On the asymptotic distribution of the maximum sample spectral coherence of Gaussian time series in the high dimensional regime

Philippe Loubaton, Alexis Rosuel, Pascal Vallet

On the asymptotic distribution of the maximum sample spectral coherence of Gaussian time series in the high dimensional regime*

Philippe Loubaton[a], Alexis Rosuel[b,], Pascal Vallet[c]

[a] *Laboratoire d'Informatique Gaspard Monge (CNRS, Univ. Gustave-Eiffel), 5 Bd. Descartes 77454 Marne-la-Vallée (France)*

[b] *Laboratoire d'Informatique Gaspard Monge (CNRS, Univ. Gustave-Eiffel), 5 Bd. Descartes 77454 Marne-la-Vallée (France)*

[c] *Laboratoire de l'Intégration du Matériau au Système (CNRS, Univ. Bordeaux, Bordeaux INP), 351, Cours de la Libération 33405 Talence (France)*

Abstract

We investigate the asymptotic distribution of the maximum of a frequency smoothed estimate of the spectral coherence of a M-variate complex Gaussian time series with mutually independent components when the dimension M and the number of samples N both converge to infinity. If B denotes the smoothing span of the underlying smoothed periodogram estimator, a type I extreme value limiting distribution is obtained under the rate assumptions $\frac{M}{N} \to 0$ and $\frac{M}{B} \to c \in (0, +\infty)$. This result is then exploited to build a statistic with controlled asymptotic level for testing independence between the M components of the observed time series. Numerical simulations support our results.

Keywords: Spectral Analysis, High Dimensional Statistics, Time Series, Independence Test.
2010 MSC: 62H15, 62H20, 62M15

*This work is funded by ANR Project HIDITSA, reference ANR-17-CE40-0003.

Email addresses: philippe.loubaton@univ-eiffel.fr (Philippe Loubaton), alexis.rosuel@univ-eiffel.fr (Alexis Rosuel), pascal.vallet@bordeaux-inp.fr (Pascal Vallet)

1. Introduction

1.1. The addressed problem and the results

We consider a zero mean M-variate complex Gaussian stationary time series $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ [1] and denote $\mathbf{S}(\nu)$ and $\mathbf{C}(\nu)$ its spectral density and spectral coherency matrices defined for each $\nu \in [0,1]$ by

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

and

$$\mathbf{C}(\nu) = \mathrm{dg}(\mathbf{S}(\nu))^{-1/2}\mathbf{S}(\nu)\,\mathrm{diag}(\mathbf{S}(\nu))^{-1/2}$$

where $\mathbf{R}(u) = \mathbb{E}(\mathbf{y}_{n+u}\mathbf{y}_n^*)$ and $\mathrm{dg}(\mathbf{S}(\nu)) = \mathbf{S}(\nu)\odot \mathbf{I}_M$, with $\odot$ denoting the Hadamard product (ie. entrywise product) and $\mathbf{I}_M$ is the M–dimensional identity matrix. Assuming N observations $(\mathbf{y}_n)_{n=1,\ldots,N}$ are available, we consider the frequency smoothed estimate $\hat{\mathbf{S}}(\nu)$ defined by

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

where B is an even integer, which represents the smoothing span, and

$$\boldsymbol{\xi}_{\mathbf{y}}(\nu) = \frac{1}{\sqrt{N}}\sum_{n=1}^{N}\mathbf{y}_n e^{-2i\pi(n-1)\nu} \tag{2}$$

is the renormalized Fourier transform of $(\mathbf{y}_n)_{n=1\ldots,N}$. The corresponding estimated spectral coherency matrix is defined as:

$$\hat{\mathbf{C}}(\nu) = \mathrm{dg}(\hat{\mathbf{S}}(\nu))^{-\frac{1}{2}}\hat{\mathbf{S}}(\nu)\,\mathrm{dg}(\hat{\mathbf{S}}(\nu))^{-\frac{1}{2}} \tag{3}$$

We denote by $(y_{1,n})_{n\in\mathbb{Z}},\ldots,(y_{M,n})_{n\in\mathbb{Z}}$ the M components of $\mathbf{y}$, and $s_{ij}(\nu)$, $c_{ij}(\nu)$, $\hat{s}_{ij}(\nu)$, $\hat{c}_{ij}(\nu)$ the entries i,j of matrices $\mathbf{S}(\nu), \mathbf{C}(\nu), \hat{\mathbf{S}}(\nu), \hat{\mathbf{C}}(\nu)$ respectively. We remark that

$$c_{ij}(\nu) = \frac{s_{ij}(\nu)}{\sqrt{s_{ii}(\nu)s_{jj}(\nu)}}, \hat{c}_{ij}(\nu) = \frac{\hat{s}_{ij}(\nu)}{\sqrt{\hat{s}_{ii}(\nu)\hat{s}_{jj}(\nu)}}$$

[1] any finite linear combination x of the components of $(\mathbf{y}_n)_{n\in\mathbb{Z}}$ is a complex Gaussian random variable, i.e. $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ are independent zero-mean Gaussian random variables having the same variance

for each ν.

Under the hypothesis

$$\mathcal{H}_0 : (y_{1,n})_{n\in\mathbb{Z}}, \ldots, (y_{M,n})_{n\in\mathbb{Z}} \text{ are mutually uncorrelated,}$$

we evaluate the behaviour of the Maximum Sample Spectral Coherence (MSSC) defined by

$$\max_{1\leq i<j\leq M} \max_{\nu\in\mathcal{G}} |\hat{c}_{ij}(\nu)|$$

where

$$\mathcal{G} := \left\{ k\frac{B+1}{N} : k \in \mathbb{N}, 0 \leq k \leq \frac{N}{B+1} \right\}$$

is the subset of the Fourier frequencies

$$\mathcal{F} := \left\{ \frac{k}{N} : k \in \mathbb{N}, 0 \leq k \leq N-1 \right\}$$

with elements spaced by a distance $(B+1)/N$. Our study is conducted in the asymptotic regime where $M = M(N)$ and $B = B(N)$ are both functions of N such that for some $\rho \in (0,1)$, $M \asymp N^\rho$ and $B \asymp N^\rho$ as $N \to \infty$ [2] , while the ratio M/B converges to some constant $c \in (0, +\infty)$. It is established that, under $\mathcal{H}_0$ and proper assumptions on the time series $(y_{1,n})_{n\in\mathbb{Z}}, \ldots, (y_{M,n})_{n\in\mathbb{Z}}$, for any $t \in \mathbb{R}$:

$$\mathbb{P}\left((B+1) \max_{(i,j,\nu)\in\mathcal{I}} |\hat{c}_{ij}(\nu)|^2 \leq t + \log\frac{N}{B+1} + \log\frac{M(M-1)}{2} \right) \xrightarrow[N\to+\infty]{} e^{-e^{-t}} \tag{4}$$

where

$$\mathcal{I} := \{(i,j,\nu) : i,j \in [M] \text{ such that } i<j,\ \nu \in \mathcal{G}\} \tag{5}$$

with $[M] = \{1, \ldots, M\}$.

In other words, under proper normalization and centering, $\max_{(i,j,\nu)\in\mathcal{I}} |\hat{c}_{ij}(\nu)|^2$ follows asymptotically a Gumbel distribution (see Embrechts et al. (2013) or Resnick (2013) for a general theory of extreme value distributions).

[2] For two sequences $(x_n)_{n\geq 1}, (y_n)_{n\geq 1}$, we denote by $x_n \asymp y_n$ if there exists $k_1, k_2 > 0$ such that $k_1|y_n| \leq |x_n| \leq k_2|y_n|$ for all large n.

1.2. Motivation

This paper is motivated by the problem of testing the independence of a large number of Gaussian time series. Since hypothesis $\mathcal{H}_0$ can be equivalently formulated as

$$\mathcal{H}_0 : \max_{1 \leq i < j \leq M} \max_{\nu \in [0,1]} |s_{ij}(\nu)|^2 = 0,$$

or by

$$\mathcal{H}_0 : \max_{1 \leq i < j \leq M} \max_{\nu \in [0,1]} |c_{ij}(\nu)|^2 = 0,$$

this suggests to compute consistent estimators of these quantities, and test their closeness to zero.

