



Largest magnitude for off-diagonal auto-correlation coefficients in high dimensional framework

Maxime Boucher, Didier Chauveau, Marguerite Zani

► To cite this version:

Maxime Boucher, Didier Chauveau, Marguerite Zani. Largest magnitude for off-diagonal auto-correlation coefficients in high dimensional framework. 2024. hal-03375111v2

HAL Id: hal-03375111

<https://hal.science/hal-03375111v2>

Preprint submitted on 20 Jun 2024

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.

Largest magnitude for off-diagonal auto-correlation coefficients in high dimensional framework

Maxime Boucher ^{*} [†], Didier Chauveau ^{*}, Marguerite Zani ^{*‡}

Abstract :

This paper studies the τ -coherence of a $(n \times p)$ -observation matrix in a Gaussian framework where both n and p are large and $p \gg n$. The τ -coherence is defined as the largest magnitude, outside a band of size $\tau = \tau(n)$, of the empirical correlation coefficients associated to the observations. Using the Chen-Stein method we show the convergence of the normalized coherence towards a Gumbel distribution. We broaden previous results by considering a 3-regime band structure for the correlation matrix, where the largest band is composed of asymptotically vanishing coefficients. We provide an hypothesis test on the correlation structure where the alternative shows a clear dichotomy on the vanishing band. Moreover, we provide numerical simulations illustrating the asymptotic behaviour of the coherence with Monte-Carlo experiment. We use a splitting strategy computing correlation matrices by blocks in order to avoid the high-dimensional memory issue.

Key words : Chen-Stein method, coherence, Gaussian high-dimensional matrices, Banded correlation.

1 Introduction

Random matrix theory has known a great amount of breakthroughs for these last decades. Developments have been made in theoretical fields as well as in various applied domains such as high-energy physics (e.g. [For10] on log-gases), electric engineering (signal and imaging, see [Don06, CT05, CRT06b, CRT06a]), statistics (see [Joh01, Joh08, BG16]). Earlier works on random matrices were focused on spectral analysis of eigenvalues and eigenvectors (see [Wig58] or [Meh04, BS10], see also [BC12] and references therein). For a reference on random matrix theory, see [BS10, Meh04, AGZ10].

In high dimensional statistics in particular, random matrices are widely used : high dimensional regression, hypothesis testing for high dimension parameters, inference for large covariance matrices. See e.g. [BS96, CT07, BJYZ09, CWX10a, CZZ10, BRT09]. In these contexts where the dimension p can be of the same or even higher magnitude than the sample size n , one faces new statistical challenges and classical methods no longer apply (see e.g. [Don00, JT09, EK18]). Empirical covariance or auto-correlation matrices are valuable to understand the dependences of the data set. In this paper we focus on the coherence, i.e. the supremum of the off-diagonal terms of (auto)correlation matrices. We consider the case where both n and p are increasing, p being much larger than n . We allow p to depend on n , this dependence is denoted by p_n . Hence in the sequel we will write indifferently p or p_n . In our model the correlation matrix has

^{*}Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, Route de Chartres, B.P. 6759, 45067, UMR 7013, Orléans cedex 2, France

[†]Département de Mathématiques, Université Libre de Bruxelles, CP201 Boulevard du Triomphe, B-1050, Bruxelles, Belgium.

[‡]Corresponding author: marguerite.zani@univ-orleans.fr

a 3-regime bandwise structure. Indeed the correlation matrix is composed of three contiguous off-diagonal bands with respectively constant, decreasing to 0 and null coefficients. Model and motivation are detailed below.

Autocorrelation matrices and coherence

Let ${}^t(X^1, X^2, \dots, X^p)$ be a p -dimensional Gaussian random vector with mean ${}^t(\mu^1, \dots, \mu^p)$ in \mathbb{R}^p and covariance matrix $\Sigma \in \mathbb{R}^p \times \mathbb{R}^p$. We consider a random sample $\{{}^t(X_i^1, X_i^2, \dots, X_i^p)\}_{i=1, \dots, n}$ issued from (X^1, X^2, \dots, X^p) , arranged in a (n, p) -matrix \mathbb{X}_n . We denote by \mathbb{X}_n^k the k^{th} column of \mathbb{X}_n . The classical empirical Pearson's correlation coefficient writes:

$$\rho_{kj} = \frac{\sum_{i=1}^n (X_i^k - \overline{\mathbb{X}_n^k}) (X_i^j - \overline{\mathbb{X}_n^j})}{\sqrt{\sum_{i=1}^n (X_i^k - \overline{\mathbb{X}_n^k})^2} \sqrt{\sum_{i=1}^n (X_i^j - \overline{\mathbb{X}_n^j})^2}} \quad (1)$$

