.i.d. In order to state the corresponding result, we recall that the Marcenko-Pastur distribution μ_d with parameters (σ^2, d) is the probability distribution defined by

$$d\mu(x) = \delta_0[1-d^{-1}]_+ + \frac{\sqrt{(x-x^-)(x^+-x)}}{2\sigma^2 d\pi x}\mathbb{1}_{[x_d^-,x_d^+]}(x)\,dx$$

with $x_d^- = \sigma^2(1-\sqrt{d})^2$ and $x_d^+ = \sigma^2(1+\sqrt{d})^2$. We denote by $m_d(z)$ its Stieltjes transform defined by $m_d(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z}$ and by $\tilde{m}_d(z)$ the function $\tilde{m}_d(z) = dm(z) - (1-d)/z$. We denote by $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ the so-called resolvent of matrices $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ and $\mathbf{W}_N^{(L)*}\mathbf{W}_N^{(L)}$ defined by $\mathbf{Q}_N(z) = \left(\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*} - z\mathbf{I}_{M-L+1}\right)^{-1}$,$\tilde{\mathbf{Q}}_N(z) = \left(\mathbf{W}_N^{(L)*}\mathbf{W}_N^{(L)} - z\mathbf{I}_{NL}\right)^{-1}$ Then, when $N \to +\infty$, the following result holds.

Proposition 1: The eigenvalue distribution of matrix $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ converges almost surely towards the Marcenko-Pastur distribution μ_{cL}. Moreover, for each fixed integer k, the k largest eigenvalues $(\lambda_l(\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}))_{l\leq k}$ verify

$$\lambda_l(\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}) \to x_{+,cL} = \sigma^2(1+\sqrt{cL})^2 \ a.s. \quad (8)$$

and the k smallest non zero eigenvalues of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ converge almost surely towards $x_{-,cL}$. If $\mathbf{a}_N, \mathbf{b}_N$ are 2 unit norm ML–dimensional deterministic vectors, then it holds that for each $z \in \mathbb{C}^+$

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - m_{cL}(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \ a.s. \quad (9)$$

Similarly, if $\tilde{\mathbf{a}}_N, \tilde{\mathbf{b}}_N$ are 2 unit norm N–dimensional deterministic vectors, then for each $z \in \mathbb{C}^+$, it holds that

$$\tilde{\mathbf{a}}_N^* \left(\tilde{\mathbf{Q}}_N(z) - \tilde{m}_{cl}(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0 \ a.s. \quad (10)$$

Moreover, for each $z \in \mathbb{C}^+$, it holds that

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z)\mathbf{W}_N^{(L)}\right)\tilde{\mathbf{b}}_N \to 0 \ a.s. \quad (11)$$

Finally, for each $\epsilon > 0$, convergence properties (9, 10, 11) hold uniformly w.r.t. z on each compact subset of $\mathbb{C} - [0, x^+ + \epsilon]$.

The proof of Proposition 1 can be found in [[9], Proposition 1]. We recall that, roughly speaking, the convergence of the eigenvalue distribution of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ towards distribution μ_{cL} means that the histogramms of the eigenvalues of any realization of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ tend to accumulate around the graph of the probability density of μ_{cL}.

The statements of Proposition 1 are well known when $L = 1$. Apart (10) and (11), Proposition 1 appears as a consequence of the results of [15]. We note that [15] is used here when L does not scale with (M, N). However, the results of [15], and thus Proposition 1, are still valid if L and M both converge towards infinity in such a way that $ML/N \to d$ with $d > 0$ and that $L = \mathcal{O}(N^\alpha)$ with $\alpha < 2/3$. In this case, parameter cL in Proposition 1 should be replaced by d. If $2/3 \leq \alpha < 1$, the eigenvalue distribution of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ still converges towards μ_d, but the almost sure convergence of the largest and smallest eigenvalues of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ towards $x_{+,d}$ and $x_{-,d}$ is not guaranteed. We finally note that if N and L converge toward ∞ at the same rate and that M remains fixed, the convergence of the eigenvalue distribution $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ towards μ_d is no longer true. Intuitively, this is because $\mathbf{W}_N^{(L)}$ depends on MN independent random variables, and that if M is fixed, this number is not sufficient to ensure nice averaging effects. In particular, if $M = 1$, it is shown in [2] that the eigenvalue distribution of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ converges towards an unbounded probability distribution that can be characterized by its moments.