Our choice of the high-dimensional regime defined above is motivated as follows. Under mild assumptions on the memory of the time series $((y_{m,n})_{n \in \mathbb{Z}})_{m \geq 1}$, in the low-dimensional regime where $N \to +\infty$ and M is fixed, it can be shown that the sample spectral coherence matrix $\hat{\mathbf{C}}(\nu)$ defined by (3) is a consistent estimate (in spectral norm for instance) of the spectral coherence matrix $\mathbf{C}(\nu)$ as long as $B \to +\infty$ and $B/N \to 0$ (up to some additional logarithmic terms). In practice, this asymptotic regime and the underlying predictions are relevant as long as the ratio M/N is small enough. If this condition is not met, test statistics based on $\hat{\mathbf{C}}(\nu)$ may be of delicate use, as the choice of the smoothing span B must meet the constraints $B \gg M$ (because B is supposed to converge towards $+\infty$) as well as $B \ll N$ (because B/N is supposed to converge towards 0). Nowadays, for many practical applications involving high dimensional signals and/or a moderate sample size, the ratio M/N may not be small enough to be able to choose B so as to meet $B \gg M$ and $B \ll N$. In this situation, one may rely on the more relevant high dimensional regime in which M, B, N converge to infinity such that M/B converges to a positive constant while B/N converges to zero.

1.3. On the literature

Correlation tests using spectral approaches have been studied in several papers, see e.g. Wahba (1971), Eichler (2008) and the references therein.

More recently, an approach similar to the one of this paper has been explored in Wu and Zaffaroni (2018), where the maximum of the sample spectral coherence, when using lag-window estimates of the spectral density, is studied. In the low-dimensional regime where M is fixed and $N \to \infty$, it is proved that the distribution of such statistic under $\mathcal{H}_0$, after proper centering and normalization, converges to the Gumbel distribution. We also mention other related papers exploring the asymptotic behaviour of various spectral density estimates in the low-dimensional regime: Woodroofe and Van Ness (1967), Rudzkis (1985), Shao et al. (2007), Lin and Liu (2009) and Liu and Wu (2010).

In the high-dimensional regime when M is a function of N such that $M := M(N) \to +\infty$ as $N \to \infty$, few results on the behaviour of correlation test statistics in the spectral domain are known. Loubaton and Rosuel (2021) proved that under $\mathcal{H}_0$ and mild assumptions on the underlying time series, the empirical eigenvalue distribution of $\hat{\mathbf{C}}(\nu)$ defined in (3) converges weakly almost surely towards the Marcenko-Pastur distribution, which can be exploited to build test statistics based on linear spectral statistics of $\hat{\mathbf{C}}(\nu)$. In Rosuel et al. (2020), a consistent test statistic based on the largest eigenvalue of $\hat{\mathbf{C}}(\nu)$ was derived for the problem of detecting the presence of a signal with low rank spectral density matrix within a noise with uncorrelated components.

In the asymptotic regime where $\frac{M}{N} \to \gamma > 0$, Pan et al. (2014) proposed to test hypothesis $\mathcal{H}_0$ when the components of $\mathbf{y}$ share the same spectral density. In this case, the rows of the $M \times N$ matrix $(\mathbf{y}_1, \ldots, \mathbf{y}_N)$ are independent and identically distributed under $\mathcal{H}_0$. Pan et al. (2014) established a central limit theorem for linear spectral statistics of the empirical covariance matrix, and deduced from this a test statistics to check whether $\mathcal{H}_0$ holds or not. We notice that the results of Pan et al. (2014) are valid in the non Gaussian case. We note also that in the specific case where the M time series have the same marginal spectral density, it is possible to directly use results on temporally white noise tests, such as those in Li et al. (2019) and Chang et al. (2017). To explain this, we denote by $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$ the $M \times N$ matrix of observations, and by $\mathbf{z}_1^T, \ldots, \mathbf{z}_M^T$ the rows of $\mathbf{Y}$, which, in the above particular case, share the same covariance matrix $\mathbf{T}$. Then, the row vectors $\mathbf{z}_1^T, \ldots, \mathbf{z}_M^T$ are uncorrelated if and only the column vectors $\mathbf{z}_1, \ldots, \mathbf{z}_M$ are extracted from a white noise sequence with covariance $\mathbf{T}$. However, in our more general model, since each time series is generated by a different spectral density, vectors $\mathbf{z}_1, \ldots, \mathbf{z}_M$ are not anymore iid, which prevents to use the results of Li et al. (2019) and Chang et al. (2017).

More results are available in the case where the time series $(y_{m,n})_{n\in\mathbb{Z}}$,

$m \in [M]$, are temporally white. To test the correlation of the M components, one can similarly consider sample estimates of the correlation matrix, and test whether it is close to the identity matrix. Under the asymptotic regime where $\frac{M}{N} \to \gamma \in (0, +\infty)$, Jiang et al. (2004) showed that the maximum off-diagonal entry of the sample correlation matrix after proper normalization is also asymptotically distributed as Gumbel. The techniques used here for proving (4) are partly based on this paper. Other works such as Mestre and Vallet (2017) studied the asymptotic distribution of linear spectral statistics of the correlation matrix, Dette and Dörnemann (2020) focused on the behaviour of the determinant of the correlation matrix, and Cai et al. (2013) considered a U-statistic and obtained minimax results over some class of alternatives. Some other papers also explored various classes of alternative $\mathcal{H}_1$, among which is Fan et al. (2019), who showed a phase transition phenomena in the behaviour of the largest off-diagonal entry of the correlation matrix driven by the magnitude of the dependence parameter defined in the alternative class $\mathcal{H}_1$. Lastly, Morales-Jimenez et al. (2018) studied asymptotic first and second order behaviour of the largest eigenvalues and associated eigenvectors of the sample correlation matrix under a specific alternative spiked model.

2. Main results

2.1. Assumptions

Throughout the paper we rely on the following assumptions.

Assumption 1 (Time series)**.** *The time series $(y_{m,n})_{n\in\mathbb{Z}}$, $m \geq 1$, are mutually independent, stationary and zero-mean complex Gaussian distributed* [3]*.*

For each $m \geq 1$, we denote by $r_m = (r_m(u))_{u\in\mathbb{Z}}$ (instead of $r_{m,m}$) the covariance sequence of $(y_{m,n})_{n\in\mathbb{Z}}$, i.e. $r_m(u) = \mathbb{E}[y_{m,n+u}\overline{y_{m,n}}]$, and we formulate the following assumption on $(r_m)_{m\geq1}$:

Assumption 2 (Memory)**.** *The covariance sequences $(r_m)_{m\geq1}$ satisfy the uniform short memory condition*

$$\sup_{m\geq1} \sum_{u\in\mathbb{Z}} (1 + |u|)|r_m(u)| < +\infty.$$

[3] A complex random variable Z is zero-mean complex Gaussian distributed with variance σ^2, denoted as $Z \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$, if $\mathrm{Re}(Z)$ and $\mathrm{Im}(Z)$ are i.i.d. $\mathcal{N}(0, \frac{\sigma^2}{2})$ random variables.

We denote by $s_m(\nu) = \sum_{u \in \mathbb{Z}} r_m(u) \mathrm{e}^{-\mathrm{i}2\pi\nu}$ (instead of $s_{m,m}(\nu)$) the spectral density of $(y_{m,n})_{n\in\mathbb{Z}}$ at frequency $\nu \in [0,1]$. Assumption 2 of course implies that the function s_m is continously differentiable and that

$$\sup_{m\geq 1} \max_{\nu\in[0,1]} s_m(\nu) < +\infty, \quad \sup_{m\geq 1} \max_{\nu\in[0,1]} \left| \frac{\mathrm{d}s_m}{\mathrm{d}\nu}(\nu) \right| < +\infty. \tag{6}$$

Eventually, as the sample spectral coherence of $(y_{i,n})_{n\in\mathbb{Z}}$ and $(y_{j,n})_{n\in\mathbb{Z}}$ involves a renormalization by the inverse of the estimates of the spectral densities s_i and s_j, we also need that s_i, s_j do not vanish. This is the substance of the next assumption.

