



HAL
open science

On the detection of low-rank signal in the presence of spatially uncorrelated noise: a frequency domain approach

Alexis Rosuel, Philippe Loubaton, Pascal Vallet, Xavier Mestre

► To cite this version:

Alexis Rosuel, Philippe Loubaton, Pascal Vallet, Xavier Mestre. On the detection of low-rank signal in the presence of spatially uncorrelated noise: a frequency domain approach. *IEEE Transactions on Signal Processing*, 2021, 69, p. 4458-4473. hal-03268782

HAL Id: hal-03268782

<https://hal.science/hal-03268782v1>

Submitted on 23 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the detection of low-rank signal in the presence of spatially uncorrelated noise: a frequency domain approach.

A. Rosuel, *Student Member, IEEE*, P. Vallet, *Member, IEEE*, P. Loubaton, *Fellow, IEEE*, and X. Mestre, *Senior Member, IEEE*

Abstract

This paper analyzes the detection of a M -dimensional useful signal modeled as the output of a $M \times K$ MIMO filter driven by a K -dimensional white Gaussian noise, and corrupted by a M -dimensional Gaussian noise with mutually uncorrelated components. The study is focused on frequency domain test statistics based on the eigenvalues of an estimate of the spectral coherence matrix (SCM), obtained as a renormalization of the frequency-smoothed periodogram of the observed signal. If N denotes the sample size and B the smoothing span, it is proved that in the high-dimensional regime where M, B, N converge to infinity while K remains fixed, the SCM behaves as a certain correlated Wishart matrix. Exploiting well-known results on the behaviour of the eigenvalues of such matrices, it is deduced that the standard tests based on linear spectral statistics of the SCM fail to detect the presence of the useful signal in the high-dimensional regime. A new test based on the SCM, which is proved to be consistent, is also proposed, and its statistical performance is evaluated through numerical simulations.

Index Terms

detection, spectral coherence matrix, periodogram, high-dimensional statistics, Random Matrix Theory

I. INTRODUCTION

DETECTING the presence of an unknown multivariate signal corrupted by noise is one of the fundamental problems in signal processing, which is found in many applications including array and radar processing, wireless communications, radio-astronomy or seismology among others. In a statistical framework, this problem is usually formulated as the following binary hypothesis test, where the objective is to discriminate between the null hypothesis \mathcal{H}_0 and the alternative hypothesis \mathcal{H}_1 defined as

$$\begin{aligned} \mathcal{H}_0 : (\mathbf{y}_n)_{n \in \mathbb{Z}} &= (\mathbf{v}_n)_{n \in \mathbb{Z}} \\ \mathcal{H}_1 : (\mathbf{y}_n)_{n \in \mathbb{Z}} &= (\mathbf{u}_n)_{n \in \mathbb{Z}} + (\mathbf{v}_n)_{n \in \mathbb{Z}} \end{aligned} \quad (1)$$

where $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ is the M -variate observed signal, and where $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ and $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ represent a non observable signal of interest and the noise respectively, both modeled in this paper as mutually independent zero-mean complex Gaussian stationary time series.

Without further knowledge on the covariance function of $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ and/or $(\mathbf{u}_n)_{n \in \mathbb{Z}}$, or access to “noise only” samples, the test problem (1) is ill-posed, even for temporally white time series $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ and $(\mathbf{u}_n)_{n \in \mathbb{Z}}$, and one needs to exploit additional information on the covariance structure of the useful signal and noise. One common assumption, widely used in the context of array processing and multi-antenna communications, is to consider that the noise $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ is spatially uncorrelated. Moreover, when the receive antennas are not calibrated, it is reasonable to assume that the spectral densities of the components of the noise may not coincide, see e.g. [2], [3], [4], [5]. This will be the context of the present paper.

A first class of tests is based on the observation that the noise is spatially uncorrelated if and only if the matrices $\mathbf{R}_{\mathbf{v}}(\ell) = \mathbb{E}[\mathbf{v}_n \mathbf{v}_{n-\ell}^*]$ are diagonal for all $\ell \in \mathbb{Z}$, whereas if the useful signal $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ is assumed spatially correlated, $\mathbf{R}_{\mathbf{u}}(\ell) = \mathbb{E}[\mathbf{u}_n \mathbf{u}_{n-\ell}^*]$ is non-diagonal for some $\ell \in \mathbb{Z}$. Under this assumption, the problem in (1) can be formulated as the following correlation test:

$$\begin{aligned} \mathcal{H}_0 : \mathbf{R}_{\mathbf{y}}(\ell) &= \text{dg}(\mathbf{R}_{\mathbf{y}}(\ell)) \text{ for all } \ell \in \mathbb{Z} \\ \mathcal{H}_1 : \mathbf{R}_{\mathbf{y}}(\ell) &\neq \text{dg}(\mathbf{R}_{\mathbf{y}}(\ell)) \text{ for some } \ell \in \mathbb{Z} \end{aligned} \quad (2)$$

where $\mathbf{R}_{\mathbf{y}}(\ell) = \mathbb{E}[\mathbf{y}_n \mathbf{y}_{n-\ell}^*]$ and $\text{dg}(\mathbf{R}_{\mathbf{y}}(\ell)) = \mathbf{R}_{\mathbf{y}}(\ell) \odot \mathbf{I}_M$, where \odot is the element-wise (Hadamard) product and \mathbf{I}_M the $M \times M$ identity matrix. A number of previous works developed lag domains tests that specifically tackle the above problem,

A. Rosuel and P. Loubaton are with Laboratoire d'Informatique Gaspard Monge (CNRS, Univ. Gustave-Eiffel), 5 Bd. Descartes 77454 Marne-la-Vallée (France), {alexis.rosuel, philippe.loubaton}@univ-eiffel.fr

P. Vallet is with Laboratoire de l'Intégration du Matériau au Système (CNRS, Univ. Bordeaux, Bordeaux INP), 351, Cours de la Libération 33405 Talence (France), pascal.vallet@bordeaux-inp.fr

X. Mestre is with Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Av. Carl Friedrich Gauss 08860 Castelldefels, Barcelona (Spain), xavier.mestre@cttc.cat

This work was partially supported by project ANR-17-CE40-0003. The material of this paper was partly presented in the conference paper [1]

see e.g. [6], [7], [8], [9], [10], [11]. Also relevant are the approaches in [2] and [3], where the possible useful signal is supposed to be the output of a filter driven by a low-dimensional white noise sequence.

Our focus here is on another type of formulation, referred to as *frequency domain approach*, which consists in rewriting problem (2) as

$$\begin{aligned}\mathcal{H}_0 : \mathbf{S}_y(\nu) &= \text{dg}(\mathbf{S}_y(\nu)) \text{ for all } \nu \in [0, 1] \\ \mathcal{H}_1 : \mathbf{S}_y(\nu) &\neq \text{dg}(\mathbf{S}_y(\nu)) \text{ for some } \nu \in [0, 1]\end{aligned}\quad (3)$$

where $\mathbf{S}_y(\nu)$ is the $M \times M$ spectral density matrix of $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ at frequency ν , defined by

$$\mathbf{S}_y(\nu) = \sum_{k \in \mathbb{Z}} \mathbf{R}_y(k) e^{-2i\pi\nu k}.$$

This problem is equivalent to testing whether the *spectral coherence matrix* (see for instance [12, Chapter 7-6], [13, Chapter 5.5])

$$\mathbf{C}_y(\nu) = \text{dg}(\mathbf{S}_y(\nu))^{-\frac{1}{2}} \mathbf{S}_y(\nu) \text{dg}(\mathbf{S}_y(\nu))^{-\frac{1}{2}} \quad (4)$$

is equal to \mathbf{I}_M for all frequencies $\nu \in [0, 1]$. In this approach, usual test statistics are mostly based on consistent sample estimates of $\mathbf{S}_y(\nu)$ or $\mathbf{C}_y(\nu)$ that are compared to a diagonal matrix or to the identity \mathbf{I}_M respectively. Previous works that developed this approach include [14], [15], [16], [17]. In particular, [14] considered the following frequency smoothed-periodogram estimator $\hat{\mathbf{S}}_y(\nu)$ defined by

$$\hat{\mathbf{S}}_y(\nu) = \frac{1}{B+1} \sum_{b=-\frac{B}{2}}^{\frac{B}{2}} \boldsymbol{\xi}_y\left(\nu + \frac{b}{N}\right) \boldsymbol{\xi}_y\left(\nu + \frac{b}{N}\right)^* \quad (5)$$

with $\boldsymbol{\xi}_y(\nu) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{y}_n e^{-2i\pi n\nu}$ the renormalized finite Fourier transform of $(\mathbf{y}_n)_{n=0, \dots, N-1}$, B the *smoothing span*, assumed to be an even number, and where $\boldsymbol{\xi}_y(\nu + \frac{b}{N})^*$ is the conjugate transpose of the vector $\boldsymbol{\xi}_y(\nu + \frac{b}{N})$. [14] was devoted to the study of the limiting distribution of

$$\log \left\{ \prod_{i=1}^P \det(\hat{\mathbf{S}}_y(\nu_i)) / \prod_{m=1}^M \hat{s}_{m,m}(\nu_i) \right\}$$

for some properly defined subset of frequencies $(\nu_i)_{i=1, \dots, P}$, where $\hat{s}_{m,m}(\nu) = (\hat{\mathbf{S}}_y(\nu))_{m,m}$. When $M = 2$, [16] considered a general kernel estimator of $\mathbf{S}_y(\nu)$:

$$\tilde{\mathbf{S}}_y(\nu) = \sum_{b=-\frac{N}{2}}^{\frac{N}{2}} w_N\left(\frac{b}{N}\right) \boldsymbol{\xi}_y\left(\nu + \frac{b}{N}\right) \boldsymbol{\xi}_y\left(\nu + \frac{b}{N}\right)^*$$

where w_N is a weight function satisfying some specific properties, and a test statistic of the form

$$\frac{1}{N} \sum_{n=1}^N \frac{|(\tilde{\mathbf{S}}_y)_{12}(\nu)|^2}{(\tilde{\mathbf{S}}_y)_{11}(\nu)(\tilde{\mathbf{S}}_y)_{22}(\nu)}$$

which is proven to be, after proper recentring and renormalization, asymptotically normally distributed. Finally, [15] and [17] considered the more general class of test statistics, defined by:

$$\int_{-1/2}^{1/2} K\left((\tilde{\mathbf{S}}_y)_{12}(\nu)\right) d\nu \text{ and } \int_{-1/2}^{1/2} \left\| \psi\left((\tilde{\mathbf{S}}_y)_{12}(\nu), \nu\right) \right\|^2 d\nu$$

for some well-defined functions K and ψ , and where $\|\cdot\|$ is the Euclidian norm. They proved that these quantities asymptotically follow normal distributions. In the present paper, we focus on the natural estimator (see e.g. [12, Chapter 7-6], [13, Chapter 8-4]) of \mathbf{C}_y , defined by

$$\hat{\mathbf{C}}_y(\nu) = \text{dg}\left(\hat{\mathbf{S}}_y(\nu)\right)^{-\frac{1}{2}} \hat{\mathbf{S}}_y(\nu) \text{dg}\left(\hat{\mathbf{S}}_y(\nu)\right)^{-\frac{1}{2}} \quad (6)$$

where $\hat{\mathbf{S}}_y(\nu)$ is the frequency-smoothed periodogram estimate defined by (5). Note that adding a weight to the matrices $\boldsymbol{\xi}_y(\nu + \frac{b}{N}) \boldsymbol{\xi}_y(\nu + \frac{b}{N})^*$ leads to a more general class of estimators of $\mathbf{S}_y(\nu)$. The study of this more general class of estimators involves different techniques and random matrix models than the ones used here, and is therefore out of the scope of this paper.

A. Low vs High-dimensional regime

The performance of the test statistics developed in the above mentioned previous works is usually studied in the *low-dimensional regime* where $N \rightarrow \infty$ and M is fixed. It is well known (see for instance [12]) that $\hat{\mathbf{S}}_{\mathbf{y}}(\nu)$ and $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ are consistent estimates if $B \rightarrow +\infty$ and $\frac{B}{N} \rightarrow 0$. Under mild assumptions on the memory of the time series $(\mathbf{y}_n)_{n \in \mathbb{Z}}$, $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ is a consistent and asymptotically normal estimate of $\mathbf{C}_{\mathbf{y}}(\nu)$, which can in turn be used to study the asymptotic performance of the various tests based on $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$. In practice, the above asymptotic regime allows to predict the actual performance of the tests quite accurately, provided the ratio $\frac{M}{N}$ is small enough. If this condition is not met, test statistics based on $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ may be of delicate use, as the choice of the smoothing span B must meet the constraints $\frac{B}{M}$ much larger than 1 (because B is supposed to converge towards $+\infty$) as well as $\frac{B}{N}$ small enough (because $\frac{B}{N}$ is supposed to converge towards 0).

Nowadays, in many practical applications involving high-dimensional signals and/or a moderate sample size, the ratio $\frac{M}{N}$ may not be small enough to be able to choose B so as to meet $\frac{B}{M}$ much larger than 1 and $\frac{B}{N}$ small enough. Therefore, the results obtained in the low-dimensional regime may fail to provide accurate predictions of the behaviour of the aforementioned test statistics. In this situation, one may rely on the more relevant *high-dimensional regime* in which M, B, N converge to infinity such that $\frac{M}{B}$ converges to a positive constant while $\frac{B}{N}$ converges to zero.

In comparison to the low-dimensional regime, the literature concerning correlation tests for the frequency domain in the high-dimensional regime is quite scarce. Recent results obtained in [18] show that under hypothesis \mathcal{H}_0 , the empirical eigenvalue distribution of the spectral coherence estimate $\hat{\mathbf{C}}(\nu)$ behaves in the high-dimensional regime as the well-known Marcenko-Pastur distribution [19]. The result of [18] allows to predict the performance under \mathcal{H}_0 of a large class of test statistics based on

$$L_\varphi(\nu) = \frac{1}{M} \sum_{m=1}^M \varphi(\lambda_m(\hat{\mathbf{C}}_{\mathbf{y}}(\nu)))$$

where $\lambda_1(\hat{\mathbf{C}}_{\mathbf{y}}(\nu)), \dots, \lambda_M(\hat{\mathbf{C}}_{\mathbf{y}}(\nu))$ are the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, and φ belongs to a certain functional class. Such family of statistics L_φ , called *linear spectral statistics (LSS)* of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, include in particular the choice $\varphi(x) = \log x$, i.e. $L_\varphi(\nu) = \frac{1}{M} \log \det \hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ and the choice $\varphi(x) = (x-1)^2$, i.e. $L_\varphi(\nu) = \frac{1}{M} \|\hat{\mathbf{C}}_{\mathbf{y}}(\nu) - \mathbf{I}_M\|_F^2$, where $\|\cdot\|_F$ represents the Frobenius norm.

In this paper, we consider the study of the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ in the high-dimensional regime under the special alternative \mathcal{H}_1 for which the useful signal $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ is modeled as the output of a stable MIMO filter driven by a K -dimensional white complex Gaussian noise. In the context where the intrinsic dimension K is fixed while $M, N, B \rightarrow \infty$, it is shown that the empirical eigenvalue distribution of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ still converges to the Marcenko-Pastur distribution, showing that the test statistic based on $L_\varphi(\nu)$ is unable to discriminate between hypotheses \mathcal{H}_0 and \mathcal{H}_1 in the high-dimensional regime. Nevertheless, we also prove that, provided that the signal-to-noise ratio is large enough, the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ asymptotically splits from the support of the Marcenko-Pastur distribution. We can therefore exploit this result to design a new frequency domain test statistic, which is shown to be consistent in the high-dimensional regime. This result is connected to the widely studied *spiked models* in Random Matrix Theory, defined as low rank perturbations of large random matrices. These models were extensively studied in the context of sample covariance matrices of independent identically distributed high-dimensional vectors, see e.g. [20]. We however notice that papers addressing the behaviour of the corresponding sample correlation matrices are quite scarce, see [21] when the low rank perturbation affects only the first components of the observations.

B. Related works

Although the asymptotic framework differs from the high-dimensional regime considered here, we also mention the series of studies [22], [23] in the econometrics field, which consider a similar model under \mathcal{H}_1 . In these works, it is assumed that $M, N \rightarrow \infty$ so the ratio $\frac{M}{N}$ remains bounded, while the K non-zero eigenvalues of the spectral density $\mathbf{S}_{\mathbf{u}}(\nu)$ of $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ are assumed to converge towards $+\infty$ at rate M . This last assumption, which ensures that the Signal-to-Noise Ratio (SNR) $\frac{\mathbb{E}\|\mathbf{u}_n\|^2}{\mathbb{E}\|\mathbf{v}_n\|^2}$ remains bounded away from 0 as $M \rightarrow \infty$, significantly facilitates the design of consistent detection methods. Nevertheless, while relevant in the domain of econometrics, this assumption may be unrealistic in several applications of array processing, where the challenge is to manage situations in which the SNR converges towards 0 at rate $\frac{1}{M}$. This situation is the one considered in this paper and, in that case, the results of [22], [23] can not be used. We discuss this point further in Section II below.

The rest of the paper is organized as follows. In Section II, we formally introduce the model of signals used in the remainder, as well as the required technical assumptions. In section III, we introduce informally the proposed test statistic, and illustrate its behaviour in order to provide some intuition before a more rigorous presentation. In section IV, we study some approximation results for the spectral coherence $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ which are useful to study the linear spectral statistics considered here. This study is then used in Section V to introduce a new test statistic that is consistent in the high-dimensional regime. Finally Section VI provides some simulations illustrating its performance and comparisons against other relevant approaches.