where $\overline{\mathbb{X}_n^k}$ is the empirical mean of the k^{th} column \mathbb{X}_n^k :

$$\overline{\mathbb{X}_n^k} = \frac{1}{n} \sum_{i=1}^n X_i^k.$$

Definition 1.1. With the notations above, we can define the largest magnitude of the off-diagonal terms of the correlation matrix :

$$L_n = \max_{1 \leq k < j \leq p} |\rho_{kj}|. \quad (2)$$

In signal processing L_n is defined as the coherence and used as an indicator of the sparsity of a matrix, appearing in the Mutual Incoherence Property (MIP). For details on this approach, see Donoho and Huo [DH01], Fuchs [Fuc04], Cai, Wang and Xu [CWX10b], and references therein.

Major theoretical results concerning coherence are due to Jiang et al. In [Jia04] he first adressed this problem and showed strong consistency of L_n and limit distribution of L_n^2 in the case where n and p are of the same order. Moments assumptions in [Jia04] and dimension for p were substantially improved by a series of papers: Li and Rosalsky [LR06], Zhou [Zho07], Liu, Lin and Shao [LLS08], Li, Liu and Rosalsky [LLR10], Li, Qi and Rosalsky [LQR12], [CJ11a]. In [CJ12] the authors consider the limiting distribution of the coherence in a spherical case. See also [CZ16] for studies on the differential correlation matrices in high dimensions.

In [CJ11a] (see also the supplement [CJ11b]) they consider more specifically the sub-exponential regime $\log p_n = o(n^\alpha)$ for some $0 < \alpha \leq 1/3$. They introduce the so-called τ -coherence aimed to test whether the covariance has a given bandwidth $\tau > 1$, and $\tau = 1$ would be a special case. The τ -coherence is defined below.

Definition 1.2. For any integer $\tau \geq 1$, the τ -coherence is:

$$L_{n,\tau} = \max_{|k-j| \geq \tau} |\rho_{kj}| \quad (3)$$

In [CJ11a] strong laws and convergence of distributions of $L_{n,\tau}$ are given as well. Recently, Shao and Zhou [SZ14] studied coherence and τ -coherence relaxing the Gaussian hypothesis, improving assumptions on the moments of the entries and on the dimension p .

As pointed out by [CJ11a], interest in τ -coherence can be relied to econometrics models, see [And91, LB95]. Moreover, recent works related to high dimensional problems adress this bandwise structure question (e.g. in [BL08b, BL08a, WD14]).

Problem addressed and Motivation

In these contexts of longitudinal data X^k and X^j being independent or less correlated as $|k - j|$ increases, we consider a correlation matrix $R = (R_{kj})_{1 \leq k, j \leq p}$ defined as follows:

$$R_{kj} = \begin{cases} r_{kj} & \text{if } 0 < |k - j| < \tau, \\ \epsilon_n & \text{if } \tau \leq |k - j| \leq \tau + K, \\ 0 & \text{if } \tau + K < |k - j|. \end{cases} \quad (4)$$

The off diagonal has a first band of size τ with $\tau = o(p_n^t)$ for any $t > 0$. Furthermore the set of large coefficients r_{kj} is small (see further Hypothesis 3 in Theorem 2.1). The second band is much larger (acting like a transitional regime) and coefficients ϵ_n are decreasing to 0. This band has size $K = O(p_n^\nu)$. Here $\nu \in]0, c]$ and c is an explicit constant depending on the coefficients r_{kj} and on the speed of convergence of ϵ_n to 0. Let us notice here that the rate of convergence to 0 of this transitional band ϵ_n can be matched to the threshold in [BL08a]. Finally, when $|k - j| > \tau + K$ the correlation coefficients R_{kj} are assumed to be 0.

This model is motivated by a more realistic approach of the covariance structure. We try to find a transition from a banded structure $1/(r_{jk})/0$ to a model $1/(r_{jk})/(\epsilon_n)/0$ where the correlation coefficients ϵ_n are spread on a larger bandwidth than the r_{jk} . This study allows us to exhibit a test where the asymptotic power is 1 when the convergence of ϵ_n to zero goes below a certain speed. In this framework, we study the asymptotic behaviour of $L_{n,\tau}$. Our work strongly relies on the Chen-Stein method which is a Poisson approximation of weakly dependent events. For references on this method, see [AGG89], [Pec12] and references therein.