Assumption 3 (Non-vanishing spectrum)**.** *The spectral densities are uniformly bounded away from zero, that is*

$$\inf_{m\geq 1} \min_{\nu\in[0,1]} s_m(\nu) > 0. \tag{7}$$

By Assumptions 2 and 3, there exist quantities $s_{\min}$ and $s_{\max}$ such that

$$0 < s_{\min} \leq \inf_{m\geq 1} \min_{\nu\in[0,1]} s_m(\nu) \leq \sup_{m\geq 1} \max_{\nu\in[0,1]} s_m(\nu) \leq s_{\max} < +\infty. \tag{8}$$

We now formulate the following assumptions on the growth rate of the quantities N, M, B, which describe the high-dimensional regime considered in this paper.

Assumption 4 (Asymptotic regime)**.** *B and M are functions of N such that there exist positive constants $C_1, C_2 \in (0,+\infty)$ and $\rho \in (0,1)$ such that:*

$$C_1 N^\rho \leq B, M \leq C_2 N^\rho$$

and

$$\frac{M}{B} := c_N \xrightarrow[N\to+\infty]{} c \in (0,+\infty).$$

Notations. Even if the subscript $\cdot_N$ is not always specified, almost all quantities should be remembered to be dependent on N. Moreover, C represents a universal constant (i.e. a positive quantity independent of N, M, B), whose precise value is irrelevant and which may change from one line to another.

2.2. Statement of the result

The main result of this paper, whose proof is deferred to Section 4, is given in the following theorem.

Theorem 1. *Under Assumptions 1 – 3, for any $t \in \mathbb{R}$:*

$$\mathbb{P}\left((B+1) \max_{(i,j,\nu)\in\mathcal{I}} |\hat{c}_{ij}(\nu)|^2 \leq t + \log \frac{N}{B+1} + \log \frac{M(M-1)}{2}\right) \xrightarrow[N\to+\infty]{} e^{-e^{-t}}.$$

Thus, Theorem 1 states that $\max_{(i,j,\nu)\in\mathcal{I}} |\hat{c}_{ij}(\nu)|^2$, atfer proper normalization and centering, converges in distribution to a type I extreme value distribution, also known as Gumbel distribution. As it will be clear in the proof, the term $\log \frac{M(M-1)}{2}$ is related to the maximum over (i,j) while the term $\log \frac{N}{B+1}$ is related to the maximum over $\nu \in \mathcal{G}$.

We now illustrate numerically the above asymptotic result. Consider M independent AR(1) processes, driven by a standard Gaussian white noise, i.e.

$$\mathbf{y}_n := \begin{pmatrix} y_{1,n} \\ \vdots \\ y_{M,n} \end{pmatrix} = \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_M \end{pmatrix} \begin{pmatrix} y_{1,n-1} \\ \vdots \\ y_{M,n-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,n} \\ \vdots \\ \epsilon_{M,n} \end{pmatrix}, \quad \epsilon_{m,n} \overset{i.i.d.}{\sim} \mathcal{N}_{\mathbb{C}}(0,1) \tag{9}$$

with $\theta_1, \ldots, \theta_M$ are iid uniformly distributed on the complex disk of radius 0.9, and $(N, M) = (20000, 500)$. The smoothed periodogram estimators are computed using $B = 1000$. We independently draw 10000 samples of the time series $(\mathbf{y}_n)_{n\in[N]}$ and compute the associated MSSC $\max_{(i,j,\nu)\in\mathcal{I}_N} |\hat{c}_{ij}(\nu)|^2$. On Figure 1 are represented the sample cumulative distribution function (cdf) and the histogram of the MSSC against the Gumbel cdf and probability density function (pdf). We indeed observe that the rescaled distribution of $\max_{(i,j,\nu)\in\mathcal{I}_N} |\hat{c}_{ij}(\nu)|^2$ is close to the Gumbel distribution.

3. Application to testing

3.1. New proposed test statistic

Theorem 1 can be used to design a new independence test statistic with controlled asymptotic level in the proposed high-dimensional regime.

Define q_α the α–quantile of the Gumbel distribution: $q_\alpha = F^{-1}(\alpha)$ where

$$F(x) = \exp(-\exp(-x)).$$

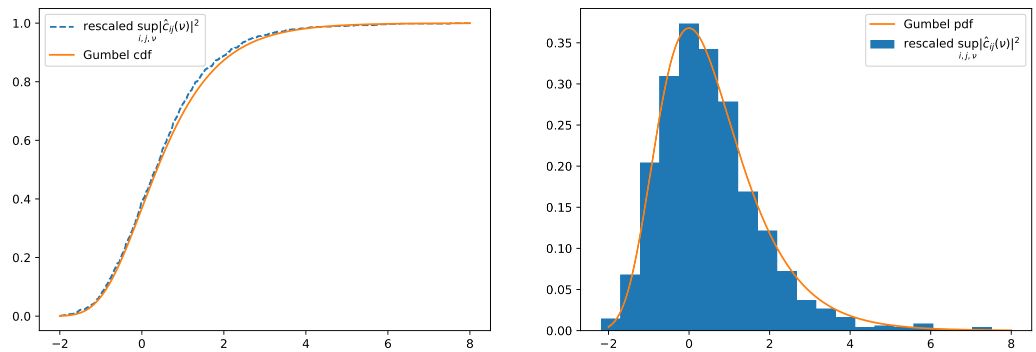

Figure 1: sample cdf and histogram of the MSSC as defined in Theorem 1 vs Gumbel distribution.

.

The test statistic $T_N^{(\mathrm{MSSC})}$ defined by

$$T_N^{(\mathrm{MSSC})} = \mathbb{1}\left(\max_{(i,j,\nu)\in\mathcal{I}} |\hat{c}_{ij}(\nu)|^2 > \frac{q_{1-\alpha} + \log\frac{N}{B+1} + \log\frac{M(M-1)}{2}}{B+1} \right) \tag{10}$$

satisfies, as a direct consequence of Theorem 1, $\lim_{N\to+\infty} \mathbb{P}[T_N^{(\mathrm{MSSC})} = 1] = \alpha$ under $\mathcal{H}_0$.

Note that in the following, the smoothing span B is set according to the dimension M in order to illustrate the asymptotic regime in which $B \asymp M$. In a practical context, the smoothing span B may be chosen via data-driven methods, such as the one in Ombao et al. (2001). Nevertheless, although it was used in the context of high-dimensional time series in Fiecas and von Sachs (2014); Fiecas et al. (2019), classical span selection procedures such as Ombao et al. (2001) usually make sense for low-dimensional time series. The development of a new data-driven span selection method, relevant for high-dimensional time series, is outside the scope of this paper but would be an interesting topic for future research.

3.2. Type I error

In order to test the independence of the signals $((y_{m,n})_{n\in\mathbb{Z}})_{m=1,\ldots,M}$, we consider the statistic $T_N^{(\mathrm{MSSC})}$ defined in (10). In Table 1 are presented the sample type I errors of $T_N^{(\mathrm{MSSC})}$ with different combinations of sample sizes and dimensions ($\rho = 0.7$ and $\frac{M}{B+1} = 0.5$), when the nominal significant level for all the tests is set at $\alpha = 0.05$, and all statistics are computed from 30000 independent replications. The time series are still generated according to model (9). One can see as expected that the type I error of $T_N^{(\mathrm{MSSC})}$ does indeed remain near 5% as M increases.