We finally remark that the almost sure convergence of the extreme eigenvalues towards $x_{+,cL} = \sigma^2(1+\sqrt{cL})^2$ and $x_{-,cL} = \sigma^2(1-\sqrt{cL})^2$ also implies that the non zero eigenvalues of $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$ are almost surely located in a neighborhood of $[\sigma^2(1-\sqrt{cL})^2, \sigma^2(1+\sqrt{cL})^2]$ when N increases. Proposition 1 thus allows to have a clear understanding of the effect of L on the spreading of the eigenvalues $\mathbf{W}_N^{(L)}\mathbf{W}_N^{(L)*}$.

In the following, we denote by $(\hat{\lambda}_{k,N}^{(L)})_{k=1,\ldots,ML}$ the eigenvalues of $\mathbf{X}_N^{(L)}\mathbf{X}_N^{(L)*}$, and by $\lambda_{1,N}^{(L)} \geq \lambda_{2,N}^{(L)} \ldots \geq \lambda_{P+L-1,N}^{(L)}$ the non zero eigenvalues of $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$. Proposition 1 allows to generalize immediately the approach used in [3], and to prove that the $P+L-1$ greatest eigenvalues of $\mathbf{X}_N^{(L)}\mathbf{X}_N^{(L)*}$ also behave as if the entries of $\mathbf{W}_N^{(L)}$ were i.i.d.

Theorem 1: We assume that:

Assumption 1: The $P+L-1$ non zero eigenvalues $(\lambda_{k,N}^{(L)})_{k=1,\ldots,K}$ of matrix $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ converge towards $\lambda_1^{(L)} \geq \lambda_2^{(L)} \geq \ldots \geq \lambda_{P+L-1}^{(L)}$ when $N \to +\infty$.

We denote by K_L, $0 \leq K_L \leq P+L-1$, the largest integer for which $\lambda_{K_L}^{(L)} > \sigma^2\sqrt{cL}$. Then, for $k = 1, \ldots, K_L$, it holds that

$$\hat{\lambda}_{k,N}^{(L)} \xrightarrow[N\to\infty]{a.s.} \rho_k^{(L)} = \phi_L(\lambda_k^{(L)}) = \frac{(\lambda_k^{(L)} + \sigma^2)(\lambda_k^{(L)} + \sigma^2 cL)}{\lambda_k^{(L)}} > x_{+,cL}.$$

while for $k = K_L+1, \ldots, P+L-1$, $\hat{\lambda}_{k,N}^{(L)} \to x_{+,cL}\ a.s.$

4. ASYMPTOTIC BEHAVIOUR OF $\eta_N^{(L)}$.

In order to simplify the following discussion, we formulate the following hypotheses on vectors $(\mathbf{h}_p)_{p=0,\ldots,P-1}$ and on signal $(s_n)_{n\in\mathbb{Z}}$:

Assumption 2:
- (i) When $N \to +\infty$, matrix $\mathbf{H}^*\mathbf{H}$ converges towards a $P \times P$ matrix $\mathbf{\Delta}$
- (ii) For each integers $i, j \geq 1$, $\frac{1}{N}\sum_{n=1}^N s_{n+i-P}s_{n+j-P}^*$ converges towards a limit. In this case, the limit only depends on $i-j$, and is denoted R_{i-j}.

As the entries of matrix $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}$ depend on the entries of $\mathbf{H}^*\mathbf{H}$, (i) implies that $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}$ converges towards a matrix $\mathbf{\Delta}^{(L)}$ whose entries depend on the entries of $\mathbf{\Delta}$. In the following, we also denote by $\mathbf{R}^{(L)}$ the $(P+L-1)\times(P+L-1)$ Toeplitz matrix defined by $\mathbf{R}_{i,j}^{(L)} = R_{i-j}$.