Notations. For a complex matrix \mathbf{A} , we denote by \mathbf{A}^* its conjugate transpose, and by $\|\mathbf{A}\|_2$ and $\|\mathbf{A}\|_F$ its spectral and Frobenius norms respectively. If \mathbf{A} is a $n \times n$ complex matrix, we denote by $\text{Tr}(\mathbf{A})$ its trace, and by $\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A})$ its eigenvalues; if moreover \mathbf{A} is Hermitian, they are sorted in decreasing order $\lambda_1(\mathbf{A}) \geq \dots \geq \lambda_n(\mathbf{A})$. The $n \times n$ identity matrix is denoted \mathbf{I}_n . The expectation of a complex random variable Z is denoted by $\mathbb{E}[Z]$. The complex circular Gaussian distribution with variance σ^2 is denoted as $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ and a random vector \mathbf{x} of \mathbb{C}^n follows the $\mathcal{N}_{\mathbb{C}^n}(\mathbf{0}, \mathbf{R})$ distribution if $\mathbf{b}^* \mathbf{x} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{b}^* \mathbf{R} \mathbf{b})$ for all deterministic (column) vector \mathbf{b} and a fixed $n \times n$ positive definite matrix \mathbf{R} . Finally, $\mathcal{C}^1(I)$ (resp. $\mathcal{C}_c^1(I)$) represents the set of continuously differentiable functions (resp. continuously differentiable functions with compact support) on an open set I .

II. MODEL AND ASSUMPTIONS

Let us consider a M -dimensional observed time series $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ defined as

$$\mathbf{y}_n = \mathbf{u}_n + \mathbf{v}_n \quad (7)$$

where $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ represents a useful signal and where $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ represents an additive noise. The useful signal is modeled as the output of an unknown stable MIMO filter $(\mathbf{H}_k)_{k \in \mathbb{Z}}$ driven by a non-observable K -dimensional complex Gaussian white noise $(\epsilon_n)_{n \in \mathbb{Z}}$ with $\mathbb{E}[\epsilon_n \epsilon_n^*] = \mathbf{I}_K$, i.e.

$$\mathbf{u}_n = \sum_{k \in \mathbb{Z}} \mathbf{H}_k \epsilon_{n-k}$$

with probability one. We notice that K represents the number of sources in the context of array processing. $(\mathbf{v}_n)_{n \in \mathbb{Z}}$ is modeled as a M -dimensional stationary complex Gaussian time series such that its component time series $(v_{1,n})_{n \in \mathbb{Z}}, \dots, (v_{M,n})_{n \in \mathbb{Z}}$ are mutually independent.

For each $m = 1, \dots, M$, we denote by $(r_m(k))_{k \in \mathbb{Z}}$ the covariance function of $(v_{m,n})_{n \in \mathbb{Z}}$, i.e. $r_m(k) = \mathbb{E}[v_{m,n} \overline{v_{m,n-k}}]$, which verifies the following memory assumption.

Assumption 1. *The covariance coefficients decay sufficiently fast in the lag domain, in the sense that*

$$\sup_{m \geq 1} \sum_{k \in \mathbb{Z}} (1 + |k|)^2 |r_m(k)| < \infty. \quad (8)$$

In particular, Assumption 1 implies that the spectral density s_m of $(v_{m,n})_{n \in \mathbb{Z}}$, given by

$$s_m(\nu) = \sum_{k \in \mathbb{Z}} r_m(k) e^{-i2\pi\nu k}$$

verifies

$$\sup_{m \geq 1} \sup_{\nu \in [0,1]} s_m(\nu) < \infty.$$

Assumption 1 is in particular verified as soon as the condition

$$|r_m(k)| \leq \frac{C}{(1 + |k|)^{3+\delta}} \quad (9)$$

holds for each $k \in \mathbb{Z}$ and each $m \geq 1$, where C and δ are positive constants. As the autocovariance function of ARMA signals decreases exponentially towards 0, Assumption 1 thus holds if the time series $(v_m)_{m \geq 1}$ are ARMA signals, provided some extra purely technical conditions that allow to manage the supremum over m in (8) are met. As the spectral coherence matrix of $(\mathbf{v}_n)_{n \in \mathbb{Z}}$, involves a renormalization by the inverse of the spectral densities s_m , we also need that s_m does not vanish for each m .

Assumption 2. *The spectral densities are uniformly bounded away from zero, that is*

$$\inf_{m \geq 1} \min_{\nu \in [0,1]} s_m(\nu) > 0.$$

Assumptions 1 and 2 also imply that the total noise power satisfies

$$0 < \inf_{M \geq 1} \frac{1}{M} \mathbb{E} \|\mathbf{v}_n\|_2^2 \leq \sup_{M \geq 1} \frac{1}{M} \mathbb{E} \|\mathbf{v}_n\|_2^2 < \infty. \quad (10)$$

The next assumption is related to the signal part $(\mathbf{u}_n)_{n \in \mathbb{Z}}$. For each $\nu \in [0, 1]$, we denote by $\mathbf{H}(\nu)$ the Fourier transform of $(\mathbf{H}_k)_{k \in \mathbb{Z}}$, i.e.

$$\mathbf{H}(\nu) = \sum_{k \in \mathbb{Z}} \mathbf{H}_k e^{-i2\pi\nu k}$$

and by $\mathbf{h}^1(\nu), \dots, \mathbf{h}^M(\nu)$ the rows of $\mathbf{H}(\nu)$.

Assumption 3. *The MIMO filter coefficient matrices are such that*

$$\sup_{M \geq 1} \sum_{k \in \mathcal{Z}} (1 + |k|) \|\mathbf{H}_k\|_2 < \infty \quad (11)$$

and

$$\lim_{M \rightarrow \infty} \max_{m=1, \dots, M} \max_{\nu \in [0, 1]} \|\mathbf{h}^m(\nu)\|_2 = 0. \quad (12)$$

When K is fixed while $M \rightarrow \infty$, condition (11) in Assumption 3 implies that the total useful signal power remains bounded, i.e.

$$\mathbb{E} \|\mathbf{u}_n\|_2^2 = \sum_{k \in \mathcal{Z}} \|\mathbf{H}_k\|_F^2 = \mathcal{O}(1) \quad (13)$$

so that, using (10), the SNR vanishes at rate $\frac{1}{M}$, i.e.

$$\frac{\mathbb{E} \|\mathbf{u}_n\|_2^2}{\mathbb{E} \|\mathbf{v}_n\|_2^2} = \mathcal{O}\left(\frac{1}{M}\right). \quad (14)$$

Likewise, condition (12) in Assumption 3 implies that the SNR per time series vanishes, i.e.

$$\frac{\mathbb{E} |u_{m,n}|^2}{\mathbb{E} |v_{m,n}|^2} = \frac{\int_0^1 \|\mathbf{h}^m(\nu)\|_2^2 d\nu}{\int_0^1 s_m(\nu) d\nu} = o(1) \quad (15)$$

as $M \rightarrow \infty$. We finally notice that (11) is stronger than (13). While $\mathbb{E} \|\mathbf{u}_n\|_2^2 = \mathcal{O}(1)$ is a rather fundamental assumption that allows to precise the behaviour of the signal to noise ratio, the extra condition $\sup_{m \geq 1} \sum_k |k| \|\mathbf{H}_k\|_2 < \infty$ is essentially motivated by technical reasons (it is needed to establish Theorem 2). However, it is clearly not restrictive in practice.

Remark 1. *Conditions (11) and (12) in Assumption 3 are especially relevant in the context of array processing, where M represents the number of sensors, which may be large [24], [25]. In this context, (14) represents the SNR before matched filtering, while (15) represents the SNR per sensor. The use of spatial filtering techniques, which combine the observations $y_{1,n}, \dots, y_{M,n}$ across the M sensors, allows to increase the SNR by a factor M when the second order statistics of $(\mathbf{y}_n)_{n \in \mathcal{Z}}$ are known, which leads to a SNR after matched filtering of the order of magnitude $\mathcal{O}(1)$. Thus, despite the apparent low SNR, reliable information on the useful signal $(\mathbf{u}_n)_{n \in \mathcal{Z}}$ can potentially still be extracted from the observed signal $(\mathbf{y}_n)_{n \in \mathcal{Z}}$.*

Let \mathbf{S}_y denote the spectral density of $(\mathbf{y}_n)_{n \in \mathcal{Z}}$, given by

$$\mathbf{S}_y(\nu) = \mathbf{H}(\nu)\mathbf{H}(\nu)^* + \mathbf{S}_v(\nu)$$

where $\mathbf{S}_v(\nu) = \text{dg}(s_1(\nu), \dots, s_M(\nu))$. To estimate \mathbf{S}_y , we consider in this paper a frequency-smoothed periodogram $\hat{\mathbf{S}}_y$, which we defined in (5). In the classical low-dimensional regime where $B, N \rightarrow \infty$ while M, K remain fixed, it is well-known [12] that

$$\mathbb{E}[\hat{\mathbf{S}}_y(\nu)] = \mathbf{S}_y(\nu) + \mathcal{O}\left(\frac{B^2}{N^2}\right)$$

and

$$\mathbb{E} \left\| \hat{\mathbf{S}}_y(\nu) - \mathbb{E}[\hat{\mathbf{S}}_y(\nu)] \right\|_2^2 = \mathcal{O}\left(\frac{1}{B}\right).$$

Thus, in this regime, $\hat{\mathbf{S}}_y(\nu)$ is a consistent estimator of $\mathbf{S}_y(\nu)$ as long as $B \rightarrow \infty$ and $\frac{B}{N} \rightarrow 0$. Likewise, the sample Spectral Coherence Matrix (SCM, not to be confused with the sample covariance matrix, which will not be used in this paper) defined in (6) is a consistent estimator of the true SCM $\mathbf{C}_y(\nu)$ defined in (4). When $M \rightarrow +\infty$ and $\frac{M}{N} \rightarrow 0$, it can be shown that, under some additional mild extra assumptions, the consistency of $\hat{\mathbf{S}}_y(\nu)$ and $\hat{\mathbf{C}}_y(\nu)$ in the spectral norm sense still holds provided that B is chosen in such a way that $\frac{B}{N} \rightarrow 0$ and $\frac{M}{B} \rightarrow 0$. In practice, for finite values of M and N , the above asymptotic regime will allow to predict the performance of various inference schemes in situations where it is possible to choose B in such a way that $\frac{M}{B}$ and $\frac{B}{N}$ are both small enough. Nevertheless, when the dimension M is large and the sample size N is not unlimited, or equivalently if $\frac{M}{N}$ is not small enough, such a choice of B may be impossible. In such a context, it seems more relevant to consider asymptotic regimes for which $\frac{M}{N} \rightarrow 0$ and $\frac{M}{B}$ converging towards a positive constant. In the following, we will consider the following asymptotic regime.

Assumption 4. $N = N(M)$ and $B = B(M)$ are both functions of M such that, for some $\alpha \in (0, 1)$,

$$M = \mathcal{O}(N^\alpha) \text{ and } \frac{M}{B} \xrightarrow{M \rightarrow \infty} c \in (0, 1)$$

while K is fixed with respect to M .

As $\frac{M}{B}$ does not converge towards 0, the consistency of $\hat{\mathbf{S}}_{\mathbf{y}}(\nu)$ and $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ is lost. This can be explained in a simple way when $\mathbf{u}_n = 0$ for each n and the signals $((v_{m,n})_{n \in \mathbb{Z}})_{m \geq 1}$ are mutually independent i.i.d. $\mathcal{N}_c(0, \sigma^2)$ distributed sequences. In this context, for each ν , the renormalized Fourier transform vectors $(\xi_{\mathbf{y}}(\nu + b/N))_{b=-B/2, \dots, B/2}$ are mutually independent $\mathcal{N}_{\mathbb{C}}(0, \sigma^2 \mathbf{I})$ random vectors. The spectral density estimate $\hat{\mathbf{S}}_{\mathbf{y}}(\nu)$ defined by (5) thus coincides with the sample covariance matrix of these $(B + 1)$ M -dimensional vectors. If B and M are of the same order to magnitude, it cannot be expected that $\|\hat{\mathbf{S}}_{\mathbf{y}}(\nu) - \mathbb{E}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))\|$ converges towards 0 because the true covariance matrix $\mathbb{E}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))$ to be estimated depends on $\mathcal{O}(M^2)$ parameters, and that the number MB of available scalar observations used to estimate $\mathbb{E}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))$ is also $\mathcal{O}(M^2)$. Despite the loss of the convergence of the estimators $\hat{\mathbf{S}}_{\mathbf{y}}(\nu)$ and $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, we will see that one can still rely on the high-dimensional structure of these matrices to design relevant test statistics.

III. INFORMAL PRESENTATION OF THE PROPOSED TEST STATISTIC

Mathematical details will reveal later that for each ν , $\hat{\mathbf{C}}(\nu)$ behaves as a spike model covariance matrix, whose eigenvalues are precisely described by [20]. More precisely, we will see that, in some sense, the eigenvalues of $\hat{\mathbf{C}}(\nu)$ that are due to the noise belong to the interval $[\lambda_-, \lambda_+]$ where $\lambda_- = (1 - \sqrt{c})^2$ and $\lambda_+ = (1 + \sqrt{c})^2$, and that in the presence of signal, some eigenvalues of $\hat{\mathbf{C}}(\nu)$ may be strictly greater than λ_+ if an SNR criteria is respected. For the remainder, we define

$$\mathcal{V}_N = \left\{ \frac{k}{N} : k = 0, \dots, N - 1 \right\} \quad (16)$$

the set of Fourier frequencies. A natural way to test for \mathcal{H}_0 against \mathcal{H}_1 is to compute the largest eigenvalue of $\hat{\mathbf{C}}(\nu)$ over the frequencies of \mathcal{V}_N , and compare it with λ_+ . This leads to the following test statistic:

$$T_\epsilon = \mathbb{1}_{[\lambda_+ + \epsilon, \infty)} \left(\max_{\nu \in \mathcal{V}_N} \lambda_1 \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu) \right) \right). \quad (17)$$

We will prove later that, under proper assumption on the SNR, this test statistic is consistent in the present high-dimensional regime. Before describing the mathematical details leading to consider T_ϵ , we now provide some numerical illustrations of its behaviour. The general settings are given as follows. The noise is generated as a Gaussian AR(1) process having spectral density

$$s_m(\nu) = \frac{1}{|1 - \theta e^{-i2\pi\nu}|^2}, \quad (18)$$

for all $m = 1, \dots, M$, with $\theta = 0.5$, whereas for the useful signal, we also consider an AR(1) process by choosing $K = 1$ and

$$\mathbf{H}_k = \sqrt{\frac{C}{M}} \beta^k (1, \dots, 1)^T \quad (19)$$

with $\beta = \frac{10}{11}$ and C being a positive constant used to adjust the SNR.

In order to understand how the test statistics T_ϵ discriminates between \mathcal{H}_0 and \mathcal{H}_1 , we show in Figure 1 the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ for $\nu \in \mathcal{V}_N$ in the presence of signal, and compare it to the threshold λ_+ . We see that for some frequencies ν around 0, the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ deviates significantly from λ_+ . As we will see later, it is possible to evaluate the asymptotic behaviour of the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, and to establish that it converges towards $\phi(\text{SNR}(\nu))$ where ϕ is a certain function, and where $\text{SNR}(\nu)$ can be interpreted as a signal-to-noise ratio at frequency ν . $\phi(\text{SNR}(\nu))$ is also represented in Figure 1, and it is seen that it is close to the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$. In Figure 2, we compare the empirical distribution of T_ϵ under \mathcal{H}_0 and \mathcal{H}_1 over 10000 repetitions. We see that the distribution of our test statistic T_ϵ is able to discriminate the scenarios where the data \mathbf{y}_n are generated under \mathcal{H}_0 or \mathcal{H}_1 , and that T_ϵ is typically over the threshold λ_+ under \mathcal{H}_1 .

IV. APPROXIMATION RESULTS FOR $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ IN THE HIGH-DIMENSIONAL REGIME

In this section we present the mathematical details which lead to the test statistic (17). More specifically, we provide useful approximation results for $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, which basically show that $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ behaves as a certain Wishart matrix in the high-dimensional regime. These approximation results are the keystone for the study of the behaviour of the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ and the detection test proposed in Section V.

We first study separately the signal-free case (i.e. $\mathbf{y}_n = \mathbf{v}_n$) as well as the noise-free case (i.e. $\mathbf{y}_n = \mathbf{u}_n$).