Table 1: Sample type I error at 5%

N	B	M	$T_N^{(\mathrm{MSSC})}$
42	20	10	0.012
316	100	50	0.035
659	180	90	0.037
1044	260	130	0.045
1459	340	170	0.046
1901	420	210	0.048
5623	1000	500	0.048
13374	2000	1000	0.049

3.3. Power

We now compare the power of our new test statistic against other independence test statistics which are designed to work in the high-dimensional regime. We define the Linear Spectral Statistic (LSS) test from Loubaton and Rosuel (2021) for any $\epsilon > 0$ by

$$T_N^{(\mathrm{LSS})} = \mathbb{1}\left(\sup_{\nu\in[0,1]} \frac{\left|\frac{1}{M}\mathrm{tr}\ f(\hat{\mathbf{C}}(\nu)) - \int_{\mathbb{R}} f\,\mathrm{d}\mu_{MP}^{(c_N)}\right|}{N^{\epsilon}(B/N)} > \kappa_{1-\alpha}\right) \tag{11}$$

where $\mu_{MP}^{(c)}$ represents the Marcenko-Pastur distribution with parameter c defined by

$$d\mu_{MP}^{(c)}(\lambda) = \left(1 - \frac{1}{c}\right)_+ \delta_0(\mathrm{d}\lambda) + \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi c\lambda}\mathbb{1}_{[\lambda_-,\lambda_+]}(\lambda)\,\mathrm{d}\lambda$$

where $\lambda_\pm = (1 \pm \sqrt{c})^2$, $(\cdot)_+ := \max(\cdot, 0)$, $c_N := \frac{M}{B+1}$ and f is some function defined on $\mathbb{R}_+$ satisfying regularity assumptions (see more details in Loubaton and Rosuel (2021)). In practice, ϵ will be taken equal to 0.1. It is proven in Loubaton and Rosuel (2021) that under $\mathcal{H}_0$, $T_N^{(\mathrm{LSS})} \to 0$ almost surely in the high-dimensional regime but the exact asymptotic distribution of the LSS test is unknown. Therefore, the detection threshold $\kappa_{1-\alpha}$ for this test is based on a sample quantile of $T_N^{(\mathrm{LSS})}$ under $\mathcal{H}_0$ computed from Monte-Carlo simulation. For fairness comparison, we also use this procedure for the new test statistic $T_N^{(\mathrm{MSSC})}$. More precisely, we compute the sample $(1-\alpha)$–quantile $\kappa_{1-\alpha}$ of a test statistic $T_N^{(\mathrm{LSS})}$ from samples under $\mathcal{H}_0$, and then reject the null hypothesis under $\mathcal{H}_1$ if $T_N^{(LSS)} > \kappa_{1-\alpha}$. It remains to choose a test function f, and we again follow Loubaton and Rosuel (2021) by considering

- the Frobenius test $T_N^{(\mathrm{FROB})}$ when $f(x) = (x-1)^2$
- the logdet test $T_N^{(\mathrm{LOG})}$ when $f(x) = \log x$

It remains to define the alternatives. For this, we consider the following multidimensional $AR(1)$ model:

$$\mathbf{y}_{n+1} = \mathbf{A}\mathbf{y}_n + \boldsymbol{\epsilon}_n \tag{12}$$

where $(\boldsymbol{\epsilon}_n)_{n\in\mathbb{Z}}$ is a sequence of independent $\mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{I})$ distributed random vectors, and $\mathbf{A}$ is a bidiagonal matrix. Three choices of $\mathbf{A}$ ($\mathbf{A}^{(\mathcal{H}_0)}$, $\mathbf{A}^{(\mathcal{H}_{1,\mathrm{loc}})}$, $\mathbf{A}^{(\mathcal{H}_{1,\mathrm{glob}})}$) allows us to define two alternatives:

1. $\mathcal{H}_0$: for $|\theta| < 1$:
$$\mathbf{A}^{(\mathcal{H}_0)} = \theta\,\mathbf{I}_M$$
so the signals $((y_{m,n})_{n\in\mathbb{Z}})_{m=1,\dots,M}$ are mutually independent.
2. $\mathcal{H}_{1,\mathrm{loc}}$: for $|\theta| < 1$ and $\beta \in \mathbb{R}$, it exists an unknown pair (i,j), $i > j$, such that :
$$\mathbf{A}^{(\mathcal{H}_{1,\mathrm{loc}})} = \theta\,\mathbf{I}_M + \beta\,\mathbf{e}_i\mathbf{e}_j^T$$
where $\mathbf{e}_1, \dots, \mathbf{e}_M$ represents the canonical basis of $\mathbf{C}^M$ so the couple of time series (i,j) is the unique correlated pair of signals. In the following experiments, we consider the case $i = 2$ and $j = 1$.
3. $\mathcal{H}_{1,\mathrm{glob}}$: for $|\theta| < 1$ and $\beta \in \mathbb{R}$:
$$\mathbf{A}^{(\mathcal{H}_{1,\mathrm{glob}})} = \begin{pmatrix} \theta & 0 & \dots & \dots & \dots & 0 \\ \beta & \theta & 0 & \dots & \dots & 0 \\ 0 & \beta & \theta & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \beta & \theta \end{pmatrix}$$

so all the signals are mutually correlated.

We now fix the value of the parameters involved under the three hypotheses. θ will always be taken equal to 0.5. Under $\mathcal{H}_{1,\text{loc}}$, $\beta = 0.1$. Concerning the alternative $\mathcal{H}_{1,\text{glob}}$, more care is required to choose β. Indeed, one can define a measure of total dependence as:

$$r := \frac{\int \|\mathbf{S}(\nu) - \mathrm{dg}\mathbf{S}(\nu)\|_F^2 \,\mathrm{d}\nu}{\int \|\mathbf{S}(\nu)\|_F^2 \,\mathrm{d}\nu} = \frac{\sum_{u\in\mathbb{Z}} \|\mathbf{R}(u) - \mathrm{dg}\mathbf{R}(u)\|_F^2}{\sum_{u\in\mathbb{Z}} \|\mathbf{R}(u)\|_F^2}$$

where $\mathbf{R}(u) := \mathbb{E}[\mathbf{y}_{n+u}\mathbf{y}_n^*]$, $\mathbf{S}(\nu) = \sum_{u\in\mathbb{Z}} \mathbf{R}(u)\mathrm{e}^{-\mathrm{i}2\pi u\nu}$ and dg denotes the diagonal part operator. Clearly, $r = 0$ under $\mathcal{H}_0$, and as $r > 0$ increases, the M–dimensional time series become correlated. We also see that for any fixed value of β, r is increasing with M. It is therefore more desirable to tune $\beta := \beta(M)$ such that r remains constant as M increases. This will enable our tests to be compared against an alternative which does not become asymptotically trivial.

The two alternatives $\mathcal{H}_{1,\text{loc}}$ and $\mathcal{H}_{1,\text{glob}}$ are useful to measure the performance of the independence tests under two different setups. Under $\mathcal{H}_{1,\text{loc}}$, each pair of time series are independent except the pair $(y_{1,n})_{n\in\mathbb{Z}}$,$(y_{2,n})_{n\in\mathbb{Z}}$, whereas under $\mathcal{H}_{1,\text{glob}}$ each time series has a small correlation with every other time series.

In Table 2 and Table 3 are presented the sample powers when the type I error is fixed at 5% for the considered tests and the two alternatives. The asymptotic regime is the same as the one considered for Table 1: $\rho = 0.7$ and $\frac{M}{B+1} = 0.2$. All statistics are computed from 30000 independent replications. We observe that under $\mathcal{H}_{1,\text{glob}}$, with $r = 0.01$, all the tests asymptotically detect the alternative, however with different performances. The LSS test statistics show better power which indicates that they may be more suited to detect alternative under $\mathcal{H}_{1,\text{glob}}$ than the MSSC test statistics. Under $\mathcal{H}_{1,\text{loc}}$ the results are opposite: the power of $T_N^{(\text{MSSC})}$ rapidly increases to 1 as M increases. These results are not surprising since the MSSC test statistic is designed to detect peaks in the off-diagonal entries of $\hat{\mathbf{C}}(\nu)$ which is exactly the class of alternative considered in $\mathcal{H}_{1,\text{loc}}$. However, when the correlations are spread among all pairs of time series under $\mathcal{H}_{1,\text{glob}}$, the test statistics based on the global behaviour of the eigenvalues of $\hat{\mathbf{C}}(\nu)$ seem more relevant.