As the non zero eigenvalues of $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ coincide with the eigenvalues of matrix $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}\frac{\mathbf{S}^{(L)}\mathbf{S}^{(L)*}}{N}$, it is clear that Assumption 2 implies that Assumption 1 holds, and that $\lambda_k^{(L)} = \lambda_k\left(\mathbf{\Delta}^{(L)}\mathbf{R}^{(L)}\right)$. We also remark that for each $L \geq 1$, matrix $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ is a sub-matrix of $\mathbf{B}_N^{(L+1)}\mathbf{B}_N^{(L+1)*}$. Therefore (see [10]), it holds that $\lambda_{k,N}^{(L)} \leq \lambda_{k,N}^{(L+1)}$, and therefore that $\lambda_k^{(L)} \leq \lambda_k^{(L+1)}$ and $\rho_k^{(L)} \leq \rho_k^{(L+1)}$ for each $k = 1, \ldots, P+L-1$.

We now use Theorem 1 in order to evaluate the behaviour of $\eta_N^{(L)}$. It is clear that $\eta_N^{(L)}$ converges almost surely $\eta^{(L)}$ defined by

$$\eta^{(L)} = \sum_{k=1}^{K_L} \rho_k^{(L)} + (P+L-1-K_L)\sigma^2(1+\sqrt{cL})^2) \tag{12}$$

We note that if $K_L = 0$, or equivalently, if the largest eigenvalue $\lambda_{1,N}^{(L)}$ of matrix $\mathbf{B}_N^{(L)}\mathbf{B}_N^{(L)*}$ is below the detectability threshold $\sigma^2\sqrt{cL}$ for each N large enough, then the first order asymptotic behaviour of $\eta_N^{(L)}$ under hypotheses H_0 and H_1 coincide. In this case, the test based on $\eta_N^{(L)}$ is not consistent, in the sense that it does not allow to distinguish between the 2 hypotheses when $N \to +\infty$. If however $\lambda_{1,N}^{(L)}$ is greater $\sqrt{cL}$ for each N large enough, the asymptotic behaviours of $\eta_N^{(L)}$ under H_0 and H_1 do not coincide and the test is consistent. In other words, the test based on $\eta_N^{(L)}$ is consistent if and only if

$$\frac{\lambda_1^{(L)}}{\sigma^2\sqrt{cL}} > 1 \tag{13}$$

This condition implies that the value of L for which $\frac{\lambda_1^{(L)}}{\sqrt{cL}}$ is maximum can be considered as optimal from the consistency of the GLRT test point of view. In order to obtain some insights on the optimal choice of L, we first assume that $(s_n)_{n\in\mathbb{Z}}$ coincides with a realization of a unit variance zero mean i.i.d. sequence and that the limit $\mathbf{\Delta}$ of matrix $\mathbf{H}^*\mathbf{H}$ is diagonal, a condition meaning that the P paths are independent. In order to simplify the notations, we denote by $\delta_0, \ldots, \delta_{P-1}$ the diagonal entries of $\mathbf{\Delta}$ which represent the powers of the various paths. We notice that $\sum_{p=0}^{P-1}\delta_p$ coincides with the power of the signal to be detected. It is easily seen that for each L, matrix $\mathbf{\Delta}^{(L)}$ is diagonal as well, and that its largest entry is equal to $\sum_{p=0}^{P-1}\delta_p$ if $L \geq P$, and to $\max_{k=0,\ldots,P-L}\sum_{p=0}^{L-1}\delta_{p+k}$ if $L \leq P$. As matrix $\mathbf{R}^{(L)}$ is equal to $\mathbf{I}_{P-L+1}$, this implies that the largest limit eigenvalue $\lambda_1^{(L)}$ is equal to $\sum_{p=0}^{P-1}\delta_p$ if $L \geq P$, and to $\max_{k=0,\ldots,P-L}\sum_{p=0}^{L-1}\delta_{p+k}$ if $L \leq P$. If $L \geq P$, the left handside of (13) is equal to $\frac{\sum_{p=0}^{P-1}\delta_p}{\sigma^2\sqrt{cL}}$ while it is equal to

$$\frac{\max_{k=0,\ldots,P-L}\sum_{p=0}^{L-1}\delta_{p+k}}{\sigma^2\sqrt{cL}}$$

if $L \leq P$. The optimal value of L of course depends on the particular values of $\delta_0, \ldots, \delta_{P-1}$. If the powers all coincide with δ, the optimal value is $L = P$, and the test based on $\eta_N^{(P)}$ is consistent if and only if $\frac{\delta}{\sigma^2}$ is greater than the detectability threshold $\sqrt{c/P}$. In this case, it is also seen that if $L > P$, then condition is $\frac{\delta}{\sigma^2} > \sqrt{c/P}\sqrt{L/P}$, i.e. the detectability threshold is multiplied by $\sqrt{L/P}$, while if $L = 1$, the detectability threshold is $\sqrt{c}$, and is thus multiplied by $\sqrt{P}$. If the channel is sparse, the $L = P$ may of course not be the optimal value.