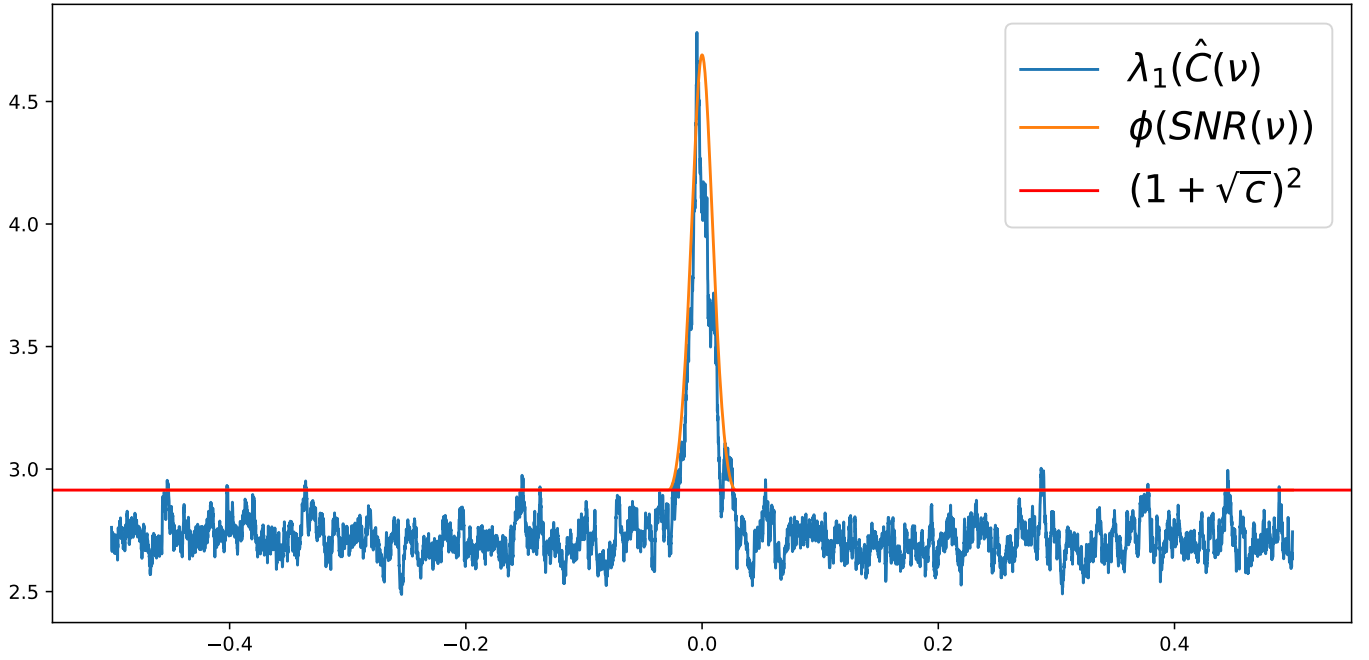


Fig. 1. Largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ for $\nu \in \mathcal{V}_N$ vs the threshold $\lambda_+ = (1 + \sqrt{\frac{M}{B+1}})^2$. $M = 60$, $c = 0.5$, $N = 6000$, $\theta = 0.5$, $C = 0.05$

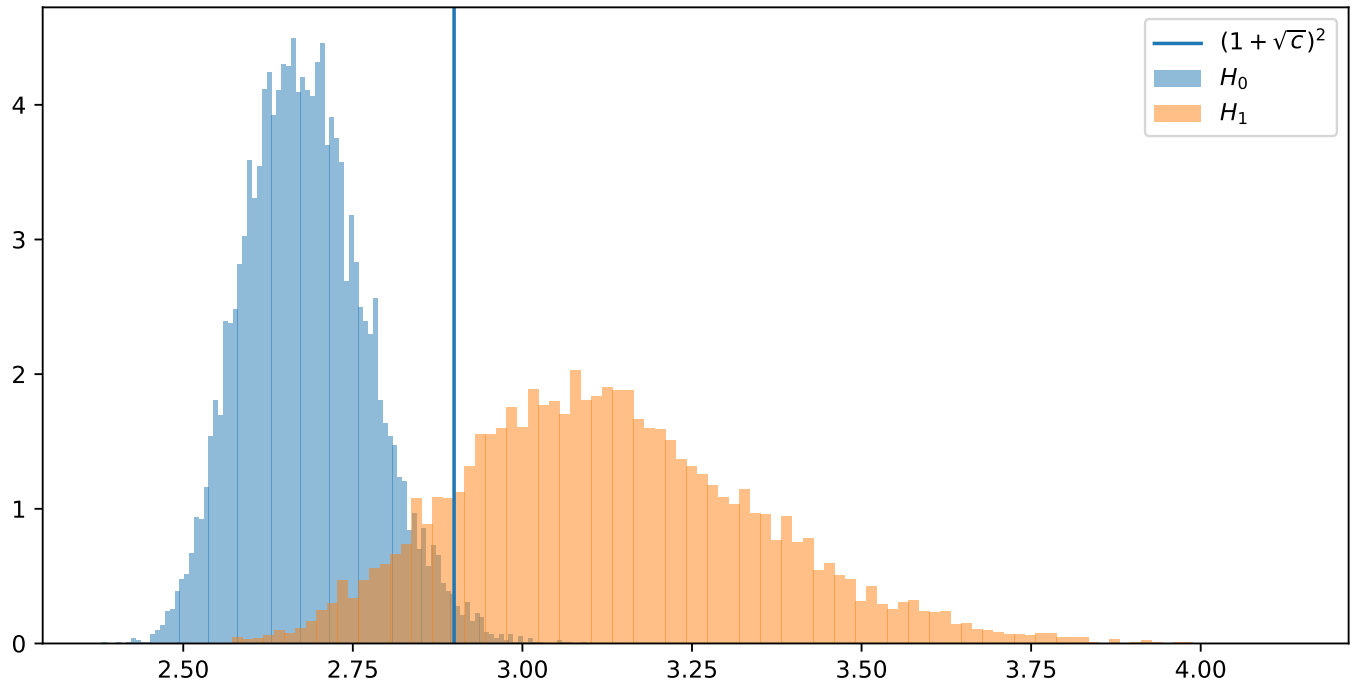


Fig. 2. Histogram of T_ϵ under \mathcal{H}_0 and \mathcal{H}_1 , over 10000 repetitions. $M = 40$, $c = 0.5$, $N = 1000$, $\theta = 0.5$, $C = 0.05$

A. Signal-free case

Let

$$\boldsymbol{\xi}_{\mathbf{v}}(\nu) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{v}_n e^{-i2\pi\nu n}$$

denote the discrete (time-limited) Fourier transform of $(\mathbf{v}_n)_{n=0,\dots,N-1}$, and define the $M \times (B+1)$ matrix $\Sigma_{\mathbf{v}}(\nu)$ as

$$\Sigma_{\mathbf{v}}(\nu) = \frac{1}{\sqrt{B+1}} \left[\xi_{\mathbf{v}} \left(\nu - \frac{B}{2N} \right), \dots, \xi_{\mathbf{v}} \left(\nu + \frac{B}{2N} \right) \right].$$

The following result, derived in [18], reveals an interesting behaviour of the frequency-smoothed periodogram of the noise.

Theorem 1. *Under Assumptions 1, 2 and 4, for all $\nu \in \mathcal{V}_N$, there exists an $M \times (B+1)$ matrix $\mathbf{Z}(\nu)$ with i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$ entries such that*

$$\max_{\nu \in \mathcal{V}_N} \left\| \Sigma_{\mathbf{v}}(\nu) - \frac{1}{\sqrt{B+1}} \mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \mathbf{Z}(\nu) \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Informally speaking, Theorem 1 shows that the random vectors $\frac{1}{\sqrt{B+1}} \xi_{\mathbf{v}} \left(\nu - \frac{B}{2N} \right), \dots, \frac{1}{\sqrt{B+1}} \xi_{\mathbf{v}} \left(\nu + \frac{B}{2N} \right)$ asymptotically behave as a family of i.i.d. $\mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{S}_{\mathbf{v}}(\nu))$ vectors, for all $\nu \in \mathcal{V}_N$. Moreover, if

$$\hat{\mathbf{S}}_{\mathbf{v}}(\nu) := \frac{1}{B+1} \sum_{b=-B/2}^{B/2} \xi_{\mathbf{v}} \left(\nu + \frac{b}{N} \right) \xi_{\mathbf{v}} \left(\nu + \frac{b}{N} \right)^*$$

denotes the frequency-smoothed periodogram of the noise observations $(\mathbf{v}_n)_{n \in \mathbb{Z}}$, we deduce that $\hat{\mathbf{S}}_{\mathbf{v}}(\nu)$ asymptotically behaves as a complex Gaussian Wishart matrix with covariance matrix $\mathbf{S}_{\mathbf{v}}(\nu)$, thanks to the following corollary.

Corollary 1. *Under the assumptions of Theorem 1, it holds that*

$$\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{v}}(\nu) - \mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \frac{\mathbf{Z}(\nu) \mathbf{Z}(\nu)^*}{B+1} \mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Proof: The proof is deferred to Appendix D-A. ■

It is worth noticing that Corollary 1 implies in particular

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{v}}(\nu) \right) - \mathbf{S}_{\mathbf{v}}(\nu) \right\| \xrightarrow[M \rightarrow \infty]{a.s.} 0$$

and consequently $\text{dg}(\hat{\mathbf{S}}_{\mathbf{v}}(\nu))$ is a consistent estimator of the noise spectral density $\mathbf{S}_{\mathbf{v}}(\nu)$ in the operator norm sense, at each Fourier frequency $\nu \in \mathcal{V}_N$. This convergence may be directly obtained using Lemma 1 in Appendix A and we omit the details since this result is well-known.

B. Noise-free case

Let

$$\xi_{\mathbf{u}}(\nu) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{u}_n e^{-i2\pi\nu n}$$

and let $\Sigma_{\mathbf{u}}(\nu)$ be the $K \times (B+1)$ matrix defined as

$$\Sigma_{\mathbf{u}}(\nu) = \frac{1}{\sqrt{B+1}} \left[\xi_{\mathbf{u}} \left(\nu - \frac{B}{2N} \right), \dots, \xi_{\mathbf{u}} \left(\nu + \frac{B}{2N} \right) \right].$$

In the same way, we also denote by ξ_{ϵ} the normalized discrete (time-limited) Fourier transform of $(\epsilon_n)_{n=0,\dots,N-1}$, and consider the $K \times (B+1)$ matrix $\Sigma_{\epsilon}(\nu)$ defined as $\Sigma_{\mathbf{u}}(\nu)$. We then have the following important approximation result.

Theorem 2. *Under Assumptions 3 and 4, it holds that*

$$\max_{\nu \in \mathcal{V}_N} \left\| \Sigma_{\mathbf{u}}(\nu) - \mathbf{H}(\nu) \Sigma_{\epsilon}(\nu) \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Proof: The proof is deferred to Appendix B. ■

As in Theorem 1, Theorem 2 shows that the random vectors $\xi_{\mathbf{u}} \left(\nu - \frac{B}{2N} \right), \dots, \xi_{\mathbf{u}} \left(\nu + \frac{B}{2N} \right)$ asymptotically behave as the i.i.d. vectors $\mathbf{H}(\nu) \xi_{\epsilon} \left(\nu - \frac{B}{2N} \right), \dots, \mathbf{H}(\nu) \xi_{\epsilon} \left(\nu + \frac{B}{2N} \right)$, for all $\nu \in \mathcal{V}_N$.

Remark 2. *The type of approximation given in Theorem 2 is well-known in the low-dimensional regime in which M, K, B are fixed while $N \rightarrow \infty$. Indeed, in that case, we have [12, Th. 4.5.2]*

$$\max_{\nu \in [0, 1]} \left\| \Sigma_{\mathbf{u}}(\nu) - \mathbf{H}(\nu) \Sigma_{\epsilon}(\nu) \right\|_2 = \mathcal{O}_P \left(\sqrt{\frac{\log(N)}{N}} \right).$$

In the high-dimensional regime where M and B also converge to infinity as described in Assumption 4, the result of Theorem 2 cannot be obtained from [12, Th. 4.5.2] and thus requires a new study.

We also deduce the following approximation result on the frequency-smoothed periodogram of the signal observations $(\mathbf{u}_n)_{n=0,\dots,N-1}$ given by

$$\hat{\mathbf{S}}_{\mathbf{u}}(\nu) := \frac{1}{B+1} \sum_{b=-B/2}^{B/2} \boldsymbol{\xi}_{\mathbf{u}} \left(\nu + \frac{b}{N} \right) \boldsymbol{\xi}_{\mathbf{u}} \left(\nu + \frac{b}{N} \right)^*.$$

Corollary 2. *Under the assumptions of Theorem 2, it holds that*

$$\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{u}}(\nu) - \mathbf{H}(\nu) \mathbf{H}(\nu)^* \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Proof: The proof is deferred to Appendix D-B. ■

As a result of Corollary 2, we deduce that the frequency-smoothed periodogram $\hat{\mathbf{S}}_{\mathbf{u}}(\nu)$ is a consistent estimator of the spectral density $\mathbf{S}_{\mathbf{u}}(\nu) = \mathbf{H}(\nu) \mathbf{H}(\nu)^*$ of $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ in the high-dimensional regime, for each $\nu \in \mathcal{V}_N$.

Having characterized the pure noise and pure signal cases, we are now in position to study the high-dimensional behaviour of the spectral coherence matrix $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$.

C. The signal-plus-noise case

First, using Corollaries 1 and 2, we deduce the high-dimensional behaviour of the frequency smoothed periodogram $\hat{\mathbf{S}}_{\mathbf{y}}(\nu)$. The following results show that, as it could be expected, the frequency smoothed periodogram essentially behaves as a colored Wishart matrix in the large asymptotic regime.

Proposition 1. *For all $\nu \in \mathcal{V}_N$, there exists an $M \times (B+1)$ matrix $\mathbf{X}(\nu)$ with i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$ entries such that*

$$\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) - \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu) \mathbf{X}(\nu)^*}{B+1} \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0. \quad (20)$$

Proof: The proof is deferred to Appendix D-C. ■

We finally consider the study of the spectral coherence $\hat{\mathbf{C}}_{\mathbf{y}}(\nu) = \text{dg}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))^{-\frac{1}{2}} \hat{\mathbf{S}}_{\mathbf{y}}(\nu) \text{dg}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))^{-\frac{1}{2}}$. From condition (12) in Assumption 3 on the SNR, it turns out that (cf. proof of Theorem 3 below where the result is shown) that

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right) - \mathbf{S}_{\mathbf{v}}(\nu) \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0. \quad (21)$$

This approximation result regarding the normalization term $\text{dg}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))$ in the SCM naturally leads to the following theorem, which is the key result of this paper.

Theorem 3. *Under Assumptions 1, 2, 3 and 4,*

$$\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{C}}_{\mathbf{y}}(\nu) - \boldsymbol{\Xi}(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu) \mathbf{X}(\nu)^*}{B+1} \boldsymbol{\Xi}(\nu)^{\frac{1}{2}} \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0$$

where

$$\boldsymbol{\Xi}(\nu) = \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \mathbf{H}(\nu) \mathbf{H}(\nu)^* \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} + \mathbf{I}_M.$$

and $\mathbf{X}(\nu)$ is the matrix defined in Proposition 1.

Proof: The proof is deferred to Appendix C. ■

Let us make a few important comments regarding the result of Theorem 3.

First, used in conjunction with Weyl's inequalities [26, Th. 4.3.1], Theorem 3 implies in particular that each eigenvalue of the SCM $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ behaves as its counterpart of the Wishart matrix

$$\mathbf{W}(\nu) = \boldsymbol{\Xi}(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu) \mathbf{X}(\nu)^*}{B+1} \boldsymbol{\Xi}(\nu)^{\frac{1}{2}}$$

for $\nu \in \mathcal{V}_N$, that is

$$\max_{m=1,\dots,M} \max_{\nu \in \mathcal{V}_N} \left| \lambda_m \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu) \right) - \lambda_m \left(\mathbf{W}(\nu) \right) \right| \xrightarrow[M \rightarrow \infty]{a.s.} 0. \quad (22)$$

Second, Theorem 3 has an important consequence regarding the behaviour of linear spectral statistics of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, that is statistics of the type

$$L_{\varphi}(\nu) = \frac{1}{M} \sum_{m=1}^M \varphi \left(\lambda_m \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu) \right) \right) \quad (23)$$

where φ belongs to a certain class of functions.

Corollary 3. Let $\varphi \in \mathcal{C}^1((0, +\infty))$. Under Assumptions 1, 2, 3 and 4, we have

$$\max_{\nu \in \mathcal{V}_N} \left| L_\varphi(\nu) - \int_{\mathbb{R}} \varphi(\lambda) f(\lambda) d\lambda \right| \xrightarrow[M \rightarrow \infty]{a.s.} 0$$

where f is the density of the Marcenko-Pastur distribution given by

$$f(\lambda) = \frac{\sqrt{(\lambda - \lambda^-)(\lambda^+ - \lambda)}}{2\pi c\lambda} \mathbb{1}_{[\lambda^-, \lambda^+]}(\lambda)$$

with $\lambda^\pm = (1 \pm \sqrt{c})^2$.

Proof: The proof is deferred to Appendix E. ■

Therefore, Corollary 3 shows that linear spectral statistics of the SCM converge to the same limit regardless of whether the observations contain only pure noise or signal-plus-noise contributions. This shows that any test statistic solely relying on linear spectral statistics of the SCM is unable to distinguish between absence or presence of useful signal, and cannot be consistent in the high-dimensional regime. Nevertheless, in the next section we will see that we can exploit Theorem 3 to build a new test statistic based on the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$, which is proved to be consistent in the high-dimensional regime.

Remark 3. Corollary 1, Corollary 2 and Theorems 3 may be interpreted in the context of array processing. Indeed, in the time model (7), usually referred to as “wideband”, the signal contribution $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ modeled as a linear process, is in general not confined to a low-dimensional subspace (i.e. with dimension less than M). However, in the frequency domain, Corollary 1 and Corollary 2 show that we can retrieve, in the high-dimensional regime, a “narrowband” model, since the useful signal is confined to a K -dimensional subspace of \mathbb{C}^M . Thus, standard narrowband techniques used in array processing for detection may be used, see e.g. [27].