On Figure 2 are represented the ROC for each test under both alternatives. We observe that $T_N^{(\text{FROB})}$ and $T_N^{(\text{LOG})}$ have similar performance and

Table 2: Power comparison under $\mathcal{H}_1$ global, type I error = 5%

N	M	B	$T_N^{(\mathrm{FROB})}$	$T_N^{(\mathrm{LOG})}$	$T_N^{(\mathrm{MSSC})}$
42	10	20	0.050	0.049	0.052
316	50	100	0.036	0.042	0.067
659	90	180	0.067	0.065	0.086
1044	130	260	0.142	0.122	0.133
1459	170	340	0.339	0.255	0.214
1901	210	420	0.601	0.462	0.328
2364	250	500	0.836	0.682	0.503
2846	290	580	0.960	0.852	0.672

Table 3: Power comparison under $\mathcal{H}_1$ local, type I error = 5%

N	M	B	$T_N^{(\mathrm{FROB})}$	$T_N^{(\mathrm{LOG})}$	$T_N^{(\mathrm{MSSC})}$
42	10	20	0.049	0.049	0.061
316	50	100	0.038	0.044	0.352
659	90	180	0.038	0.041	0.881
1044	130	260	0.034	0.038	0.999
1459	170	340	0.034	0.038	1.000
1901	210	420	0.035	0.039	1.000
2364	250	500	0.031	0.039	1.000
2846	290	580	0.032	0.036	1.000

outperform $T_N^{(\mathrm{MSSC})}$ for the alternative $\mathcal{H}_{1,\mathrm{glob}}$, while $T_N^{(\mathrm{MSSC})}$ has better performance for $\mathcal{H}_{1,\mathrm{loc}}$.

To conclude this section, we mention that the work of Fan et al. (2019), which extends the study of Jiang et al. (2004) on the asymptotic distribution of the largest entry of the sample covariance matrix under a specific $\mathcal{H}_1$ scenario, shows that a phase transition phenomenon occurs, as the asymptotic distribution may take three different forms depending on the behaviour of the distance between the $\mathcal{H}_0$ and $\mathcal{H}_1$ sample distributions. Extending such results to the MSSC test statistic for a general class of alternative hypotheses $\mathcal{H}_1$ would be a deep and interesting perspective requiring a significant work.

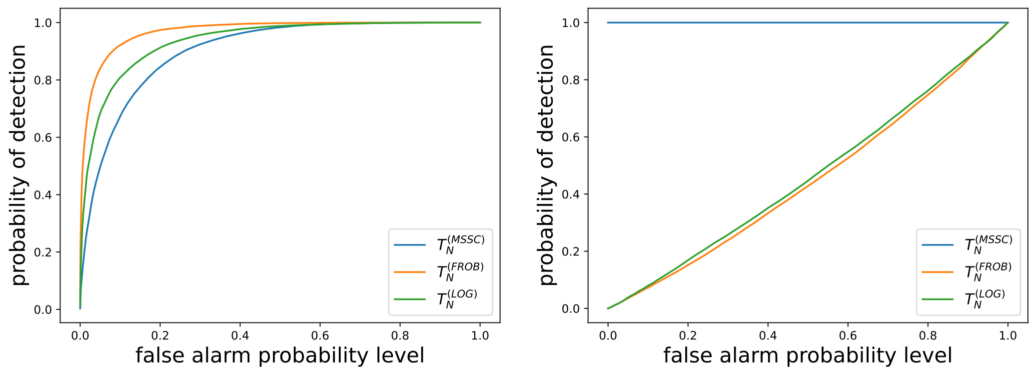

Figure 2: ROC associated to each test under $\mathcal{H}_1^{(glob)}$ with $r = 0.01$ (left) and $\mathcal{H}_1^{(loc)}$ with $\beta = 0.1$ (right) when $(N, M, B) = (2846, 290, 580))$

.

4. Proof of Theorem 1

We will detail in this section the main steps to prove Theorem 1, while some details will be left in the Appendix.

4.1. General approach

First, we notice that the frequency smoothed estimate $\hat{s}_{i,j}(\nu)$ can be written as

$$\hat{s}_{ij}(\nu) = \frac{1}{B+1} \boldsymbol{\xi}_{y_j}(\nu)^* \boldsymbol{\xi}_{y_i}(\nu) \tag{13}$$

where

$$\boldsymbol{\xi}_{y_i}(\nu) = \left(\xi_{y_i}\left(\nu - \frac{B}{2N}\right), \ldots, \xi_{y_i}\left(\nu + \frac{B}{2N}\right)\right)^T.$$

This is a sesquilinear form of the finite Fourier transform of the M time series samples $(y_{i,1}, \ldots, y_{i,N})_{i\in[M]}$. To handle the statistical dependence between the components of $\boldsymbol{\xi}_{y_i}(\nu)$, we use the well-known Bartlett decomposition (see for instance Walker (1965)) whose procedure is described hereafter.

From Assumptions 2 and 3, the spectral distribution of $(y_{m,n})_{n\in\mathbb{Z}}$ is absolutely continuous with density s_m being uniformly bounded and bounded away from 0. Therefore, from Wold's Theorem (Brockwell and Davis, 2006, Th. 5.7.1, Th. 5.7.2), each time series $(y_{m,n})_{n\in\mathbb{Z}}$ admits a causal and causally invertible linear representation in terms of its normalized innovation sequence:

$$y_{m,n} = \sum_{k=0}^{+\infty} a_{m,k} \epsilon_{m,n-k}, \tag{14}$$

where $(\epsilon_{1,k})_{k\in\mathbb{Z}}, \ldots, (\epsilon_{M,k})_{k\in\mathbb{Z}}$ are mutually independent sequences of $\mathcal{N}_{\mathbb{C}}(0,1)$ i.i.d. random variables, and $(a_{1,k})_{k\in\mathbb{N}}, \ldots, (a_{M,k})_{k\in\mathbb{N}} \in \ell^2(\mathbb{N})$ such that if

$$h_m(\nu) = \sum_{k=0}^{+\infty} a_{m,k} e^{-2i\pi k\nu} \tag{15}$$

then $|h_m(\nu)|^2 = s_m(\nu)$ and $h_m(\nu)$ coincides with the outer causal spectral factor of $s_m(\nu)$. Define now $\tilde{s}_{ij}(\nu)$, an approximation of $\hat{s}_{ij}(\nu)$, as:

$$\tilde{s}_{ij}(\nu) = \frac{1}{B+1} \sum_{b=-B/2}^{B/2} h_i\left(\nu + \frac{b}{N}\right) \overline{h_j\left(\nu + \frac{b}{N}\right)} \xi_{\epsilon_i}\left(\nu + \frac{b}{N}\right) \overline{\xi_{\epsilon_j}\left(\nu + \frac{b}{N}\right)}$$

or equivalently

$$\tilde{s}_{ij}(\nu) = \boldsymbol{\xi}_{\epsilon_j}(\nu)^* \frac{\boldsymbol{\Pi}_{ij}(\nu)}{B+1} \boldsymbol{\xi}_{\epsilon_i}(\nu) \tag{16}$$

where

$$\boldsymbol{\Pi}_{ij}(\nu) = \mathrm{dg}\left(h_i\left(\nu + \frac{b}{N}\right) \overline{h_j\left(\nu + \frac{b}{N}\right)}\right)_{b=-B/2,\ldots,B/2} \tag{17}$$

and

$$\boldsymbol{\xi}_{\epsilon_i}(\nu) = \left(\xi_{\epsilon_i}\left(\nu - \frac{B}{2N}\right), \ldots, \xi_{\epsilon_i}\left(\nu + \frac{B}{2N}\right)\right)^T.$$

Instead of working directly with $|\hat{c}_{ij}(\nu)|^2 = \frac{|\hat{s}_{ij}(\nu)|^2}{\hat{s}_i(\nu)\hat{s}_j(\nu)}$, it turns out that it is more convenient to show the limiting Gumbel distribution for $\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)}$ where

$$\begin{aligned}
\sigma_{ij}^2(\nu) &= \frac{1}{B+1} \sum_{b=-B/2}^{B/2} \left|h_i\left(\nu + \frac{b}{N}\right)\right|^2 \left|h_j\left(\nu + \frac{b}{N}\right)\right|^2 \\
&= \frac{1}{B+1} \sum_{b=-B/2}^{B/2} s_i\left(\nu + \frac{b}{N}\right) s_j\left(\nu + \frac{b}{N}\right) \\
&:= \frac{\mathrm{tr}\ \boldsymbol{\Sigma}_{ij}(\nu)}{B+1}
\end{aligned} \tag{18}$$

and where

$$\mathbf{\Sigma}_{ij}(\nu) := \mathbf{\Pi}_{ij}^*(\nu)\mathbf{\Pi}_{ij}(\nu) = \mathrm{dg}\left(\left|h_i\left(\nu + \frac{b}{N}\right)\right|^2 \left|h_j\left(\nu + \frac{b}{N}\right)\right|^2, b = -\frac{B}{2}, \ldots, \frac{B}{2}\right). \quad (19)$$

This is the aim of Proposition 1 below.