We now consider a more realistic scenario in which matrix $\mathbf{\Delta}$ is not diagonal. We assume that the signal to be detected is a sampled version of a continuous time linearly modulated signal $\sum_n s_n g_a(t - nT)$ where $(s_n)_{n\in\mathbb{Z}}$ is an i.i.d. sequence of symbols and where $g_a(t)$ is the classical continuous time shaping filter. The propagation channel is a Rayleigh multipath channel with Q uncorrelated paths with time-delays $\tau_0, \ldots, \tau_{Q-1}$. In this context, vectors $(\mathbf{h}_p)_{p=0,\ldots,P-1}$ are given by

$$\mathbf{h}_p = \sum_{q=0}^{Q-1} \boldsymbol{\lambda}_q \, g_a(pT - \tau_q) \tag{14}$$

where vectors $(\boldsymbol{\lambda}_q)_{q=0,\ldots,Q-1}$ are the realizations of independent zero-mean random Gaussian vectors. We denote by $\mathbf{\Lambda}$ the $M \times Q$ matrix $\mathbf{\Lambda} = (\boldsymbol{\lambda}_0, \ldots, \boldsymbol{\lambda}_{Q-1})$, and assume that matrix $\mathbf{\Lambda}^*\mathbf{\Lambda}$ converges towards $\mu\mathbf{I}_Q$. In practice, this hypothesis means that the Q paths share the same power. As $\mathbf{H} = (\mathbf{h}_{P-1}, \ldots, \mathbf{h}_0)$ is given by $\mathbf{H} = \mathbf{\Lambda}\mathbf{G}$ where $\mathbf{G} = (\mathbf{g}_{P-1}, \ldots, \mathbf{g}_0)$ and where each Q–dimensional vector $\mathbf{g}_p$ is given by $\mathbf{g}_p = (g_a((P-1)T - \tau_0), \ldots, g_a((P-1)T - \tau_{Q-1}))^T$, it is clear that matrix $\mathbf{H}^*\mathbf{H}$ converges towards $\mathbf{\Delta} = \mu\mathbf{G}^*\mathbf{G}$, and that $\mathbf{H}^{(L)*}\mathbf{H}^{(L)}$ converges towards $\mathbf{\Delta}^{(L)} = \mu\mathbf{G}^{(L)*}\mathbf{G}^{(L)}$ where matrix $\mathbf{G}^{(L)}$ is the $QL \times (P+L-1)$ block-Toeplitz matrix with first block line $(\mathbf{g}_{P-1}, \ldots, \mathbf{g}_0, 0, \ldots, 0)$. Therefore, the largest eigenvalue $\lambda_1^{(L)}$ of $\mathbf{\Delta}^{(L)}$ is equal to $\mu\lambda_1(\mathbf{G}^{(L)*}\mathbf{G}^{(L)})$. The optimal value of L thus depends on the way the largest eigenvalue of $\mathbf{G}^{(L)*}\mathbf{G}^{(L)}$ increases with L. As the optimal value of L cannot be found using analytical arguments, we give a numerical example. We assume that $g_a(t)$ is a square root Nyquist filter with excess bandwidth 0.5 which is truncated to interval $[-2.5T, 2.5T]$. Moreover, $Q = 2$, $\tau_0 = 0$, $\tau_1 = 2T$, $c = 1/2$ and the SNR $\frac{\mu}{\sigma^2}$ is equal to 2 dB. In figure 1, we plot the largest eigenvalue of $\mathbf{G}^{(L)*}\mathbf{G}^{(L)}$ and the lefthandside of (13) versus L. It is seen that the optimal value of L is equal to 3, it is thus different from P, which, is the present context is equal to $P = 7$.