V. A NEW CONSISTENT TEST STATISTIC

As we have seen in Theorem 3 and the related comments, the SCM $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ behaves in the high-dimensional regime as a Wishart matrix with scale $\Xi(\nu) = \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \mathbf{H}(\nu) \mathbf{H}(\nu)^* \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} + \mathbf{I}_M$ being a fixed rank K perturbation of the identity matrix. The behaviour of the eigenvalues for each ν of such matrix model is well-known since [20] (and other related works such as the well-known BBP-phase-transition [28] or [29]), and the rest of this section is devoted to the application of the results from [20] in our frequency-domain detection context. A crucial point is to choose the particular frequency at which the above mentioned results will be used in order to obtain information on the behaviour of $\max_{\nu \in \mathcal{V}_N} \lambda_1(\hat{\mathbf{C}}_{\mathbf{y}}(\nu))$. For this, we have first to define some notations. We consider the fundamental function ϕ which already appears in [20]:

$$\phi(x) = \begin{cases} \frac{(x+1)(x+c)}{x} & \text{if } x > \sqrt{c} \\ \lambda^+ & \text{if } x \leq \sqrt{c} \end{cases}$$

where we recall that $\lambda^+ = (1 + \sqrt{c})^2$ (see Corollary 3). We notice that for all $x > \sqrt{c}$, $\phi(x) > \phi(\sqrt{c}) = \lambda^+$. Define as $\gamma(\nu)$ the maximum eigenvalue of the finite rank perturbation for each ν , that is

$$\gamma(\nu) = \lambda_1 \left(\mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \mathbf{H}(\nu) \mathbf{H}(\nu)^* \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \right) \quad (24)$$

and let $\nu_N^* \in \mathcal{V}_N$ such that

$$\nu_N^* \in \operatorname{argmax}_{\nu \in \mathcal{V}_N} \gamma(\nu).$$

We remark that $\gamma(\nu_N^*)$ may be interpreted as a certain SNR metric in the frequency domain. In the following, we study the behaviour of the largest eigenvalue of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu_N^*)$, which requires the following additional assumption on $\gamma(\nu_N^*)$.

Assumption 5. There exists $\gamma_\infty \geq 0$ such that

$$\gamma(\nu_N^*) \xrightarrow[M \rightarrow \infty]{} \gamma_\infty.$$

Theorem 3 implies that the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu_N^*)$ have the same asymptotic behaviour as the corresponding eigenvalues of matrix $\Xi(\nu_N^*)^{\frac{1}{2}} \frac{\mathbf{X}(\nu_N^*) \mathbf{X}(\nu_N^*)^*}{B+1} \Xi(\nu_N^*)^{\frac{1}{2}}$. Under Assumption 5, [28], [20] or [29] immediately imply the following result. Note that since ν_N^* is unknown in practice, this proposition is an intermediate theoretical result that will justify the detection test statistic introduced below.

Proposition 2. Under Assumptions 1, 2, 3, 4 and 5, we have

$$\lambda_1 \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu_N^*) \right) \xrightarrow[M \rightarrow \infty]{a.s.} \phi(\gamma_\infty) \quad (25)$$

while

$$\lambda_{K+1} \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu_N^*) \right) \xrightarrow[M \rightarrow \infty]{a.s.} \lambda^+ \quad (26)$$

and

$$\lambda_M \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu_N^*) \right) \xrightarrow[M \rightarrow \infty]{a.s.} \lambda^-. \quad (27)$$

Moreover, if $\gamma_\infty = 0$,

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1 \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu) \right) \leq \lambda^+ \quad a.s. \quad (28)$$

Proof: It just remains to establish (28), see Appendix F. ■

Since neither the intrinsic dimensionality K of the useful signal $(\mathbf{u}_n)_{n \in \mathbb{Z}}$ nor the frequency ν_N^* are known in practice, we use the largest eigenvalue of the SCM maximized over all Fourier frequencies as a test statistic. This leads to the test statistic T_ϵ defined previously in (17) which we recall here:

$$T_\epsilon = \mathbb{1}_{[\lambda^+ + \epsilon, \infty)} \left(\max_{\nu \in \mathcal{V}_N} \lambda_1 \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu) \right) \right).$$

It turns out that this test statistics is consistent in the high-dimensional regime, as stated in the following result.

Proposition 3. *Under Assumptions 1, 2, 3, 4 and 5, and if under Hypothesis \mathcal{H}_1 ,*

$$\gamma_\infty > \sqrt{c}$$

then for all $0 < \epsilon < \phi(\gamma_\infty) - \lambda^+$ and $i \in \{0, 1\}$,

$$\mathbb{P}_i \left(\lim_{M \rightarrow \infty} T_\epsilon = i \right) = 1$$

where \mathbb{P}_i is the underlying probability measure under Hypothesis \mathcal{H}_i .

Proof: Under Hypothesis \mathcal{H}_0 , since $\gamma_\infty = 0$, we directly apply (28) in Proposition 2 to obtain that for all $\epsilon > 0$, $T_\epsilon = 0$ with probability one, for all large M . Under Hypothesis \mathcal{H}_1 , we get

$$\liminf_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1 \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu) \right) \geq \lim_{M \rightarrow \infty} \lambda_1 \left(\hat{\mathbf{C}}_{\mathbf{y}}(\nu_N^*) \right) = \phi(\gamma_\infty)$$

with probability one. Since by assumption, $\phi(\gamma_\infty) > \lambda^+ + \epsilon$, we deduce that $T_\epsilon = 1$ with probability one for all large M . ■

VI. SIMULATIONS

In this section, we provide some numerical illustrations of the approximation results of Section IV. We will consider the case where the rank K of the signal is equal to one and then the case where K is strictly greater than one.

A. Case $K = 1$

As in the numerical simulation presented in Section III, each component of the noise \mathbf{v}_n is generated as a Gaussian AR(1) process with $\theta = 0.5$. The expression of its spectral density s_m for all $m = 1, \dots, M$ is still given in (18). The useful signal is generated as an AR(1) process with $K = 1$, \mathbf{H}_k defined by (19) and $\beta = \frac{10}{11}$. C is again a positive constant used to tune the SNR. Note that, in this context, the SNR $\gamma(\nu)$ at frequency ν defined in (24) takes the form

$$\gamma(\nu) = C \left| \frac{1 - \theta e^{-i2\pi\nu}}{1 - \beta e^{-i2\pi\nu}} \right|^2.$$

Figures 3 and 4 illustrate the signal-free case $C = 0$, and where $(N, M, B) = (20000, 100, 200)$. In Figure 3, we plot the histogram of the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ for $\nu = 0$. As predicted by Corollary 3 in the signal-free case, the empirical eigenvalue distribution of $\hat{\mathbf{C}}_{\mathbf{y}}(\nu)$ is well approximated by the Marcenko-Pastur distribution with shape parameter $c = 0.5 \approx M/(B + 1)$. Figure 4 further illustrates this convergence, where the cumulative distribution function (cdf) of the Marcenko-Pastur distribution is plotted against the two following quantities:

$$F_{\min}(t) = \min_{\nu \in \mathcal{V}_N} \frac{1}{M} \sum_{\lambda_i(\hat{\mathbf{C}}(\nu)) < t} \delta_{\lambda_i(\hat{\mathbf{C}}(\nu))}$$

$$F_{\max}(t) = \max_{\nu \in \mathcal{V}_N} \frac{1}{M} \sum_{\lambda_i(\hat{\mathbf{C}}(\nu)) < t} \delta_{\lambda_i(\hat{\mathbf{C}}(\nu))}.$$

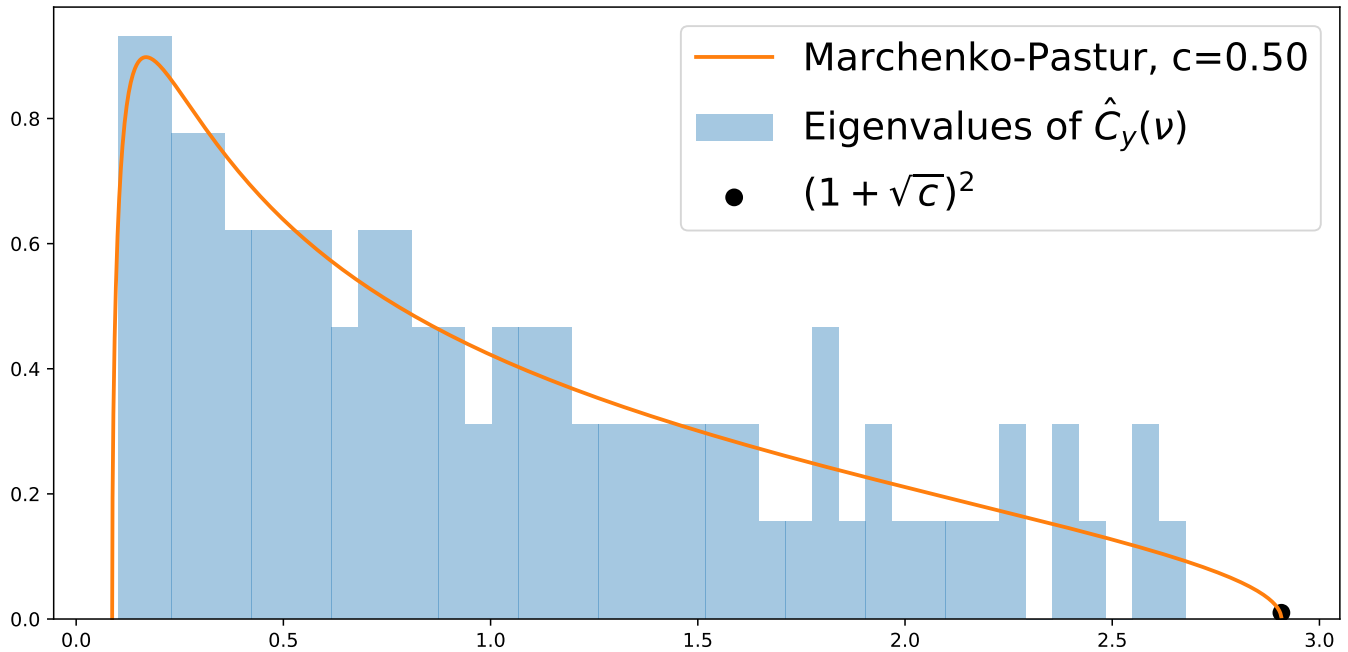


Fig. 3. Eigenvalue distribution of $\hat{C}_y(0)$ vs the density of the Marcenko-Pastur distribution with parameter $c = 1/2$.

These two functions represent the maximum deviations (from above and below) over the frequencies $\nu \in \mathcal{V}_N$ of the empirical spectral distribution of $\hat{C}(\nu)$ against the Marcenko-Pastur distribution. As suggested by the uniform convergence in the frequency domain in Corollary 3, the Marcenko-Pastur approximation in the high-dimensional regime is reliable over the whole set of Fourier frequencies. Note that the statement of Corollary 3 does not exactly match the setting used in Figure 4, as the test function used here is not in $C^1((0, +\infty))$.

To illustrate the signal-plus-noise case and the results of Corollary 3 and Proposition 2, we plot in Figure 5, the histogram of the eigenvalues of $\hat{C}_y(\nu)$ for $\nu = 0$, with $\gamma(0) = 2.9$. We see that the largest eigenvalue deviates from the right edge $(1 + \sqrt{c})^2$ and is located around the value $\phi(\gamma(0)) = 4.5$, as predicted by Proposition 2, while all the other eigenvalues spread as the Marcenko-Pastur distribution, as predicted by Corollary 3.

In order to compare the test statistic (17) with other frequency domain methods based on the SCM, we consider:

- the new test statistic (17), denoted as LE (for largest eigenvalue),
- two tests based on LSS of the SCM given by

$$T'_\epsilon = \mathbb{1}_{[\epsilon, +\infty)} \left(\max_{\nu \in \mathcal{V}_N} \left| L_\varphi(\nu) - \int_{\mathbb{R}} \varphi(\lambda) f(\lambda) d\lambda \right| \right)$$

where L_φ and density f are defined in (23) and Corollary 3 respectively, and with $\varphi(x) = (x - 1)^2$ for the Frobenius norm test (denoted as LSS Frob.) and $\varphi(x) = \log(x)$ for the logdet test (denoted as LSS logdet),

- a test statistic based on the largest off-diagonal entry of the SCM:

$$T''_\epsilon = \mathbb{1}_{[\epsilon, +\infty)} \left(\max_{\nu \in \mathcal{V}_N} \max_{\substack{i,j=1,\dots,M \\ i < j}} \left| [\hat{C}_y(\nu)]_{i,j} \right| \right)$$

denoted as MCC (for Maximum of Cross Coherence),

and where $\epsilon > 0$ is some threshold. In Table I, we provide, via Monte-Carlo simulations (10000 draws), the power of each of the four tests, calibrated so that the empirical type I error is equal to 0.05. The results are provided for various values of (N, M, B) chosen so that $M \in \{20, 40, \dots, 180\}$, $N = M^2$ and $B = 2M$. We set the SNR in the frequency domain as $\max_{\nu \in \mathcal{V}_N} \gamma(\nu) = 2\sqrt{\frac{M}{B}} = 1.41$.

The LE test presents the best detection performance among the four candidates, whereas the MCC test does not seem to be adapted to the detection of this alternative. While it is proved in Corollary 3 that the test statistics based on the LSS of

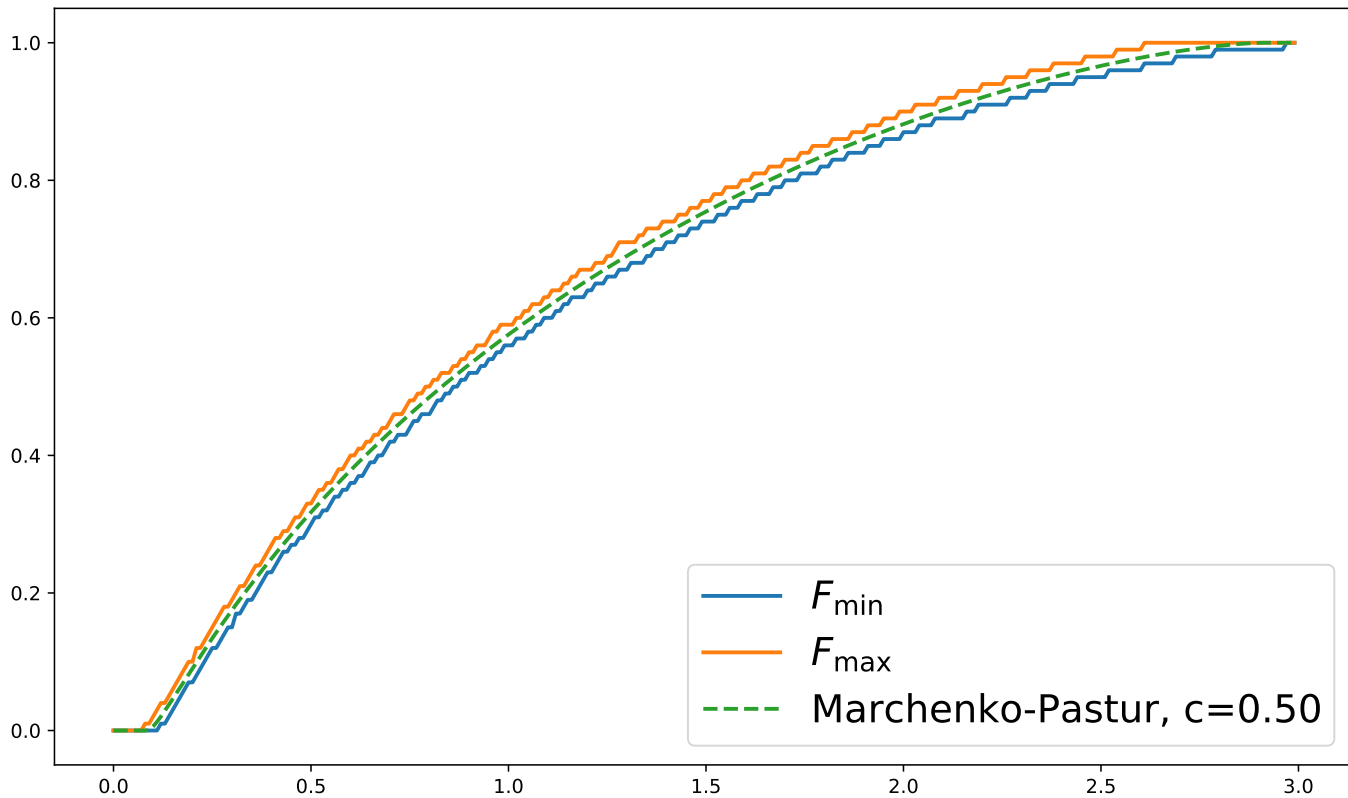


Fig. 4. Uniform convergence of the eigenvalue distribution of $\hat{C}_y(\nu)$ over $\nu \in \mathcal{V}_N$ toward the Marcenko-Pastur distribution with parameter $c = 1/2$.