Proposition 1 (Gumbel limit for $\max_{(i,j,\nu)\in\mathcal{I}} |\tilde{s}_{ij}(\nu)|^2$)**.** *Under Assumptions 1 – 3, for any $t \in \mathbb{R}$, we have*

$$\mathbb{P}\left(\max_{(i,j,\nu)\in\mathcal{I}}(B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} \leq t + \log\frac{N}{B+1} + \log\frac{M(M-1)}{2}\right) \xrightarrow[N\to+\infty]{} e^{-e^{-t}}. \quad (20)$$

Once equipped with Proposition 1, it remains then to show that $\max_{(i,j,\nu)\in\mathcal{I}} \frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)}$ is close enough to $\max_{(i,j,\nu)\in\mathcal{I}} |\hat{c}_{ij}(\nu)|^2$ to prove that these quantities have the same limiting distribution. This result is given by the following Proposition.

Proposition 2. *Under Assumptions 1 – 3, as $N \to \infty$,*

$$\max_{(i,j,\nu)\in\mathcal{I}}(B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} - \max_{(i,j,\nu)\in\mathcal{I}}(B+1)|\hat{c}_{ij}(\nu)|^2 = o_P(1).$$

As Theorem 1 is directly obtained by Proposition 1, Proposition 2 and an application of Slutsky's lemma, the two remaining subsections are devoted to the proofs of Proposition 1 and Proposition 2.

4.2. Proof of Proposition 1

To prove Proposition 1, the main tool is Lemma A.4 from Jiang et al. (2004), which is a special case of Poisson approximation from Arratia et al. (1989). We rewrite it here for the sake of completeness.

Lemma 2. *Let $(X_\alpha)_{\alpha\in\mathcal{I}}$ be a finite collection of Bernoulli random variables, and for each $\alpha \in \mathcal{I}$, let $\mathcal{I}_\alpha \subset \mathcal{I}$ such that $\alpha \in \mathcal{I}_\alpha$. Then,*

$$\left|\mathbb{P}\left(\sum_{\alpha\in\mathcal{I}} X_\alpha = 0\right) - \exp\left(-\sum_{\alpha\in\mathcal{I}} \mathbb{P}(X_\alpha = 1)\right)\right| \leq \Delta_1 + \Delta_2 + \Delta_3$$

where

$$\Delta_1 = \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in \mathcal{I}_\alpha} \mathbb{P}\left(X_\alpha = 1\right) \mathbb{P}\left(X_\beta = 1\right)$$
$$\Delta_2 = \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in \mathcal{I}_\alpha \setminus \{\alpha\}} \mathbb{P}\left(X_\alpha = 1,\, X_\beta = 1\right)$$
$$\Delta_3 = \sum_{\alpha \in \mathcal{I}} \mathbb{E} \left| \mathbb{P}\left(X_\alpha = 1 | \left(X_\beta\right)_{\beta \in \mathcal{I} \setminus \mathcal{I}_\alpha}\right) - \mathbb{P}(X_\alpha = 1) \right|$$

In particular, if for each $\alpha \in \mathcal{I}$, X_α is independent of $\{X_\beta : \beta \in \mathcal{I} \setminus \mathcal{I}_\alpha\}$, then $\Delta_3 = 0$.

Lemma 2 is the keystone for the proof of Proposition 1, and is a standard tool for analyzing distributions of maxima of dependent random variables. We now prove Proposition 1.

Proof. We start by proving (20). Define

$$t_N = \sqrt{x + \log \frac{M(M-1)}{2} + \log \frac{N}{B+1}} \tag{21}$$

and for $(i, j, \nu) \in \mathcal{I}$ (recall that $\mathcal{I}$ is defined in (5), and that it depends on N, but in order to avoid cumbersome notations we do not recall this dependency) the Bernoulli random variables $X_{ij}(\nu)$ as

$$X_{ij}(\nu) := \mathbb{1}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2\right). \tag{22}$$

Define the set $\mathcal{I}_{(i,j,\nu)}$

$$\mathcal{I}_{(i,j,\nu)} = \{(i', j', \nu) : 1 \leq i' < j' \leq M, i = i' \text{ or } j = j'\}. \tag{23}$$

From (16) and under Assumption 1, if $(i', j', \nu') \in \mathcal{I} \setminus \mathcal{I}_{(i,j,\nu)}$, then $\tilde{s}_{i'j'}(\nu')$ is independent from $\tilde{s}_{ij}(\nu)$ since we have either

(1) $i' \neq i$, $j' \neq j$, $\nu' = \nu$;
(2) $i' = i$ or $j' = j$, and $\nu' \neq \nu$ (implying $|\nu - \nu'| > \frac{B}{N}$ by assumption), in which case $\left(\boldsymbol{\xi}_{\epsilon_{i'}}(\nu'), \boldsymbol{\xi}_{\epsilon_{j'}}(\nu')\right)$ is independent from $\left(\boldsymbol{\xi}_{\epsilon_i}(\nu), \boldsymbol{\xi}_{\epsilon_j}(\nu)\right)$.

From the definition of $X_{ij}(\nu)$ in (22),

$$\mathbb{P}\left((B+1)\max_{(i,j,\nu)\in\mathcal{I}}\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} \le t_N^2\right) = \mathbb{P}\left(\sum_{(i,j,\nu)\in\mathcal{I}} X_{ij}(\nu) = 0\right)$$

which can be estimated by Lemma 2 as:

$$\left|\mathbb{P}\left(\sum_{(i,j,\nu)\in\mathcal{I}} X_{ij}(\nu) = 0\right) - e^{-\lambda}\right| \le \Delta_1 + \Delta_2 + \Delta_3$$

where

$$\lambda = \sum_{(i,j,\nu)\in\mathcal{I}} \mathbb{P}\left(X_{ij}(\nu) = 1\right) = \sum_{(i,j,\nu)\in\mathcal{I}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2\right)$$

and

$$\Delta_1 = \sum_{(i,j,\nu)\in\mathcal{I}} \sum_{(i',j',\nu)\in\mathcal{I}_{(i,j,\nu)}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{i,j}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2\right) \mathbb{P}\left((B+1)\frac{|\tilde{s}_{i'j'}(\nu)|^2}{\sigma_{i'j'}^2(\nu)} > t_N^2\right) \tag{24}$$

$$\Delta_2 = \sum_{(i,j,\nu)\in\mathcal{I}} \sum_{\substack{(i',j',\nu)\in\mathcal{I}_{(i,j,\nu)} \\ (i',j')\neq(i,j)}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{i,j}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2,\ (B+1)\frac{|\tilde{s}_{i'j'}(\nu)|^2}{\sigma_{i'j'}^2(\nu)} > t_N^2\right) \tag{25}$$

$$\begin{aligned}\Delta_3 = \sum_{(i,j,\nu)\in\mathcal{I}} \mathbb{E}\Bigg|\mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2 \middle|\, (\tilde{s}_{i'j'}(\nu'))_{(i',j',\nu')\in\mathcal{I}\setminus\mathcal{I}_{(i,j,\nu)}}\right) \\ -\mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2\right)\Bigg|.\end{aligned} \tag{26}$$

We now have to control the four quantities λ, Δ_1, Δ_2 and Δ_3, which requires studying moderate deviations results for

$$\mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2\right)$$

as well as

$$\mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma^2_{ij}(\nu)} > t_N^2,\ (B+1)\frac{|\tilde{s}_{i'j'}(\nu)|^2}{\sigma^2_{i'j'}(\nu)} > t_N^2\right)$$

for all $(i', j', \nu) \in \mathcal{I}_{(i,j,\nu)}$. The following Proposition 3, proved in Appendix C from Rosuel et al. (2021), provides exactly this.