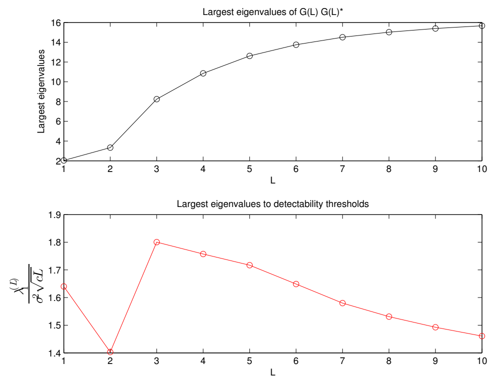

Fig. 1. Largest eigenvalues to detectability thresholds versus L

5. SIMULATION RESULTS.

In this section, we provide numerical simulations illustrating the results given in the previous sections. We first consider the case where matrix $\mathbf{H}$ coincides with a realization of Gaussian random matrix with i.i.d. $\mathcal{N}_{\mathbb{C}}(0, \mathbf{I}/(MP))$ entries. In this context, matrix $\mathbf{H}^*\mathbf{H}$ converges towards $\mathbf{I}_P$. Sequence $(s_n)_{n=2-P,\ldots,N-P+1}$ is a realization of an i.i.d. sequence taking values ± 1 with probability $1/2$. In this context, we have shown before that the optimal value of L is equal to P. In order to illustrate this behaviour, we consider the case $M = 80$, $N = 160$ and $P = 5$, and represent in figure 2 the ROC curves, evaluated using Monte-Carlo simulations, corresponding to the statistics $\eta_N^{(5)}, \eta_N, \lambda_1(\frac{\mathbf{Y}^{(5)}\mathbf{Y}^{(5)*}}{N})$ and $\lambda_1(\frac{\mathbf{Y}\mathbf{Y}*}{N})$, referred to as spatio-temporal, spatial, lmax-st and lmax-s in figure 2. The numerical results confirm that the use of $\eta_N^{(5)}$ leads to much better results than the use of η_N which corresponds to $L = 1$, and that it is indeed beneficial to take into account the $P + L - 1$ largest eigenvalues of the empirical spatio-temporal covariance matrix, and not only the largest one.

We now generate vectors $(\mathbf{h}_p)_{p=0,\ldots,P-1}$ according to model (14) for $Q = 2$, $\tau_0 = 0$, $\tau_1 = 2T$, $P = 7$ and when $g_a(t)$ is a square

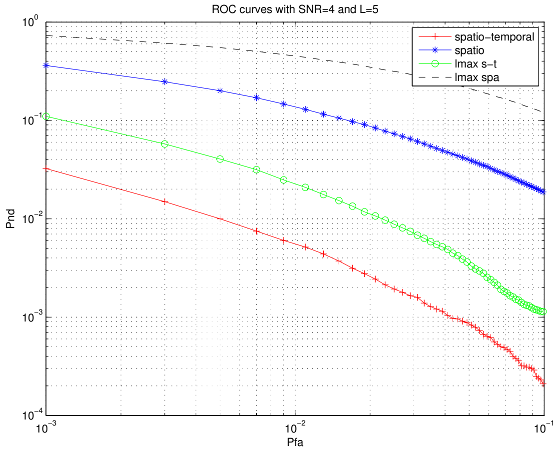

Fig. 2. ROC curves of different statistics,$\boldsymbol{\Delta}$ diagonal

root Nyquist filter with excess bandwidth 0.5 which is truncated to interval $[-2.5T, 2.5T]$. In figure 3, we assume that $M = 80, N = 160$ and again represent the ROC curves corresponding to the statistics $\eta_N^{(L)}, \eta_N, \lambda_1(\frac{\mathbf{Y}^{(L)}\mathbf{Y}^{(L))*}}{N})$ and $\lambda_1(\frac{\mathbf{Y}\mathbf{Y}*}{N})$ for $L = 3$. This time, it is seen that it is not beneficial to take into account the $L + P - 1$ largest eigenvalues of $\mathbf{Y}^{(L)}\mathbf{Y}^{(L)*}/N$, and that the best strategy is to consider the largest eigenvalue, which, for $L = 3$, provides the best results.