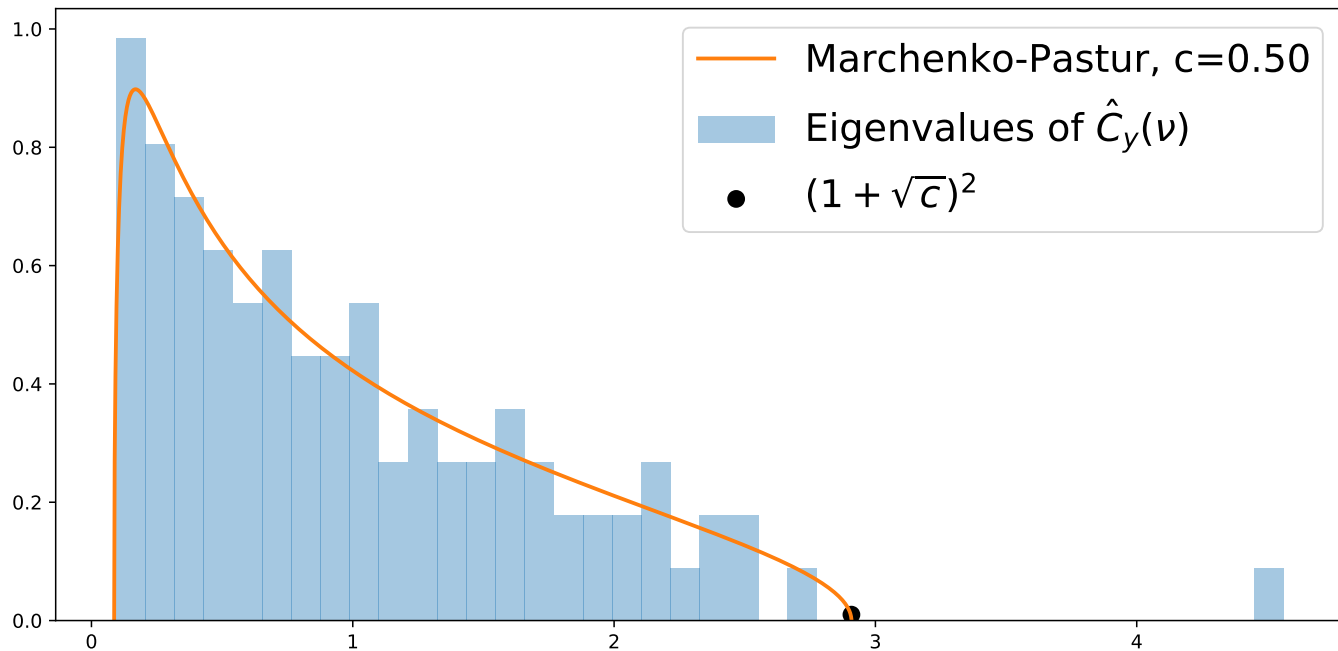


Fig. 5. Eigenvalue distribution of $\hat{C}(\nu)$ vs Marcenko-Pastur distribution with parameter $c = 1/2$ in the signal case.

$\hat{C}(\nu)$ can not asymptotically distinguish between \mathcal{H}_0 and \mathcal{H}_1 , they remain sensible to a large variation of a single eigenvalue

TABLE I
POWER COMPARISON, $K=1$, $\gamma(\nu_N^*) = 2\sqrt{\frac{1}{2}}$, TYPE I ERROR = 5%

N	M	B	LSS Frob.	LSS logdet	MCC	LE
400	20	40	0.09	0.07	0.06	0.15
1600	40	80	0.15	0.08	0.06	0.37
3600	60	120	0.19	0.08	0.06	0.68
6400	80	160	0.25	0.08	0.06	0.87
10000	100	200	0.26	0.07	0.06	0.96
14400	120	240	0.25	0.06	0.06	0.99
19600	140	280	0.28	0.06	0.06	1.00
25600	160	320	0.30	0.06	0.06	1.00
32400	180	360	0.31	0.06	0.06	1.00

for finite values of M . Consider for instance the Frobenius LSS test, where the test statistic is based on :

$$\max_{\nu \in \mathcal{V}_N} \left| \frac{1}{M} \sum_{m=1}^M (\lambda_m(\hat{\mathbf{C}}(\nu)) - 1)^2 - \int (\lambda - 1)^2 f(\lambda) d\lambda \right|$$

where an explicit computation shows that $\int (\lambda - 1)^2 f(\lambda) d\lambda = c$. An $\mathcal{O}(1)$ variation of $\lambda_1(\hat{\mathbf{C}}(\nu))$, the largest eigenvalue of $\hat{\mathbf{C}}(\nu)$, will lead to a variation of order $\mathcal{O}(\frac{1}{M})$ of the above term. Therefore, the power of a LSS based test asymptotically converge towards zero, while having non-zero power for finite values of M as it is visible on the results of Table I.

B. Case $K > 1$

We eventually consider a model which have the flexibility to consider a signal with an arbitrary value of $K \geq 1$. We assume that matrices $(\mathbf{H}_l)_{l \geq 0}$ verify $\mathbf{H}_l = 0$ if $l > L$ for a certain integer L , and that the sequence of $M \times K$ matrices $(\mathbf{H}_l)_{0 \leq l \leq L}$ is defined by:

$$\mathbf{H}_l = (C_1 \mathbf{w}_{l,1}, \dots, C_K \mathbf{w}_{l,K})$$

where the vectors $((\mathbf{w}_{l,k})_{l=0, \dots, L})_{k=1, \dots, K}$ are generated as independent realisations of M -dimensional vectors uniformly distributed on the unit sphere of \mathbb{C}^M and where the $C_1 \geq C_2 \geq \dots \geq C_K$ are positive constants used to tune the SNR of each of the K sources at the desired level. Moreover, as the K columns of each matrix \mathbf{H}_l coincide with the realisations of mutually independent random vectors, the columns of $\mathbf{H}(\nu)$ are easily seen to be nearly orthogonal and to nearly share the same norm for each ν if M is large enough. More precisely, for each ν , it holds that $\mathbf{H}(\nu)^* \mathbf{H}(\nu) \rightarrow (L+1) \text{Diag}(C_1, \dots, C_K)$ when $M \rightarrow +\infty$. As the spectral densities of the components of the noise all coincide with $s(\nu) = \frac{1}{|1 - \theta e^{-2i\pi\nu}|^2}$, the non-zero eigenvalues of $\mathbf{S}_\nu(\nu)^{-\frac{1}{2}} \mathbf{H}(\nu) \mathbf{H}(\nu)^* \mathbf{S}_\nu(\nu)^{-\frac{1}{2}}$ converge towards the $((L+1)C_k/s(\nu))_{k=1, \dots, K}$ when M increases. Therefore, the signal obtained by this model satisfies Assumption 3. Rather than just providing the performance of the test T_ϵ based on the maximum of the largest eigenvalue of $\hat{\mathbf{C}}_y(\nu)$ proposed in this paper, we compare in the following T_ϵ with $T_{K,\epsilon}$ defined by

$$T_{K,\epsilon} = \mathbb{1}_{[K\lambda^+ + \epsilon, \infty)} \left(\max_{\nu \in \mathcal{V}_N} \sum_{k=1}^K \lambda_k \left(\hat{\mathbf{C}}_y(\nu) \right) \right)$$

which depends on the K largest eigenvalues of $\hat{\mathbf{C}}_y(\nu)$ rather than on the largest one. It is easy to generalize Proposition 2 and Proposition 3 in order to study the asymptotic properties of $T_{K,\epsilon}$. More precisely, for each $k = 1, \dots, K$, we define $\gamma_k(\nu)$ by

$$\gamma_k(\nu) = \lambda_k \left(\mathbf{S}_\nu(\nu)^{-\frac{1}{2}} \mathbf{H}(\nu) \mathbf{H}(\nu)^* \mathbf{S}_\nu(\nu)^{-\frac{1}{2}} \right) \quad (29)$$

and denote $\nu_{K,N}^*$ one of the frequency such that $\max_{\nu \in \mathcal{V}_N} \sum_{k=1}^K \gamma_k(\nu) = \sum_{k=1}^K \gamma_k(\nu_{K,N}^*)$. $\gamma_k(\nu)$ can of course be seen as a generalization of $\gamma(\nu)$ defined by (24). Then, under the extra assumption that for $k = 1, \dots, K$, $\gamma_k(\nu_{K,N}^*)$ converges towards a finite limit $\gamma_{k,\infty}$ (a condition which holds in the context of the present experiment because it is easily seen that $\gamma_k(\nu_{K,N}^*) \rightarrow (L+1)(1+\theta)^2 C_k$), $\lambda_k \left(\hat{\mathbf{C}}_y(\nu_{K,N}^*) \right)$ converges towards λ^+ if $\gamma_{k,\infty} \leq \sqrt{c}$ and towards $\phi(\gamma_{k,\infty}) > \lambda^+$ if $\gamma_{k,\infty} > \sqrt{c}$. It is easy to check that if $\gamma_\infty = \gamma_{1,\infty} > \sqrt{c}$, then the statistics $T_{K,\epsilon}$ also leads to a consistent test provided $0 < \epsilon < \phi(\gamma_\infty) - \lambda_+$. While in practice the number of sources K is unknown, it is interesting to evaluate the performance provided by $T_{K,\epsilon}$ which can be considered as an ideal reference. Intuitively, $T_{K,\epsilon}$ could lead to a better performance than T_ϵ when $\gamma_{k,\infty} > \sqrt{c}$ for $k = 1, \dots, K$, because, in this context, if $\hat{\nu}_{K,N}^*$ is a frequency that maximises $\sum_{k=1}^K \lambda_k \left(\hat{\mathbf{C}}_y(\nu) \right)$, then $\liminf_{M \rightarrow +\infty} \lambda_k \left(\hat{\mathbf{C}}_y(\hat{\nu}_{K,N}^*) \right) > \lambda^+$. Therefore, the K largest eigenvalues of $\hat{\mathbf{C}}_y(\hat{\nu}_{K,N}^*)$ bring useful information to the

TABLE II
POWER COMPARISON, $\frac{C_1}{C_2} = 4$, $(\gamma_1 + \gamma_2)(\nu_{2,N}^*) = 3\sqrt{\frac{1}{2}}$, TYPE I ERROR = 5%

N	M	B	LSS Fr.	LSS ld	MCC	LE(1)	LE(2)
100	10	20	0.31	0.18	0.16	0.42	0.37
400	20	40	0.79	0.39	0.45	0.94	0.89
900	30	60	0.94	0.49	0.53	1.00	0.99
1600	40	80	0.98	0.50	0.55	1.00	1.00
2500	50	100	0.99	0.52	0.55	1.00	1.00
3600	60	120	1.00	0.51	0.43	1.00	1.00
4900	70	140	1.00	0.55	0.37	1.00	1.00
6400	80	160	1.00	0.54	0.28	1.00	1.00

TABLE III
POWER COMPARISON, $\frac{C_1}{C_2} = 1$, $(\gamma_1 + \gamma_2)(\nu_{2,N}^*) = 3\sqrt{\frac{1}{2}}$, TYPE I ERROR = 5%

N	M	B	LSS Fr.	LSS ld	MCC	LE(1)	LE(2)
100	10	20	0.38	0.22	0.16	0.48	0.46
400	20	40	0.58	0.30	0.30	0.75	0.73
900	30	60	0.67	0.30	0.28	0.91	0.89
1600	40	80	0.74	0.29	0.18	0.96	0.97
2500	50	100	0.79	0.30	0.16	0.99	0.99
3600	60	120	0.79	0.24	0.13	1.00	1.00
4900	70	140	0.85	0.28	0.12	1.00	1.00

detection of the useful signal.

In order to evaluate numerically the compared performance of T_ϵ and $T_{K,\epsilon}$ when K is known, we first consider the case $K = 2$, $L = 3$, and where $(\gamma_1 + \gamma_2)(\nu_{2,N}^*) = 3\sqrt{c}$. Concerning the value of (C_1, C_2) , we consider the two following cases: $\frac{C_1}{C_2} = 1$ and $\frac{C_1}{C_2} = 4$. This corresponds respectively to the case where both sources contributes exactly the same on each sensor, and where the first source contributes much more than the second one. Tables II, III report the power of the proposed test (LE(1) represents T_ϵ and LE(2) represents $T_{2,\epsilon}$) against the LSS tests and the MCC test, with a type I error fixed at 5%. When $\frac{C_1}{C_2} = 4$, it can be expected that the most powerful source is dominant, and that $\gamma_2(\nu_{2,N}^*) < \sqrt{c}$. Therefore, $\lambda_2(\hat{C}_y(\nu))$ is likely to stay close to λ^+ for each ν , so that the use of $T_{2,\epsilon}$ should not bring any extra performance. This intuition is confirmed by Table II. When $\frac{C_1}{C_2} = 1$, $\gamma_1(\nu_{2,N}^*)$ and $\gamma_2(\nu_{2,N}^*)$ should be both close to $\frac{3}{2}\sqrt{c}$, thus suggesting that the two largest eigenvalues of \hat{C}_y at the maximizing frequency $\hat{\nu}_{2,N}^*$ should also nearly coincide, and should escape from $[\lambda^-, \lambda^+]$. While the second eigenvalue brings here some information, Table III tends to indicate that T_ϵ has better performance than $T_{2,\epsilon}$. In the next experiment, $(\gamma_1 + \gamma_2)(\nu_{2,N}^*) = 2\sqrt{c}$. For $\frac{C_1}{C_2} = 4$, the largest eigenvalue of $\hat{C}_y(\nu)$ is likely to be still dominant for each ν , and Table IV confirms the better performance of T_ϵ . When $\frac{C_1}{C_2} = 1$, $\gamma_1(\nu_{2,N}^*)$ and $\gamma_2(\nu_{2,N}^*)$ should be both close to the detectability threshold \sqrt{c} , and Table V this time shows that the use of $T_{2,\epsilon}$ leads to some improvement. For comparison, we also report the results of T_ϵ for $C_2 = 0$ in Table VI.

This discussion tends to indicate that, even when $K > 1$ is assumed known, the use of the maximum over \mathcal{V}_N of the largest eigenvalue of $\hat{C}_y(\nu)$ does not introduce any significant loss of performance.

TABLE IV
POWER COMPARISON, $\frac{C_1}{C_2} = 4$, $(\gamma_1 + \gamma_2)(\nu_{2,N}^*) = 2\sqrt{\frac{1}{2}}$, TYPE I ERROR = 5%

N	M	B	LSS Fr.	LSS ld	MCC	LE(1)	LE(2)
100	10	20	0.15	0.10	0.10	0.21	0.20
400	20	40	0.33	0.15	0.12	0.55	0.50
900	30	60	0.39	0.15	0.17	0.75	0.71
1600	40	80	0.52	0.16	0.14	0.94	0.90
2500	50	100	0.54	0.15	0.14	0.98	0.97
3600	60	120	0.56	0.13	0.13	1.00	0.99
4900	70	140	0.55	0.13	0.10	1.00	1.00
6400	80	160	0.62	0.11	0.10	1.00	1.00

TABLE V
POWER COMPARISON, $\frac{C_1}{C_2} = 1$, $(\gamma_1 + \gamma_2)(\nu_{2,N}^*) = 2\sqrt{\frac{1}{2}}$, TYPE I ERROR = 5%

N	M	B	LSS Fr.	LSS ld	MCC	LE(1)	LE(2)
400	20	40	0.17	0.11	0.08	0.27	0.27
1600	40	80	0.18	0.10	0.08	0.45	0.48
3600	60	120	0.15	0.07	0.07	0.58	0.62
6400	80	160	0.16	0.07	0.08	0.69	0.75
10000	100	200	0.13	0.05	0.07	0.76	0.83
14400	120	240	0.10	0.03	0.07	0.82	0.86
19600	140	280	0.09	0.04	0.07	0.86	0.89
25600	160	320	0.10	0.03	0.06	0.89	0.93
32400	180	360	0.09	0.03	0.06	0.87	0.93

TABLE VI
POWER COMPARISON, $C_2 = 0$, $\gamma(\nu_N^*) = 2\sqrt{\frac{1}{2}}$, TYPE I ERROR = 5%

N	M	B	LSS Fr.	LSS ld	MCC	LE(1)	LE(2)
100	10	20	0.19	0.12	0.12	0.26	0.22
400	20	40	0.43	0.19	0.14	0.66	0.59
900	30	60	0.51	0.19	0.19	0.88	0.83
1600	40	80	0.62	0.20	0.15	0.97	0.95
2500	50	100	0.65	0.18	0.17	0.99	0.99
3600	60	120	0.68	0.16	0.12	1.00	1.00
4900	70	140	0.71	0.16	0.13	1.00	1.00
6400	80	160	0.75	0.17	0.12	1.00	1.00

VII. CONCLUSION

In this paper, we have studied the statistical behaviour of certain frequency-domain detection test statistics, based on the eigenvalues of a sample estimate of the SCM, in the high-dimensional regime in which both the dimension M of the underlying signals and the number of samples N converge to infinity at certain rates. In particular, we have proved various approximation results showing that the sample SCM asymptotically behaves as a Wishart matrix. These results have been exploited to prove that test statistics based on LSS of the sample SCM are not consistent in the high-dimensional regime. A new test statistic relying on the largest eigenvalue of the sample SCM has also been proposed and proved to be consistent in the high-dimensional regime. Finally, numerical results have demonstrated that this new test statistic provides reasonable performance and outperforms other standard test statistics in situations where the dimension M and the number of samples N are large.