Proposition 3. *Under Assumptions 1 – 3, there exists a constant $\eta > 0$ such that for any $C > 0$, we have*

$$\max_{t\in[0,CB^\eta]} \max_{(i,j,\nu)\in\mathcal{I}} \left|\mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma^2_{ij}(\nu)} > t^2\right) e^{t^2} - 1\right| \xrightarrow[N\to\infty]{} 0 \tag{27}$$

and

$$\max_{t,s\in[0,CB^\eta]} \max_{\substack{(i,j,\nu)\in\mathcal{I} \\ (i',j',\nu)\in\mathcal{I}_{(i,j,\nu)}}} \left|\mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma^2_{ij}(\nu)} > t^2, (B+1)\frac{|\tilde{s}_{i'j'}(\nu)|^2}{\sigma^2_{i'j'}(\nu)} > s^2\right) \times e^{t^2+s^2} - 1\right| \xrightarrow[N\to\infty]{} 0. \tag{28}$$

First, concerning $\exp(-\lambda)$, since t_N as defined in (21) is $\mathcal{O}(\log N)$, one can use Proposition 3 to get

$$\begin{aligned}
\exp(-\lambda) &= \exp\left(-\sum_{(i,j,\nu)\in\mathcal{I}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma^2_{ij}(\nu)} > t_N^2\right)\right) \\
&= \exp\left(-\frac{N}{B+1}\frac{M(M-1)}{2} e^{-t_N^2}(1+o(1))\right) \\
&\xrightarrow[N\to\infty]{} \exp\left(-\exp(-x)\right).
\end{aligned}$$

We now turn to the control of Δ_1, Δ_2, and Δ_3. Regarding Δ_3, since under Assumption 1 the random variables $\tilde{s}_{ij}(\nu)$ and $\tilde{s}_{i'j'}(\nu')$ for $(i', j', \nu') \in \mathcal{I}\backslash\mathcal{I}_{(i,j,\nu)}$ are independent, we clearly have $\Delta_3 = 0$. Consider now (24) and (25). The aim is to show that $\Delta_1 = o(1)$ and $\Delta_2 = o(1)$ when t_N is defined by (21). Using the moderate deviation result (27) from Proposition 3, and recalling that C represents a universal constant independent of N whose value can change from one line to another, we get:

$$\begin{aligned}
\Delta_1 &\leq \underbrace{|\mathcal{I}|}_{\mathcal{O}(\frac{N}{B}M^2)} \underbrace{\max_{(i,j,\nu)\in\mathcal{I}} |\mathcal{I}_{(i,j,\nu)}|}_{\mathcal{O}(M)} \max_{(i,j,\nu)\in\mathcal{I}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{i,j}(\nu)|^2}{\sigma_{i,j}(\nu)^2} > t_N^2\right)^2 \\
&\leq C\frac{N}{B}M^3 \underbrace{e^{-2t_N^2}}_{\mathcal{O}(\frac{1}{M^4}\frac{B^2}{N^2})} \underbrace{\max_{(i,j,\nu)\in\mathcal{I}} \left(\mathbb{P}\left[(B+1)\frac{|\tilde{s}_{i,j}(\nu)|^2}{\sigma_{i,j}(\nu)^2} > t_N^2\right] e^{t_N^2}\right)^2}_{=1+o(1)} \\
&= \mathcal{O}\left(\frac{1}{N}\right).
\end{aligned}$$

Δ_2 is handled similarly with equation (28) from Proposition 3:

$$\begin{aligned}
\Delta_2 = &\sum_{(i,j,\nu)\in\mathcal{I}} \sum_{(i',j',\nu)\in\mathcal{I}_{(i,j,\nu)}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} > t_N^2,\ (B+1)\frac{|\tilde{s}_{i'j'}(\nu)|^2}{\sigma_{i'j'}^2(\nu)} > t_N^2\right) \\
&\leq |\mathcal{I}| \max_{(i,j,\nu)\in\mathcal{I}} |\mathcal{I}_{(i,j,\nu)}|\ e^{-2t_N^2} \\
&\quad \times \underbrace{\max_{(i,j,\nu)\in\mathcal{I}} \max_{(i',j',\nu)\in\mathcal{I}_{(i,j,\nu)}} \mathbb{P}\left((B+1)\frac{|\tilde{s}_{i,j}(\nu)|^2}{\sigma_{i,j}(\nu)^2} > t_N^2, (B+1)\frac{|\tilde{s}_{i'j'}(\nu)|^2}{\sigma_{i'j'}^2(\nu)} > t_N^2\right) e^{2t_N^2}}_{=1+o(1)} \\
&= \mathcal{O}\left(\frac{1}{N}\right).
\end{aligned}$$

The proof of (20) is complete. □

4.3. Proof of Proposition 2

To prove Proposition 2, ie. the fact that $\max_{(i,j,\nu)\in\mathcal{I}} \frac{|\hat{s}_{ij}(\nu)|^2}{\hat{s}_i(\nu)\hat{s}_j(\nu)}$ and $\max_{(i,j,\nu)\in\mathcal{I}} \frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)}$ are close enough in probability, we work separately on the numerator and the denominator. This constitutes the statement of the two following propositions.

Proposition 4 (Change of numerator)**.** *Under Assumptions 1 – 3, there exists $\delta > 0$ such that as $N \to \infty$,*

$$\sqrt{B+1} \max_{(i,j,\nu)\in\mathcal{I}} |\hat{s}_{ij}(\nu) - \tilde{s}_{ij}(\nu)| = \mathcal{O}_P(N^{-\delta}). \tag{29}$$

The proof is deferred to Appendix B from Rosuel et al. (2021). A consequence of Proposition 4 and Proposition 1 is that

$$\begin{aligned}
&\sqrt{B+1} \max_{(i,j,\nu)\in\mathcal{I}} |\hat{s}_{ij}(\nu)| \\
&\qquad \leq \sqrt{B+1} \max_{(i,j,\nu)\in\mathcal{I}} |\tilde{s}_{i,j}(\nu)| + \sqrt{B+1} \max_{(i,j,\nu)\in\mathcal{I}} |\hat{s}_{ij}(\nu) - \tilde{s}_{ij}(\nu)| \\
&\qquad = \mathcal{O}_P\left(\sqrt{\log N}\right). \qquad (30)
\end{aligned}$$

Proposition 5 (Change of denominator)**.** *Under Assumption 2, for any* $\epsilon > 0$*, as* $N \to \infty$*,*

$$\max_{(i,j,\nu)\in\mathcal{I}} \left|\hat{s}_i(\nu)\hat{s}_j(\nu) - \sigma_{ij}^2(\nu)\right| = \mathcal{O}_P\left(\frac{B}{N} + \frac{N^\epsilon}{\sqrt{B}}\right). \qquad (31)$$

Moreover,

$$0 < \inf_{N\geq 1} \min_{(i,j,\nu)\in\mathcal{I}} \sigma_{ij}^2(\nu) \leq \sup_{N\geq 1} \max_{(i,j,\nu)\in\mathcal{I}} \sigma_{ij}^2(\nu) < +\infty \qquad (32)$$

and

$$\max_{i\in[M]} \max_{\nu\in\mathcal{G}} \frac{1}{\hat{s}_i(\nu)} = O_P(1), \quad \max_{i\in[M]} \max_{\nu\in\mathcal{G}} \hat{s}_i(\nu) = O_P(1). \qquad (33)$$

The proof is deferred to Appendix A in Rosuel et al. (2021). We recall that for any sequences (a_n) and (b_n), the following inequality holds:

$$\left|\sup_n a_n - \sup_n b_n\right| \leq \sup_n |a_n - b_n|.$$

Therefore, to show that Proposition 2 holds, it is enough to show that

$$\max_{(i,j,\nu)\in\mathcal{I}} \left|(B+1)\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} - (B+1)|\hat{c}_{ij}(\nu)|^2\right| = o_P(1).$$