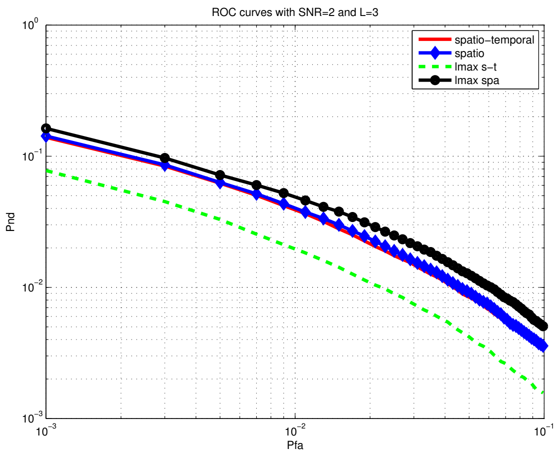

Fig. 3. ROC curves of different statistics, $\boldsymbol{\Delta}$ not diagonal

6. CONCLUSION

In this paper, we have studied the detection problem of a signal corrupted by an unknown multipath propagation channel using a sensor array in the case where the number of sensors M and the number of observations N are large and of the same order of magnitude. In this asymptotic regime, we have obtained a clear understading of the advantages of using statistics based on the largest eigenvalues of empirical spatio-temporal covariance matrices.

REFERENCES

[1] Z. Bai, J.W. Silverstein, ”Spectral analysis of large dimensional random matrices”, Springer Series in Statistics, 2nd ed., 2010.

[2] A. Basak, A. Bose, S. Sen, ”Limiting spectral distribution of sample autocovariance matrices”, to appear in *Bernouilli*, can be downloaded on Arxiv, arXiv:1108.3147v1.

[3] F. Benaych-Georges, R.R. Nadakuditi,”The singular values and vectors of low rank perturbations of large rectangular random matrices”, J. Multivariate Anal., Vol. 111 (2012), 120–135.

[4] O. Besson, S. Kraut, L.L. Sharf, ”Detection of an unknown rank-one component in white noise”, *IEEE Trans. on Signal Process.*, vol. 54, no. 7, pp. 2835-2839, Jul. 2006.

[5] P. Bianchi, M. Debbah, M. Maëda and J. Najim, ”Performance of Statistical Tests for Single Source Detection using Random Matrix Theory”, *IEEE Inf. Theory*, vol. 57, no.4, pp. 2400–2419, april 2011

[6] A. Leshem, A.J. Van der Veen, ”Multichannel detection of Gaussian signals with uncalibrated receivers”, *IEEE Signal Process. Letters*, vol. 8, pp. 120-122, 2001

[7] A. Taherpour, M. Naisri-Kenari, S. Gazor, ”Multiple antenna based spectrum sensing in cognitive radios, *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814-823, Feb. 2010.

[8] S. Hiltunen, P. Loubaton, P. Chevalier, "Asymptotic analysis of a GLRT test for detection with large sensor arrays", can be downloaded on Arxiv.

[9] G.T.Pham, P.Loubaton, P.Vallet, ”Performance analysis of spatial smoothing schemes in the context of large arrays”, arXiv:1503.08196.

[10] R. Horn, C.R. Johnson, ”Matrix Analysis”, 2nd. Ed., Cambridge Univ. Press, 2013.

[11] Y. Hua, ”Fast blind maximum likelihood for blind identification of multiple FIR channels”, *IEEE Trans. on Signal Process.*, vol. 44, no.3, pp. 661-672, Mar. 1996.

[12] S. Kritchman, B. Nadler, ”Non-parametric Detection of the Number of Signals, Hypothesis Testing and Random Matrix Theory, *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3930–3941, 2009.

[13] S. Kay, ”Fundamentals of statistical signal processing, volume II: detection theory”, Prentice Hall, New Jersey 1998.

[14] P. Wang, J. Fang, N. Han, H. Li, ”Multiantenna-assisted spectrum sensing for cognitive radio”, *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1791-1800, May 2010.

[15] P. Loubaton, ”On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices”. ArXiv:1405.2006. To appear on J. of Theor. Prob.

[16] R.R. Nadakuditi, A. Edelman, ”Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples”, *IEEE Transactions of Signal Processing*, vol. 56, no. 7, pp. 2625-2637, July 2008.

[17] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, I. Santamaria, ”Detection of rank P signals in cognitive radio networks with uncalibrated multiple antennas”, *IEEE Trans. on Signal Process.*, vol. 59, no. 8, August 2011.

[18] D.T.M. Slock, ”Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction”, *in Proc. ICASSP 1994*, 19th IEEE International Conference on Acoustics, Speech, and Signal Processing, 19-22 April 1994, Adelaide, Australia.