REFERENCES

- [1] A. Rosuel, P. Vallet, P. Loubaton, and X. Mestre, "On the frequency domain detection of high dimensional time series," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 8782–8786.
- [2] D. Ramírez, G. Vázquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-p signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, 2011.
- [3] J. Sala-Alvarez, G. Vázquez-Vilar, R. López-Valcarce, S. Sedighi, and A. Taherpour, "Multiantenna glr detection of rank-one signals with known power spectral shape under spatially uncorrelated noise," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6269–6283, 2016.
- [4] A.-J. Boonstra and A.-J. Van der Veen, "Gain calibration methods for radio telescope arrays," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 25–38, 2003.
- [5] A. Leshem and A.-J. Van der Veen, "Multichannel detection of gaussian signals with uncalibrated receivers," *IEEE Signal Process. Lett.*, vol. 8, no. 4, pp. 120–122, 2001.
- [6] L. D. Haugh, "Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach," *Journal of the American Statistical Association*, vol. 71, no. 354, pp. 378–385, 1976.
- [7] Y. Hong, "Testing for independence between two covariance stationary time series," *Biometrika*, vol. 83, no. 3, pp. 615–625, 1996.
- [8] W. Li and Y. Hui, "Robust residual cross correlation tests for lagged relations in time series," *Journal of Statistical Computation and Simulation*, vol. 49, no. 1-2, pp. 103–109, 1994.
- [9] K. El Himdi, R. Roy, and P. Duchesne, "Tests for non-norrelation of two multivariate time series: A nonparametric approach," *Lecture Notes-Monograph Series*, vol. 42, pp. 397–416, 2003.
- [10] D. Ramírez, J. Vía, I. Santamaría, and L. Scharf, "Detection of spatially correlated gaussian time series," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5006–5015, 2010.
- [11] N. Klausner, M. Azimi-Sadjadi, and L. Scharf, "Detection of spatially correlated time series from a network of sensor arrays," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1396–1407, 2014.
- [12] D. Brillinger, *Time series: data analysis and theory*. Classics in Applied Mathematics, Siam, 2001, vol. 36.
- [13] L. Koopmans, *The spectral analysis of time series*. Probability and Mathematical Statistics, Academic Press, 1995, vol. 22.
- [14] G. Wahba, "Some tests of independence for stationary multivariate time series," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 33, no. 1, pp. 153–166, 1971.

- [15] M. Taniguchi, M. L. Puri, and M. Kondo, "Nonparametric approach for non-gaussian vector stationary processes," *Journal of Multivariate Analysis*, vol. 56, no. 2, pp. 259–283, 1996.
- [16] M. Eichler, "A frequency-domain based test for non-correlation between stationary time series," *Metrika*, vol. 65, no. 2, pp. 133–157, 2007.
- [17] —, "Testing nonparametric and semiparametric hypotheses in vector stationary processes," *Journal of Multivariate Analysis*, vol. 99, no. 5, pp. 968–1009, 2008.
- [18] P. Loubaton and A. Rosuel, "Large random matrix approach for testing independence of a large number of gaussian time series," *arXiv preprint arXiv:2007.08806*, 2020.
- [19] V. Marcenko and L. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, p. 457, 1967.
- [20] J. Baik and J. Silverstein, "Eigenvalues of large sample covariance matrices of spiked population models," *J. Multivariate Anal.*, vol. 97, no. 6, pp. 1382–1408, 2006.
- [21] D. Morales-Jimenez, I. Johnstone, M. MacKay, and J. Yang, "Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models," *To appear in Statistica Sinica*, 2019, preprint arXiv:1810.10214v3.
- [22] M. Forni, M. Hallin, M. Lippi, and L. Reichlin, "The generalized dynamic-factor model: Identification and estimation," *Rev. Econ. Stat.*, vol. 82, no. 4, pp. 540–554, 2000.
- [23] —, "The generalized dynamic factor model consistency and rates," *J. Econom.*, vol. 119, no. 2, pp. 231–255, 2004.
- [24] P. Vallet, P. Loubaton, and X. Mestre, "Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, 2012.
- [25] P. Vallet, X. Mestre, and P. Loubaton, "Performance Analysis of an Improved MUSIC DoA Estimator," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6407–6422, Dec. 2015.
- [26] R. Horn and C. Johnson, *Matrix analysis*. Cambridge university press, 2005.
- [27] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 33, no. 2, pp. 387–392, 1985.
- [28] J. Baik, G. B. Arous, S. Péché *et al.*, "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices," *Annals of Probability*, vol. 33, no. 5, pp. 1643–1697, 2005.
- [29] F. Benaych-Georges and R. R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices," *Advances in Mathematics*, vol. 227, no. 1, pp. 494–521, 2011.
- [30] M. Rudelson and R. Vershynin, "Hanson-wright inequality and sub-gaussian concentration," *Electron. Commun. Probab.*, vol. 18, 2013.
- [31] U. Haagerup and S. Thorbjørnsen, "Random matrices with complex gaussian entries," *Expo. Math.*, vol. 21, no. 4, pp. 293–337, 2003.
- [32] T. Tao, *Topics in random matrix theory*. American Mathematical Soc., 2012, vol. 132.
- [33] A. Guionnet and O. Zeitouni, "Concentration of the spectral measure for large matrices," *Electron. Commun. Probab.*, vol. 5, pp. 119–136, 2000. [Online]. Available: <https://doi.org/10.1214/ECP.v5-1026>
- [34] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, 2008.

APPENDIX A USEFUL RESULTS

In this section, we recall some useful results which will be constantly used in the proofs developed in the following sections.

The first result is based on a Chernoff bound for the χ^2 distribution, and is also a special case of the well-known Hanson-Wright inequality describing the concentration of sub-Gaussian quadratic forms around their means (see [30]).

Lemma 1. *Let $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}^n}(\mathbf{0}, \mathbf{I}_n)$ and Ξ a deterministic $n \times n$ complex matrix. Then there exists a constant $\kappa > 0$ independent of n and Ξ such that for all $t \geq 0$,*

$$\mathbb{P}(|\mathbf{z}^* \Xi \mathbf{z} - \mathbb{E}[\mathbf{z}^* \Xi \mathbf{z}]| > t) \leq 2 \exp\left(-\kappa \min\left\{\frac{t^2}{\|\Xi\|_F^2}, \frac{t}{\|\Xi\|_2}\right\}\right).$$

The second following result describes the behaviour of the largest and smallest eigenvalues of a standard Wishart matrix.

Lemma 2 ([31, Proof of Lemma 7.3]). *Let \mathbf{Z} be a $M \times (B+1)$ matrix with i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$ entries. Then under Assumption 4, there exists a constant $C > 0$ independent of M, B such that for all $t > 0$,*

$$\mathbb{P}\left(\lambda_1\left(\frac{\mathbf{Z}\mathbf{Z}^*}{B+1}\right) > \left(1 + \sqrt{\frac{M}{B+1}}\right)^2 + t\right) \leq (B+1)\exp(-C(B+1)t^2)$$

and

$$\mathbb{P}\left(\lambda_M\left(\frac{\mathbf{Z}\mathbf{Z}^*}{B+1}\right) < \left(1 - \sqrt{\frac{M}{B+1}}\right)^2 - t\right) \leq (B+1)\exp(-C(B+1)t^2).$$

We will mainly use Lemma 2 as follows; let $(\mathbf{Z}(\nu))_{\nu \in \mathcal{V}_N}$ be a family of $M \times (B+1)$ random matrices such that $\mathbf{Z}(\nu)$ has i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$ (recall the definition of the index set \mathcal{V}_N in (16)), then from the union bound

$$\begin{aligned} & \mathbb{P}\left(\max_{\nu \in \mathcal{V}_N} \lambda_1\left(\frac{\mathbf{Z}(\nu)\mathbf{Z}(\nu)^*}{B+1}\right) > \left(1 + \sqrt{\frac{M}{B+1}}\right)^2 + t\right) \\ & \leq \sum_{\nu \in \mathcal{V}_N} \mathbb{P}\left(\lambda_1\left(\frac{\mathbf{Z}(\nu)\mathbf{Z}(\nu)^*}{B+1}\right) > \left(1 + \sqrt{\frac{M}{B+1}}\right)^2 + t\right) \\ & \leq \exp\left(-C(B+1)t^2 + \log(N(B+1))\right). \end{aligned}$$

Using Assumption 4 and Borel-Cantelli lemma, we deduce that

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1 \left(\frac{\mathbf{Z}(\nu) \mathbf{Z}(\nu)^*}{B+1} \right) \leq (1 + \sqrt{c})^2$$

with probability one.

APPENDIX B PROOF OF THEOREM 2

A. Reduction to $K = 1$

First, note that we may assume $K = 1$ without loss of generality. Indeed, consider the decomposition

$$\mathbf{u}_n = \sum_{\ell=1}^K \mathbf{u}_n^{(\ell)}$$

where $\mathbf{u}_n^{(\ell)} = \sum_{k=0}^{+\infty} \mathbf{h}_{\ell,k} \epsilon_{n-k}^{(\ell)}$ and where $\mathbf{h}_{\ell,k}$ and $\epsilon_n^{(\ell)}$ are the ℓ -th column of \mathbf{H}_k and the ℓ -th entry of ϵ_n respectively. Moreover, Assumption 3 implies that

$$\sup_{M \geq 1} \sum_{k \in \mathbb{Z}} (1 + |k|) \|\mathbf{h}_{\ell,k}\|_2 < \infty$$

From the fact that K is fixed with respect to N (Assumption 4) and

$$\max_{\nu \in \mathcal{V}_N} \|\Sigma_{\mathbf{u}}(\nu) - \mathbf{H}(\nu) \Sigma_{\epsilon}(\nu)\|_2 \leq \sum_{\ell=1}^K \max_{\nu \in \mathcal{V}_N} \|\Sigma_{\mathbf{u}^{(\ell)}}(\nu) - \mathbf{h}_{\ell}(\nu) \Sigma_{\epsilon^{(\ell)}}(\nu)\|_2$$

where $\Sigma_{\mathbf{u}^{(\ell)}}(\nu)$, $\mathbf{h}_{\ell}(\nu)$, $\Sigma_{\epsilon^{(\ell)}}(\nu)$ are defined as $\Sigma_{\mathbf{u}}(\nu)$, $\mathbf{H}(\nu)$, $\Sigma_{\epsilon}(\nu)$ respectively, Theorem 2 is proved if we can show that

$$\max_{\nu \in \mathcal{V}_N} \|\Sigma_{\mathbf{u}^{(\ell)}}(\nu) - \mathbf{h}_{\ell}(\nu) \Sigma_{\epsilon^{(\ell)}}(\nu)\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0$$

for all $\ell = 1, \dots, K$. Therefore, we assume for the remainder of the proof that

$$\mathbf{u}_n = \sum_{k \in \mathbb{Z}} \mathbf{h}_k \epsilon_{n-k},$$

where

- $(\mathbf{h}_k)_{k \in \mathbb{Z}}$ is a filter, with $\mathbf{h}_k \in \mathbb{C}^M$ and such that

$$\sup_{M \geq 1} \sum_{k \in \mathbb{Z}} (1 + |k|) \|\mathbf{h}_k\|_2 < \infty. \quad (30)$$

- $(\epsilon_n)_{n \in \mathbb{Z}}$ is a scalar standard complex Gaussian white noise.

B. Reduction to $B = 1$

Let $\mathbf{h}(\nu) = \sum_{k \in \mathbb{Z}} \mathbf{h}_k e^{-i2\pi\nu k}$ and

$$\xi_{\epsilon}(\nu) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \epsilon_n e^{-i2\pi\nu n}.$$

From (30) and Assumption 4, a first-order Taylor expansion of $b \mapsto \mathbf{h}(\nu + \frac{b}{N})$ at 0 leads to

$$\sup_{\nu \in [0,1]} \max_{b \in \{-\frac{B}{2}, \dots, \frac{B}{2}\}} \left\| \mathbf{h}(\nu) - \mathbf{h}\left(\nu + \frac{b}{N}\right) \right\|_2 = \mathcal{O}\left(\frac{B}{N}\right) = \mathcal{O}\left(\frac{1}{N^{1-\alpha}}\right).$$

Moreover, from Lemma 1 applied to the random vector

$$\mathbf{z} = \left(\xi_{\epsilon}\left(\nu - \frac{B}{2N}\right), \dots, \xi_{\epsilon}\left(\nu + \frac{B}{2N}\right) \right)^T \sim \mathcal{N}_{\mathbb{C}^{B+1}}(\mathbf{0}, \mathbf{I}_{B+1})$$

and matrix $\Xi = \frac{\mathbf{I}_{B+1}}{B+1}$, there exists some constant κ independent of M such that for all $t \geq 2$,

$$\mathbb{P}\left(\max_{\nu \in \mathcal{V}_N} \frac{1}{B+1} \sum_{b=-B/2}^{B/2} \left| \xi_{\epsilon}\left(\nu + \frac{b}{N}\right) \right|^2 > t\right) \leq N \mathbb{P}\left(\frac{1}{B+1} \sum_{b=-B/2}^{B/2} \left| \xi_{\epsilon}\left(\frac{b}{N}\right) \right|^2 > t\right) \leq N \exp(-\kappa B)$$

and Borel-Cantelli lemma together with Assumption 4 imply

$$\max_{\nu \in \mathcal{V}_N} \frac{1}{B+1} \sum_{b=-B/2}^{B/2} \left| \xi_\epsilon \left(\nu + \frac{b}{N} \right) \right|^2 = \mathcal{O}(1)$$

with probability one. Defining

$$\Sigma_\epsilon(\nu) = \frac{1}{\sqrt{B+1}} \left(\xi_\epsilon \left(\nu - \frac{B}{2N} \right), \dots, \xi_\epsilon \left(\nu + \frac{B}{2N} \right) \right)$$

as well as

$$\Phi(\nu) = \frac{1}{\sqrt{B+1}} \left[\phi \left(\nu - \frac{B}{2N} \right), \dots, \phi \left(\nu + \frac{B}{2N} \right) \right]$$

with $\phi(\nu) = \mathbf{h}(\nu)\xi_\epsilon(\nu)$, we therefore have the control

$$\begin{aligned} \max_{\nu \in \mathcal{V}_N} \|\mathbf{h}(\nu)\Sigma_\epsilon(\nu) - \Phi(\nu)\|_2 &\leq \sup_{\nu \in [0,1]} \max_{b \in \{-\frac{B}{2}, \dots, \frac{B}{2}\}} \left\| \mathbf{h}(\nu) - \mathbf{h} \left(\nu + \frac{b}{N} \right) \right\|_2 \sqrt{\max_{\nu \in \mathcal{V}_N} \frac{1}{B+1} \sum_{b=-B/2}^{B/2} \left| \xi_\epsilon \left(\nu + \frac{b}{N} \right) \right|^2} \\ &= \mathcal{O} \left(\frac{1}{N^{1-\alpha}} \right) \text{ a.s.} \\ &\xrightarrow[M \rightarrow \infty]{a.s.} 0. \end{aligned}$$

Finally, since the spectral norm of a matrix is bounded by its Frobenius norm,

$$\begin{aligned} \max_{\nu \in \mathcal{V}_N} \|\Sigma_{\mathbf{u}}(\nu) - \Phi(\nu)\|_2 &\leq \sqrt{\frac{1}{B+1} \sum_{b=-B/2}^{B/2} \left\| \xi_{\mathbf{u}} \left(\nu + \frac{b}{N} \right) - \phi \left(\nu + \frac{b}{N} \right) \right\|_2^2} \\ &\leq \max_{\nu \in \mathcal{V}_N} \|\xi_{\mathbf{u}}(\nu) - \phi(\nu)\|_2. \end{aligned}$$

Theorem 1 is proven if we show that

$$\max_{\nu \in \mathcal{V}_N} \|\xi_{\mathbf{u}}(\nu) - \mathbf{h}(\nu)\xi_\epsilon(\nu)\|_2 \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

C. Periodization

For all integer n , let $[n]$ denotes the integer contained in $\{0, \dots, N-1\}$ such that $[n] \equiv n \pmod{N}$ and define

$$\tilde{\mathbf{u}}_n = \sum_{k \in \mathbb{Z}} \mathbf{h}_k \epsilon_{[n-k]}$$

where $(\tilde{\mathbf{u}}_n)_{n \in \mathbb{Z}}$ represents the circular convolution between $(\mathbf{h}_k)_{k \in \mathbb{Z}}$ and $(\epsilon_n)_{n \in \mathbb{Z}}$. If $\xi_{\tilde{\mathbf{u}}}(\nu) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \tilde{\mathbf{u}}_n e^{-i2\pi n \nu}$, then the equality

$$\xi_{\tilde{\mathbf{u}}}(\nu) = \mathbf{h}(\nu)\xi_\epsilon(\nu)$$

holds for all $\nu \in \mathcal{V}_N$. It is straightforward to check that

$$\xi_{\tilde{\mathbf{u}}}(\nu) - \xi_{\mathbf{u}}(\nu) = \delta(\nu) + \check{\delta}(\nu)$$

where

$$\begin{aligned} \delta(\nu) &= \frac{1}{\sqrt{N}} \sum_{k=1}^{N-1} \mathbf{h}_k \sum_{p=1}^k (\epsilon_{[-p]} - \epsilon_{-p}) e^{-i2\pi \nu(k-p)} \\ &\quad + \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \mathbf{h}_k \sum_{p=0}^{N-1} (\epsilon_{[p-k]} - \epsilon_{p-k}) e^{-i2\pi \nu p} \end{aligned}$$

and

$$\begin{aligned} \check{\delta}(\nu) &= \frac{1}{\sqrt{N}} \sum_{k=1}^{N-1} \mathbf{h}_{-k} \sum_{p=1}^k (\epsilon_{[N+p-1]} - \epsilon_{N+p-1}) e^{-i2\pi \nu(N-1+p-k)} \\ &\quad + \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \mathbf{h}_{-k} \sum_{p=0}^{N-1} (\epsilon_{[p+k]} - \epsilon_{p+k}) e^{-i2\pi \nu p} \end{aligned}$$