This result could be proved by writing the following decomposition:

$$(B+1) \max_{(i,j,\nu)\in\mathcal{I}} \left|\frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} - \frac{|\hat{s}_{ij}(\nu)|^2}{\hat{s}_i(\nu)\hat{s}_j(\nu)}\right| \leq \Psi_3(\Psi_1 + \Psi_2).$$

where

$$\Psi_1 := (B+1) \max_{(i,j,\nu)\in\mathcal{I}} \left| |\hat{s}_{ij}(\nu)|^2 - |\tilde{s}_{ij}(\nu)|^2 \right| \hat{s}_i(\nu)\hat{s}_j(\nu)$$
$$\Psi_2 := (B+1) \max_{(i,j,\nu)\in\mathcal{I}} |\hat{s}_{ij}(\nu)|^2 |\hat{s}_i(\nu)\hat{s}_j(\nu) - \sigma_{ij}^2(\nu)|$$
$$\Psi_3 := \max_{(i,j,\nu)\in\mathcal{I}} \frac{1}{\hat{s}_i(\nu)\hat{s}_j(\nu)\sigma_{ij}^2(\nu)}.$$

It is clear by (20) that

$$\max_{(i,j,\nu)\in\mathcal{I}} (B+1)|\tilde{s}_{ij}(\nu)|^2 = \mathcal{O}_P(\log N).$$

Combining this with Proposition 5 and equation (29) from Proposition 4, there exists $\delta > 0$ such that

$$(B+1) \max_{(i,j,\nu)\in\mathcal{I}} \left| |\hat{s}_{ij}(\nu)|^2 - |\tilde{s}_{ij}(\nu)|^2 \right| \leq \underbrace{\sqrt{B+1} \max_{(i,j,\nu)\in\mathcal{I}} (|\hat{s}_{ij}(\nu)| + |\tilde{s}_{ij}(\nu)|)}_{=\mathcal{O}_P(\sqrt{\log N})} \times \underbrace{\sqrt{B+1} \max_{(i,j,\nu)\in\mathcal{I}} ||\hat{s}_{ij}(\nu)| - |\tilde{s}_{ij}(\nu)||}_{=\mathcal{O}_P(N^{-\delta})}$$

which is $\mathcal{O}_P(\sqrt{\log N} N^{-\delta})$. Using (33), this implies that

$$\Psi_1 = \mathcal{O}_P\left(\sqrt{\log N} N^{-\delta}\right).$$

Similarly, using Proposition 5, for any $\epsilon > 0$,

$$\Psi_2 = \mathcal{O}_P\left(\log N \left(\frac{B}{N} + \frac{N^\epsilon}{\sqrt{B}}\right)\right)$$
$$\Psi_3 = \mathcal{O}_P(1).$$

Combining the estimates of Ψ_1, Ψ_2 and Ψ_3 we get that for any $\epsilon > 0$:

$$(B+1) \max_{(i,j,\nu)\in\mathcal{I}} \left| \frac{|\tilde{s}_{ij}(\nu)|^2}{\sigma_{ij}^2(\nu)} - \frac{|\hat{s}_{ij}(\nu)|^2}{\hat{s}_i(\nu)\hat{s}_j(\nu)} \right| = \mathcal{O}_P\left(N^{-\delta}\sqrt{\log N} + \log N \left(\frac{B}{N} + \frac{N^\epsilon}{\sqrt{B}}\right)\right).$$

This quantity is $o_P(1)$ if $\frac{N^\epsilon}{\sqrt{B}} = o(1)$ which is satisfied by choosing $\epsilon < \frac{\rho}{2}$ from Assumption (4).

References

Arratia, R., Goldstein, L., Gordon, L., et al., 1989. Two moments suffice for poisson approximations: the chen-stein method. The Annals of Probability 17, 9–25.

Brockwell, P.J., Davis, R.A., 2006. Time series: theory and methods. Springer Series in Statistics, Springer, New York. Reprint of the second (1991) edition.

Cai, T.T., Ma, Z., et al., 2013. Optimal hypothesis testing for high dimensional covariance matrices. Bernoulli 19, 2359–2388.

Chang, J., Yao, Q., Zhou, W., 2017. Testing for high-dimensional white noise using maximum cross-correlations. Biometrika 104, 111–127. URL: https://doi.org/10.1093/biomet/asw066, doi:10.1093/biomet/asw066, arXiv:https://academic.oup.com/biomet/article-pdf/104/1/

Dette, H., Dörnemann, N., 2020. Likelihood ratio tests for many groups in high dimensions. Journal of Multivariate Analysis , 104605.

Eichler, M., 2008. Testing nonparametric and semiparametric hypotheses in vector stationary processes. Journal of Multivariate Analysis 99, 968–1009.

Embrechts, P., Klüppelberg, C., Mikosch, T., 2013. Modelling extremal events: for insurance and finance. volume 33. Springer Science & Business Media.

Fan, J., Jiang, T., et al., 2019. Largest entries of sample correlation matrices from equi-correlated normal populations. The Annals of Probability 47, 3321–3374.

Fiecas, M., Leng, C., Liu, W., Yu, Y., 2019. Spectral analysis of high-dimensional time series. Electronic Journal of Statistics 13, 4079–4101.

Fiecas, M., von Sachs, R., 2014. Data-driven shrinkage of the spectral density matrix of a high-dimensional time series. Electronic Journal of Statistics 8, 2975–3003.

Jiang, T., et al., 2004. The asymptotic distributions of the largest entries of sample correlation matrices. The Annals of Applied Probability 14, 865–880.

Li, Z., Lam, C., Yao, J., Yao, Q., 2019. On testing for high-dimensional white noise. The Annals of Statistics 47, 3382 – 3412. URL: https://doi.org/10.1214/18-AOS1782, doi:10.1214/18-AOS1782.

Lin, Z., Liu, W., 2009. On maxima of periodograms of stationary processes. The Annals of Statistics , 2676–2695.

Liu, W., Wu, W.B., 2010. Asymptotics of spectral density estimates. Econometric Theory , 1218–1245.

Loubaton, P., Rosuel, A., 2021. Large random matrix approach for testing independence of a large number of gaussian time series, v4. arXiv:2007.08806.

Mestre, X., Vallet, P., 2017. Correlation tests and linear spectral statistics of the sample correlation matrix. IEEE Transactions on Information Theory 63, 4585–4618.

Morales-Jimenez, D., Johnstone, I.M., McKay, M.R., Yang, J., 2018. Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models. arXiv preprint arXiv:1810.10214 .

Ombao, H.C., Raz, J.A., Strawderman, R.L., Von Sachs, R., 2001. A simple generalised crossvalidation method of span selection for periodogram smoothing. Biometrika 88, 1186–1192.

Pan, G., Gao, J., Yang, Y., 2014. Testing independence among a large number of high-dimensional random vectors. Journal of the American Statistical Association 109, 600–612.

Resnick, S.I., 2013. Extreme values, regular variation and point processes. Springer.

Rosuel, A., Loubaton, P., Vallet, P., 2021. On the asymptotic distribution of the maximum sample spectral coherence of gaussian time series in the high dimensional regime. arXiv:2107.02891.

Rosuel, A., Vallet, P., Loubaton, P., Mestre, X., 2020. On the frequency domain detection of high dimensional time series, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE. pp. 8782–8786.

Rudzkis, R., 1985. On the distribution of the maximum deviation of the gaussian stationary time series spectral density estimate. Lithuanian Mathematical Journal 25, 18–130.

Shao, X., Wu, W.B., et al., 2007. Asymptotic spectral theory for nonlinear time series. The Annals of Statistics 35, 1773–1801.

Wahba, G., 1971. Some tests of independence for stationary multivariate time series. Journal of the Royal Statistical Society: Series B (Methodological) 33, 153–166.

Walker, A., 1965. Some asymptotic results for the periodogram of a stationary time series. Journal of the Australian Mathematical Society 5, 107–128.

Woodroofe, M.B., Van Ness, J.W., 1967. The maximum deviation of sample spectral densities. The Annals of Mathematical Statistics , 1558–1569.

Wu, W.B., Zaffaroni, P., 2018. Asymptotic theory for spectral density estimates of general multivariate time series. Econometric Theory 34, 1–22.