Theorem 2 is proved if we can show that

$$\max_{\nu \in \mathcal{V}_N} \|\delta(\nu)\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (31)$$

and

$$\max_{\nu \in \mathcal{V}_N} \|\check{\delta}(\nu)\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0. \quad (32)$$

In the remainder, we only prove (31) and omit the details for (32) whose treatment is similar. To that end, we define

$$\begin{aligned} \delta_1(\nu) &= \frac{1}{\sqrt{N}} \sum_{k=1}^{N-1} \mathbf{h}_k \sum_{p=1}^k (\epsilon_{[-p]} - \epsilon_{-p}) e^{-i2\pi\nu(k-p)} \\ \delta_2(\nu) &= \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \mathbf{h}_k \sum_{p=0}^{N-1} (\epsilon_{[p-k]} - \epsilon_{p-k}) e^{-i2\pi\nu p}. \end{aligned}$$

D. Control of $\delta_1(\nu)$

For $p \in \{1, \dots, N-1\}$, let

$$z_p(\nu) = (\epsilon_{[-p]} - \epsilon_{-p}) e^{i2\pi\nu p} = (\epsilon_{N-p} - \epsilon_{-p}) e^{i2\pi\nu p}.$$

Then $z_1(\nu), \dots, z_{N-1}(\nu)$ are i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 2)$ and by rearranging the sums in $\delta_1(\nu)$, we have

$$\delta_1(\nu) = \sum_{p=1}^{N-1} z_p(\nu) \mathbf{g}_p(\nu)$$

with

$$\mathbf{g}_p(\nu) = \frac{1}{\sqrt{N}} \sum_{k=p}^{N-1} \mathbf{h}_k e^{-i2\pi k\nu}.$$

Therefore, $\delta_1(\nu) \sim \mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{G}(\nu))$ with

$$\mathbf{G}(\nu) = 2 \sum_{p=1}^{N-1} \mathbf{g}_p(\nu) \mathbf{g}_p(\nu)^*.$$

Moreover,

$$\mathbb{E} \|\delta_1(\nu)\|_2^2 = \text{Tr} \mathbf{G}(\nu) \leq \frac{2}{N} \sum_{p=1}^{N-1} \left(\sum_{k=p}^{N-1} \|\mathbf{h}_k\|_2^2 + 2 \sum_{p \leq k < k' \leq N-1} \|\mathbf{h}_k\|_2 \|\mathbf{h}_{k'}\|_2 \right)$$

and a straightforward rearrangement together with (30) leads to

$$\begin{aligned} \max_{\nu \in [0,1]} \mathbb{E} \|\delta_1(\nu)\|_2^2 &\leq \frac{2}{N} \sum_{k=1}^{N-1} k \|\mathbf{h}_k\|_2^2 + \frac{4}{N} \sum_{1 \leq k < k' \leq N-1} \sqrt{k} \sqrt{k'} \|\mathbf{h}_k\|_2 \|\mathbf{h}_{k'}\|_2 \\ &= \frac{2}{N} \left(\sum_{k=1}^{N-1} \sqrt{k} \|\mathbf{h}_k\|_2 \right)^2 \\ &= \mathcal{O}\left(\frac{1}{N}\right). \end{aligned}$$

where we used that $k \leq \sqrt{k} \sqrt{k'}$ for $k' \geq k$. Additionally,

$$\max_{\nu \in [0,1]} \|\mathbf{G}(\nu)\|_2 \leq \max_{\nu \in [0,1]} \text{Tr} \mathbf{G}(\nu) = \mathcal{O}\left(\frac{1}{N}\right)$$

and

$$\max_{\nu \in [0,1]} \|\mathbf{G}(\nu)\|_F \leq \sqrt{M} \max_{\nu \in [0,1]} \|\mathbf{G}(\nu)\|_2 = \mathcal{O}\left(\frac{\sqrt{M}}{N}\right).$$

Using Lemma 1, there exists a constant $\kappa > 0$ independent of $M, (\mathbf{h}_k)_{k \in \mathbb{Z}}$ such that for all $t > 0$,

$$\mathbb{P} \left(\max_{\nu \in \mathcal{V}_N} \left| \|\boldsymbol{\delta}_1(\nu)\|_2^2 - \mathbb{E} \|\boldsymbol{\delta}_1(\nu)\|_2^2 \right| > t \right) \leq 2N \max_{\nu \in \mathcal{V}_N} \exp \left(-\kappa \min \left(\frac{t^2}{\|\mathbf{G}(\nu)\|_F^2}, \frac{t}{\|\mathbf{G}(\nu)\|_2} \right) \right).$$

Applying Assumption 4 and Borel-Cantelli lemma, it follows that

$$\max_{\nu \in \mathcal{V}_N} \left| \|\boldsymbol{\delta}_1(\nu)\|_2^2 - \mathbb{E} \|\boldsymbol{\delta}_1(\nu)\|_2^2 \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

Finally, we deduce that

$$\max_{\nu \in \mathcal{V}_N} \|\boldsymbol{\delta}_1(\nu)\|_2^2 \leq \max_{\nu \in \mathcal{V}_N} \mathbb{E} \|\boldsymbol{\delta}_1(\nu)\|_2^2 + \max_{\nu \in \mathcal{V}_N} \left| \|\boldsymbol{\delta}_1(\nu)\|_2^2 - \mathbb{E} \|\boldsymbol{\delta}_1(\nu)\|_2^2 \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

E. Control of $\boldsymbol{\delta}_2(\nu)$

We first split $\boldsymbol{\delta}_2(\nu)$ in the following two parts

$$\boldsymbol{\delta}_2(\nu) = \boldsymbol{\delta}_{2,1}(\nu) + \boldsymbol{\delta}_{2,2}(\nu)$$

where

$$\begin{aligned} \boldsymbol{\delta}_{2,1}(\nu) &= \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \mathbf{h}_k \sum_{p=0}^{N-1} \epsilon_{[p-k]} e^{-i2\pi p\nu} \\ \boldsymbol{\delta}_{2,2}(\nu) &= \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \mathbf{h}_k \sum_{p=0}^{N-1} \epsilon_{p-k} e^{-i2\pi p\nu}. \end{aligned}$$

We remark that $\boldsymbol{\delta}_{2,1}(\nu)$ only involves the N i.i.d. random variables $\epsilon_0, \dots, \epsilon_{N-1}$ and that

$$\boldsymbol{\delta}_{2,1}(\nu) = \sum_{p=0}^{N-1} \epsilon_p \tilde{\mathbf{g}}_p(\nu)$$

with $\tilde{\mathbf{g}}_p(\nu)$ defined as

$$\tilde{\mathbf{g}}_p(\nu) = \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \mathbf{h}_k e^{-i2\pi\nu[p+k]}.$$

It is clear that

$$\max_{p=1, \dots, N} \max_{\nu \in [0,1]} \|\tilde{\mathbf{g}}_p(\nu)\|_2 \leq \frac{1}{\sqrt{N}} \sum_{k=N}^{+\infty} \|\mathbf{h}_k\|_2 \leq \frac{1}{N^{3/2}} \sum_{k=N}^{+\infty} k \|\mathbf{h}_k\|_2$$

and from (30),

$$\max_{p=1, \dots, N} \max_{\nu \in [0,1]} \|\tilde{\mathbf{g}}_p(\nu)\|_2 = o \left(\frac{1}{N^{3/2}} \right)$$

Thus $\boldsymbol{\delta}_{2,1}(\nu) \sim \mathcal{N}_{\mathbb{C}^M} \left(\mathbf{0}, \tilde{\mathbf{G}}(\nu) \right)$ with $\tilde{\mathbf{G}}(\nu) = \sum_{p=0}^{N-1} \tilde{\mathbf{g}}_p(\nu) \tilde{\mathbf{g}}_p(\nu)^*$ and

$$\max_{\nu \in [0,1]} \text{Tr} \tilde{\mathbf{G}}(\nu) = o \left(\frac{1}{N^2} \right)$$

as $M \rightarrow \infty$. Using Lemma 1 as for the control of $\boldsymbol{\delta}_1(\nu)$ in the previous section, we end up with

$$\max_{\nu \in \mathcal{V}_N} \|\boldsymbol{\delta}_{2,1}(\nu)\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

We now consider the term $\boldsymbol{\delta}_{2,2}(\nu)$, which involves the sequence of random variables $(\epsilon_{-n})_{n \geq 1}$. For all $k \geq N$, set

$$\boldsymbol{\chi}_k = \frac{1}{\sqrt{N}} \mathbf{h}_k \sum_{p=0}^{N-1} \epsilon_{p-k} e^{-i2\pi p\nu}$$

and consider the sequence $(\chi_p)_{k \geq N}$. Using Assumption 3,

$$\begin{aligned} \sum_{k=N}^{+\infty} \|\chi_k\|_2 &\leq \sum_{k=N}^{+\infty} \sqrt{k} \|\mathbf{h}_k\|_2 \left| \frac{1}{N} \sum_{p=0}^{N-1} \epsilon_{p-k} e^{-i2\pi p\nu} \right| \\ &\leq \left(\sup_{k \geq N} \left| \frac{1}{N} \sum_{p=0}^{N-1} \epsilon_{p-k} e^{-i2\pi p\nu} \right| \right) \sum_{k=N}^{+\infty} \sqrt{k} \|\mathbf{h}_k\|_2 \\ &< +\infty \text{ a.s.} \end{aligned}$$

since for any k , by the gaussianity of the ϵ_{p-k} , $\sup_{\nu \in \mathcal{V}_N} \left| \frac{1}{N} \sum_{p=0}^{N-1} \epsilon_{p-k} e^{-i2\pi p\nu} \right|$ converges almost surely towards 0 as $N \rightarrow +\infty$ by the law of the large numbers, so it remains almost surely bounded for any finite N . This implies that the family $(\chi_k)_{k \geq N}$ is a.s. absolutely summable. Therefore, we can rearrange the series defining $\delta_{2,2}(\nu)$ and write

$$\delta_{2,2}(\nu) = \sum_{p=1}^{+\infty} \epsilon_{-p} \check{\mathbf{g}}_p(\nu)$$

with probability one, where this time $\mathbf{g}_p(\nu)$ is defined for all $p \geq 1$ as

$$\check{\mathbf{g}}_p(\nu) = \begin{cases} \frac{1}{\sqrt{N}} \sum_{k=0}^{p-1} \mathbf{h}_{k+N} e^{-i2\pi(N+k-p)\nu} & \text{if } p \in \{1, \dots, N\} \\ \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \mathbf{h}_{p+k} e^{-i2\pi k\nu} & \text{if } p \geq N+1. \end{cases}$$

Again,

$$\sup_{p \geq 1} \max_{\nu \in [0,1]} \|\check{\mathbf{g}}_p(\nu)\|_2 = o\left(\frac{1}{N}\right) \quad (33)$$

$\delta_{2,2}(\nu) \sim \mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \check{\mathbf{G}}(\nu))$, where

$$\check{\mathbf{G}}(\nu) = \sum_{p=1}^{+\infty} \check{\mathbf{g}}_p(\nu) \check{\mathbf{g}}_p(\nu)^*$$

and such that $\text{Tr } \check{\mathbf{G}}(\nu) = o\left(\frac{1}{N}\right)$. Thus, using Lemma 1 also yields

$$\max_{\nu \in \mathcal{V}_N} \|\delta_{2,2}(\nu)\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

This concludes the proof of Theorem 2.

APPENDIX C PROOF OF THEOREM 3

To prove Theorem 3, we need as a preliminary step to study the behaviour of the renormalization by $\text{dg}(\hat{\mathbf{S}}_{\mathbf{y}}(\nu))^{-\frac{1}{2}}$ in the SCM.

Lemma 3. *Under Assumptions 1, 3 and 4, we have*

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right) - \mathbf{S}_{\mathbf{v}}(\nu) \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (34)$$

as well as

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right)^{-\frac{1}{2}} - \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (35)$$

Proof: To prove (34), we establish successively

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right) - \text{dg} \left(\mathbf{S}_{\mathbf{y}}(\nu) \right) \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (36)$$

as well as

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\mathbf{S}_{\mathbf{y}}(\nu) \right) - \mathbf{S}_{\mathbf{v}}(\nu) \right\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (37)$$

Using (20), we have the bound

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right) - \text{dg} \left(\mathbf{S}_{\mathbf{y}}(\nu) \right) \right\|_2 \leq \Delta_1 + \Delta_2,$$

with

$$\Delta_1 = \max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) - \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \right\|_2$$

$$\xrightarrow[M \rightarrow \infty]{a.s.} 0,$$

and

$$\Delta_2 = \max_{\nu \in \mathcal{V}_N} \max_{m=1, \dots, M} \left| \left[\mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} - \mathbf{I}_M \right) \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \right]_{m,m} \right|.$$

Denoting $\mathbf{u}_m(\nu) = \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \mathbf{e}_m$, where \mathbf{e}_m is the m -th vector of the canonical basis of \mathbb{C}^M , as well as $\mathbf{x}_1(\nu), \dots, \mathbf{x}_{B+1}(\nu)$ the i.i.d. $\mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{I}_M)$ column vectors of $\mathbf{X}(\nu)$, we have for all $t > 0$,

$$\mathbb{P}(\Delta_2 > t) \leq \sum_{\nu \in \mathcal{V}_N} \sum_{m=1}^M \mathbb{P} \left(\left| \frac{1}{B+1} \sum_{b=1}^{B+1} |\mathbf{u}_m(\nu)^* \mathbf{x}_b(\nu)|^2 - \|\mathbf{u}_m(\nu)\|_2^2 \right| > t \right).$$

From Assumption 1, Assumption 2 and condition (11) from Assumption 3, we have

$$0 < \inf_{M \geq 1} \min_{m=1, \dots, M} \min_{\nu \in [0,1]} \|\mathbf{u}_m(\nu)\|_2 \leq \sup_{M \geq 1} \max_{m=1, \dots, M} \max_{\nu \in [0,1]} \|\mathbf{u}_m(\nu)\|_2 < \infty.$$

Setting in the statement of Lemma 1

$$\mathbf{z} = (\mathbf{x}_1(\nu)^T, \dots, \mathbf{x}_{B+1}(\nu)^T)^T \sim \mathcal{N}_{\mathbb{C}^{M(B+1)}}(\mathbf{0}, \mathbf{I}_{M(B+1)})$$

and Ξ as the $M(B+1) \times M(B+1)$ block-diagonal matrix

$$\Xi = \frac{\mathbf{I}_{B+1} \otimes (\mathbf{u}_m(\nu)\mathbf{u}_m(\nu)^*)}{B+1}$$

with \otimes denoting the Kronecker product, we obtain

$$\mathbb{P}(\Delta_2 > t) \leq 2MN \max_{\nu \in \mathcal{V}_N} \exp \left(-C \min \left\{ \frac{Bt^2}{\|\mathbf{u}_m(\nu)\|_2^4}, \frac{Bt}{\|\mathbf{u}_m(\nu)\|_2^2} \right\} \right)$$

where $C > 0$ is a constant independent of M , which in turn implies that

$$\Delta_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0$$

and that (36) holds. In order to check (37), we use Assumption 3 eq. (12) to get that

$$\max_{\nu \in \mathcal{V}_N} \|\text{dg}(\mathbf{H}(\nu)\mathbf{H}(\nu)^*)\|_2 = \max_{\nu \in \mathcal{V}_N} \max_{m=1, \dots, M} \|\mathbf{h}_m(\nu)\|_2^2 \xrightarrow[M \rightarrow \infty]{} 0$$

and from the fact that

$$\text{dg}(\mathbf{S}_{\mathbf{y}}(\nu)) = \text{dg}(\mathbf{H}(\nu)\mathbf{H}(\nu)^*) + \mathbf{S}_{\mathbf{v}}(\nu)$$

we obtain (37) and, in turn, (34).

To prove (35), we write (using that $|\sqrt{a} - \sqrt{b}| < \sqrt{|a-b|}$ for $a, b > 0$)

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right)^{-\frac{1}{2}} - \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \right\|_2 \leq \max_{\nu \in \mathcal{V}_N} \max_{m=1, \dots, M} \sqrt{\frac{|\hat{[\mathbf{S}}_{\mathbf{y}}(\nu)]_{m,m} - s_m(\nu)|}{[\hat{\mathbf{S}}_{\mathbf{y}}(\nu)]_{m,m} s_m(\nu)}}.$$

From Assumption 2, there exists $\epsilon > 0$ such that

$$\inf_{M \geq 1} \min_{m=1, \dots, M} \min_{\nu \in \mathcal{V}_N} s_m(\nu) \geq \epsilon > 0.$$

Using (34) and denoting

$$\Delta = \max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right) - \mathbf{S}_{\mathbf{v}}(\nu) \right\|_2$$

we have that

$$\max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right)^{-\frac{1}{2}} - \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \right\|_2 \leq \sqrt{\frac{\Delta}{\epsilon(\epsilon - \Delta)}}$$

with probability one for all large M , which proves (35). ■

We also need the following lemma on the boundedness of matrix $\hat{\mathbf{S}}_{\mathbf{y}}(\nu)$.

Lemma 4. *Under Assumptions 1, 3 and 4, we have*

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_M} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right\|_2 < \infty$$

with probability one.

Proof: From (20), we have

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right\|_2 \leq \limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \|\mathbf{S}_{\mathbf{y}}(\nu)\|_2 \max_{\nu \in \mathcal{V}_N} \left\| \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \right\|_2.$$

From Assumptions 1 and 3, it is clear that

$$\sup_{M \geq 1} \max_{\nu \in \mathcal{V}_N} \|\mathbf{S}_{\mathbf{y}}(\nu)\|_2 < \infty.$$

Finally, from Lemma 2 and the remarks below this lemma, we have

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \left\| \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \right\|_2 < \infty$$

with probability one, and Lemma 4 is proved. ■

Equipped with Lemmas 3 and 4, we are now in position to prove Theorem 3. Define

$$\tilde{\Delta} = \max_{\nu \in \mathcal{V}_N} \left\| \text{dg} \left(\hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right)^{-\frac{1}{2}} - \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \right\|_2$$

and recall the definition of the random matrix $\mathbf{X}(\nu)$ in (20). Let us write

$$\hat{\mathbf{C}}_{\mathbf{y}}(\nu) - \Xi(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \Xi(\nu)^{\frac{1}{2}} = \Psi_1(\nu) + \Psi_2(\nu)$$

where the two error terms are defined by:

$$\Psi_1(\nu) = \hat{\mathbf{C}}_{\mathbf{y}}(\nu) - \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \hat{\mathbf{S}}_{\mathbf{y}}(\nu) \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}}$$

$$\Psi_2(\nu) = \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} \hat{\mathbf{S}}_{\mathbf{y}}(\nu) \mathbf{S}_{\mathbf{v}}(\nu)^{-\frac{1}{2}} - \Xi(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \Xi(\nu)^{\frac{1}{2}}$$

which satisfies:

$$\max_{\nu \in \mathcal{V}_N} \|\Psi_1(\nu)\|_2 \leq \tilde{\Delta} \max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) \right\|_2 \left(\tilde{\Delta} + \frac{2}{\sqrt{\min_{\nu \in \mathcal{V}_N} \lambda_M(\mathbf{S}_{\mathbf{v}}(\nu))}} \right)$$

and

$$\max_{\nu \in \mathcal{V}_N} \|\Psi_2(\nu)\|_2 \leq \frac{\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) - \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \mathbf{S}_{\mathbf{y}}(\nu)^{\frac{1}{2}} \right\|_2}{\min_{\nu \in \mathcal{V}_N} \lambda_M(\mathbf{S}_{\mathbf{v}}(\nu))}.$$

From Assumption 2, we have

$$\inf_{M \geq 1} \min_{\nu \in \mathcal{V}_N} \lambda_M(\mathbf{S}_{\mathbf{v}}(\nu)) > 0.$$

Using Lemmas 3 and 4, we directly deduce that

$$\max_{\nu \in \mathcal{V}_N} \|\Psi_1(\nu)\|_2 \xrightarrow{M \rightarrow \infty} 0.$$

Likewise, using (20), we deduce that

$$\max_{\nu \in \mathcal{V}_N} \|\Psi_2(\nu)\|_2 \xrightarrow{M \rightarrow \infty} 0,$$

which concludes the proof of Theorem 3.

APPENDIX D
PROOF OF COROLLARY 1, COROLLARY 2 AND PROPOSITION 1

A. *Proof of Corollary 1*

Write $\hat{\mathbf{S}}_{\mathbf{v}}(\nu) = \mathbf{\Sigma}_{\mathbf{v}}(\nu)\mathbf{\Sigma}_{\mathbf{v}}(\nu)^*$, and denote

$$\Delta_{\mathbf{v}}(\nu) = \left\| \mathbf{\Sigma}_{\mathbf{v}}(\nu) - \frac{1}{\sqrt{B+1}} \mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \mathbf{Z}(\nu) \right\|_2.$$

Using the fact that for any two matrices \mathbf{A}, \mathbf{B} of appropriate dimensions, we have

$$\mathbf{A}\mathbf{A}^* - \mathbf{B}\mathbf{B}^* = (\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^* + (\mathbf{A} - \mathbf{B})\mathbf{B}^* + \mathbf{B}(\mathbf{A} - \mathbf{B})^*$$

and

$$\|\mathbf{A}\mathbf{B}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{B}\|_2$$

we see that

$$\left\| \hat{\mathbf{S}}_{\mathbf{v}}(\nu) - \frac{1}{B+1} \mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \mathbf{Z}(\nu) \mathbf{Z}(\nu)^* \mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \right\|_2 \leq \Delta_{\mathbf{v}}(\nu) \left(\Delta_{\mathbf{v}}(\nu) + 2\sqrt{\frac{\|\mathbf{S}_{\mathbf{v}}(\nu)\|_2}{B+1}} \|\mathbf{Z}(\nu)\|_2 \right).$$

Assumption 1 implies that

$$\sup_{M \geq 1} \max_{\nu \in [0,1]} \|\mathbf{S}_{\mathbf{v}}(\nu)\|_2 < \infty$$

while from Lemma 2 from Appendix A, since $\mathbf{Z}(\nu)$ has i.i.d. complex Gaussian entries,

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \frac{\|\mathbf{Z}(\nu)\|_2}{\sqrt{B+1}} < \infty \quad (38)$$

with probability one. This concludes the proof of Corollary 1.

B. *Proof of Corollary 2*

The proof of Corollary 2 is similar to the one of Corollary 1. Denoting $\Delta_{\mathbf{u}}(\nu) = \|\mathbf{\Sigma}_{\mathbf{u}}(\nu) - \mathbf{H}(\nu)\mathbf{\Sigma}_{\epsilon}(\nu)\|_2$, and noticing that $\sup_{M \geq 1} \max_{\nu \in [0,1]} \|\mathbf{H}(\nu)\|_2 < \infty$ from Assumption 3 eq. (11), we obtain that

$$\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{u}}(\nu) - \mathbf{H}(\nu)\mathbf{\Sigma}_{\epsilon}(\nu)\mathbf{\Sigma}_{\epsilon}(\nu)^*\mathbf{H}(\nu)^* \right\|_2 \leq \max_{\nu \in \mathcal{V}_N} \Delta_{\mathbf{u}}(\nu) (\Delta_{\mathbf{u}}(\nu) + 2\|\mathbf{H}(\nu)\|_2 \|\mathbf{\Sigma}_{\epsilon}(\nu)\|_2), \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Since K is fixed with respect to M from Assumption 4, we also have

$$\max_{\nu \in \mathcal{V}_N} \|\mathbf{\Sigma}_{\epsilon}(\nu)\mathbf{\Sigma}_{\epsilon}(\nu)^* - \mathbf{I}_M\|_2 \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (39)$$

using Lemma 1, which proves Corollary 2.

C. *Proof of Proposition 1*

To prove Proposition 1, let us write

$$\mathbf{Y}(\nu) = \mathbf{H}(\nu)\mathbf{\Sigma}_{\epsilon}(\nu) + \frac{\mathbf{S}_{\mathbf{v}}(\nu)^{1/2} \mathbf{Z}(\nu)}{\sqrt{B+1}}.$$

Then, from (38) and (39), we have

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \|\mathbf{Y}(\nu)\|_2 < \infty$$

with probability one, which implies the following convergence

$$\max_{\nu \in \mathcal{V}_N} \left\| \hat{\mathbf{S}}_{\mathbf{y}}(\nu) - \mathbf{Y}(\nu)\mathbf{Y}(\nu)^* \right\|_2 \leq \max_{\nu \in \mathcal{V}_N} (\Delta_{\mathbf{u}}(\nu) + \Delta_{\mathbf{v}}(\nu)) (\Delta_{\mathbf{u}}(\nu) + \Delta_{\mathbf{v}}(\nu) + 2\|\mathbf{Y}(\nu)\|_2) \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Finally, since the columns of $\sqrt{B+1}\mathbf{Y}(\nu)$ are i.i.d. $\mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{S}_{\mathbf{y}}(\nu))$ with $\mathbf{S}_{\mathbf{y}}(\nu) = \mathbf{H}(\nu)\mathbf{H}(\nu)^* + \mathbf{S}_{\mathbf{v}}(\nu)$, it follows that

$$\mathbf{Y}(\nu) = \mathbf{S}_{\mathbf{y}}(\nu)^{1/2} \frac{\mathbf{X}(\nu)}{\sqrt{B+1}}$$

for some $M \times (B+1)$ matrix $\mathbf{X}(\nu)$ having i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$ entries and the proof of Proposition 1 is complete.

APPENDIX E
PROOF OF COROLLARY 3

We first prove that all the eigenvalues of the SCM asymptotically concentrate in a compact set with probability one for all large M . Indeed, considering matrix

$$\mathbf{W}(\nu) = \Xi(\nu)^{\frac{1}{2}} \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \Xi(\nu)^{\frac{1}{2}}$$

defined through Theorem 3 and using Lemma 2 in conjunction with Borel-Cantelli lemma, we deduce that there exists constants C_1, C_2 such that

$$\liminf_{M \rightarrow \infty} \min_{\nu \in \mathcal{V}_N} \lambda_M(\mathbf{W}(\nu)) \geq C_1(1 - \sqrt{c})^2$$

and

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1(\mathbf{W}(\nu)) \leq C_2(1 + \sqrt{c})^2 \quad (40)$$

with probability one, where C_1, C_2 verify, thanks to Assumption 3,

$$0 < C_1 < 1 = \inf_{M \geq 1} \min_{\nu \in \mathcal{V}_N} \lambda_M(\Xi(\nu))$$

and

$$\sup_{M \geq 1} \max_{\nu \in \mathcal{V}_N} \lambda_M(\Xi(\nu)) < C_2 < \infty.$$

Using (22), we obtain similarly

$$\liminf_{M \rightarrow \infty} \min_{\nu \in \mathcal{V}_N} \lambda_M(\hat{\mathbf{C}}_{\mathbf{y}}(\nu)) \geq C_1(1 - \sqrt{c})^2$$

and

$$\limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1(\hat{\mathbf{C}}_{\mathbf{y}}(\nu)) \leq C_2(1 + \sqrt{c})^2$$

with probability one. Let $0 < \epsilon < \frac{C_1}{2}(1 - \sqrt{c})^2$ and $h \in \mathcal{C}_c^1(\mathbb{R})$ such that

$$h(\lambda) = \begin{cases} 1 & \text{if } \lambda \in [C_1(1 - \sqrt{c})^2 - \epsilon, C_2(1 + \sqrt{c})^2 + \epsilon] \\ 0 & \text{if } \lambda \notin [C_1(1 - \sqrt{c})^2 - 2\epsilon, C_2(1 + \sqrt{c})^2 + 2\epsilon] \end{cases}.$$

Then it follows that

$$\max_{\nu \in \mathcal{V}_N} |L_\varphi(\nu) - L_{\varphi h}(\nu)| \xrightarrow{M \rightarrow \infty} 0 \text{ a.s.}$$

Thus, without loss of generality, we may assume for the remainder of the proof that $\varphi \in \mathcal{C}_c^1((0, +\infty))$. Using (22), we deduce that

$$\max_{\nu \in \mathcal{V}_N} \frac{1}{M} \sum_{m=1}^M \left| \varphi(\lambda_m(\hat{\mathbf{C}}_{\mathbf{y}}(\nu))) - \varphi(\lambda_m(\mathbf{W}(\nu))) \right| \xrightarrow{M \rightarrow \infty} 0 \text{ a.s.}$$

Next, consider the two functions

$$\hat{m}(z, \nu) = \frac{1}{M} \sum_{m=1}^M \frac{1}{\lambda_m(\mathbf{W}(\nu)) - z} = \int_{\mathbb{R}} \frac{d\hat{\mu}(\lambda, \nu)}{\lambda - z}$$

and

$$\tilde{m}(z, \nu) = \frac{1}{M} \sum_{m=1}^M \frac{1}{\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right) - z} = \int_{\mathbb{R}} \frac{d\tilde{\mu}(\lambda, \nu)}{\lambda - z}$$

defined for all $z \in \mathbb{C}^+ := \{\zeta \in \mathbb{C} : \text{Im}(\zeta) > 0\}$, and where for all Borel set $A \subset \mathbb{R}$,

$$\hat{\mu}(A, \nu) = \frac{1}{M} \sum_{m=1}^M \delta_{\lambda_m(\mathbf{W}(\nu))}(A)$$

and

$$\tilde{\mu}(A, \nu) = \frac{1}{M} \sum_{m=1}^M \delta_{\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right)}(A)$$

denote the empirical eigenvalue distributions of matrices $\mathbf{W}(\nu)$ and $\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}$ respectively, and δ_x is the Dirac measure at point x . Functions $z \mapsto \hat{m}(z, \nu)$ and $z \mapsto \tilde{m}(z, \nu)$ coincide with the Stieltjes transforms of measures $\hat{\mu}(\cdot, \nu)$ and $\tilde{\mu}(\cdot, \nu)$ respectively (see [32] for a review of the main properties of the Stieltjes transform). Since

$$\hat{m}(z, \nu) - \tilde{m}(z, \nu) = \frac{1}{M} \text{Tr} \left((\mathbf{W}(\nu) - z\mathbf{I})^{-1} - \left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} - z\mathbf{I} \right)^{-1} \right)$$

and using the fact that $\mathbf{A}^{-1} - \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B} - \mathbf{A})\mathbf{B}^{-1}$ for non-singular matrices \mathbf{A}, \mathbf{B} , we have

$$|\hat{m}(z, \nu) - \tilde{m}(z, \nu)| \leq \frac{1}{|\text{Im}(z)|^2} \frac{K}{M} \|\mathbf{H}(\nu)\|_2^2 \|\mathbf{S}_\nu(\nu)^{-1}\|_2 \left\| \frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1} \right\|_2$$

it follows from Assumptions 2, 3, 4 and Lemma 2 that

$$\max_{\nu \in \mathcal{V}_N} |\hat{m}(z, \nu) - \tilde{m}(z, \nu)| \xrightarrow[M \rightarrow \infty]{a.s.} 0 \quad (41)$$

for all $z \in \mathbb{C}^+$. In the following, we fix a realization in an event of probability one for which (41) holds for all $z \in \mathbb{C}^+$ and consider

$$\nu^* \in \operatorname{argmax}_{\nu \in \mathcal{V}_N} \left| \int_{\mathbb{R}} \varphi(\lambda) d\hat{\mu}(\lambda, \nu) - \int_{\mathbb{R}} \varphi(\lambda) d\tilde{\mu}(\lambda, \nu) \right|.$$

Then $|\hat{m}(z, \nu^*) - \tilde{m}(z, \nu^*)| \rightarrow 0$ as $M \rightarrow \infty$, for all $z \in \mathbb{C}^+$. From the fact that the pointwise convergence on \mathbb{C}^+ of a sequence of Stieltjes transforms is equivalent to the weak convergence of the related sequence of probability measures (see e.g. [32, Ex.2.4.10]), we deduce that

$$\max_{\nu \in \mathcal{V}_N} \left| \frac{1}{M} \sum_{m=1}^M \left(\varphi(\lambda_m(\mathbf{W}(\nu))) - \varphi\left(\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right)\right) \right) \right| = \left| \int_{\mathbb{R}} \varphi(\lambda) d\hat{\mu}(\lambda, \nu^*) - \int_{\mathbb{R}} \varphi(\lambda) d\tilde{\mu}(\lambda, \nu^*) \right| \xrightarrow[M \rightarrow \infty]{} 0.$$

To conclude the proof of Corollary 3, it remains to prove that

$$\max_{\nu \in \mathcal{V}_N} \left| \frac{1}{M} \sum_{m=1}^M \varphi\left(\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right)\right) - \int_{\mathbb{R}} \varphi(\lambda) f(\lambda) d\lambda \right| \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Consider the decomposition

$$\max_{\nu \in \mathcal{V}_N} \left| \frac{1}{M} \sum_{m=1}^M \varphi\left(\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right)\right) - \int_{\mathbb{R}} \varphi(\lambda) f(\lambda) d\lambda \right| \leq \Delta_1 + \Delta_2,$$

where

$$\Delta_1 = \max_{\nu \in \mathcal{V}_N} \left| \frac{1}{M} \sum_{m=1}^M \left(\varphi\left(\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right)\right) - \mathbb{E} \left[\varphi\left(\lambda_m\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right)\right) \right] \right) \right|$$

and

$$\Delta_2 = \left| \frac{1}{M} \sum_{m=1}^M \mathbb{E} \left[\varphi\left(\lambda_m\left(\frac{\mathbf{X}(0)\mathbf{X}(0)^*}{B+1}\right)\right) \right] - \int_{\mathbb{R}} \varphi(\lambda) f(\lambda) d\lambda \right|.$$

Using the concentration inequality of [33, Cor. 1.8(b)], it is straightforward to show that

$$\Delta_1 \xrightarrow[M \rightarrow \infty]{a.s.} 0.$$

Moreover, using again the properties of the Stieltjes transform, it can be deduced from e.g. [34] that

$$\Delta_2 \xrightarrow[M \rightarrow \infty]{} 0.$$

This concludes the proof of Corollary 3.

APPENDIX F
PROOF OF PROPOSITION 2

Convergences (25), (26) and (27) are straightforward consequences of (22) and the results of [20, Th. 1.1] on the behaviour of the largest eigenvalues for the so-called multiplicative spike model random matrices. To prove (28), we use the bound

$$\lambda_1(\mathbf{W}(\nu)) \leq \lambda_1\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right) \lambda_1(\mathbf{\Xi}(\nu))$$

Then, from the fact that $\gamma_\infty = 0$ and Lemma 2, we finally obtain

$$\begin{aligned} \limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1(\mathbf{W}(\nu)) &\leq \limsup_{M \rightarrow \infty} \max_{\nu \in \mathcal{V}_N} \lambda_1\left(\frac{\mathbf{X}(\nu)\mathbf{X}(\nu)^*}{B+1}\right) \\ &\leq (1 + \sqrt{c})^2. \end{aligned}$$

The proof is concluded by invoking again convergence (22).