To conclude the bibliographical review we cite Mikosh et al [HMY21] which recently used the point process convergence of some i.i.d. random walks to show the limiting Gumbel distribution highlighted in the different papers mentioned above. Moreover, a recent work of Fan and Jiang [FJ19] explores the transitional behaviour of the coherence when the variables go from equi-correlated to non correlated.

Our paper is organized as follows: Section 2 presents the main convergence results while Section 3 is devoted to the statistical test on the structure of the correlation. Section 4 gives some simulation results for our model. Section 5 presents some discussion and perspectives. Section 6 is devoted to the proof of the main result. Finally, an Appendix gather some technical results and proofs of technical lemmas.

The usual notations $u_n = o(v_n)$ and $u_n = O(v_n)$ stand for u_n negligible with respect to v_n and u_n of the same order of v_n respectively and asymptotically when $n \rightarrow \infty$. Moreover, without loss of generality, we can assume that ${}^t(\mu^1, \dots, \mu^p) = 0_{\mathbb{R}^p}$

2 Main result

Before stating our main result, we precise our assumptions. In particular we fix the number of "big" correlations coefficients : we assume that the correlation matrix has few off-diagonal coefficients having values close to 1 or -1 . For this purpose, we define the following set:

Definition 2.1. For any $\delta \in]0, 1[$ we define by

$$\Gamma_{p,\delta} = \{k \in \llbracket 1; p \rrbracket : |r_{kj}| > 1 - \delta \text{ for some } j \in \llbracket 1; p \rrbracket \text{ and } k \neq j\},$$

Assumption 1. Let n be an integer, $p = p_n$ a sequence such that $p_n \xrightarrow{n \rightarrow +\infty} +\infty$. Let $(\epsilon_n)_{n \in \mathbb{N}^*}$ be a sequence of real number in $] -1, 1[$. Let us assume the following conditions :

Hyp 1 : $\log(p_n) = o(n^{\frac{1}{3}})$ as $n \rightarrow +\infty$

Hyp 2 : $\tau = \tau(n) = o(p_n^t)$ as $n \rightarrow +\infty$ for any $t > 0$.

Hyp 3 : $\exists \delta \in]0, 1[$ such that $|\Gamma_{p,\delta}| = o(p_n)$ where $|\cdot|$ denotes the cardinality of the set.

Hyp 4 : $\epsilon_n \sim \gamma \sqrt{\frac{\log(p_n)}{n}}$ as $n \rightarrow +\infty$, $\gamma \in] -2 + \sqrt{2}, 2 - \sqrt{2}[$

Hyp 5 : $K = K(n) = O(p_n^\nu)$ where $\nu \in]0, c(\gamma, \delta)[$ and $c(\gamma, \delta) = \min\left(\frac{1}{3}(\frac{1}{2}\gamma^2 - 2|\gamma| + 1), \frac{\delta^2(2-\delta)^2}{36}\right)$.

Theorem 2.1. Under Assumptions (1) we have

$$nL_{n,\tau}^2 - 4\log(p_n) + \log(\log(p_n)) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z \quad (5)$$

where Z has the cdf $F(y) = e^{-\frac{1}{\sqrt{8\pi}}e^{-\frac{y}{2}}}$ for all $y \in \mathbb{R}$.

A direct corollary is the convergence, in probability, of the τ -coherence :

Corollaire 2.1. Under the framework of Theorem 2.1, we have

$$L_{n,\tau} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0 \quad (6)$$

and

$$\frac{n}{\log(p)} L_{n,\tau}^2 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 4 \quad (7)$$

Remark 1. We can sharpen the above results as follows: under Hypothesis 1 to 3 of Assumption 1 and with

$$\epsilon_n \sim \sqrt{\frac{\log(p_n)}{n^h}} \text{ with } h \geq 1$$

$$K = K(n) = O(p_n^\nu) \quad \nu \in]0, \frac{\delta^2(2-\delta)^2}{36}[$$

asymptotics (5), (6), (7) still hold.

Now we focus on the case where some coefficient (ϵ_n) are decreasing to 0 with a faster speed. We have the following proposition.

Theorem 2.2. We assume a bandwise structure as in 4, and that there exists (j, k) in $\{|j - k| \geq \tau\}$ such that (X^j, X^k) is a bivariate normal distribution $\mathcal{N}_2(0, \begin{pmatrix} 1 & \epsilon_n \\ \epsilon_n & 1 \end{pmatrix})$ where

$$\epsilon_n \sim \sqrt{\frac{\log(p_n)}{n^h}} \text{ with } 0 < h < 1$$

Then

$$\frac{n}{\log(p)} L_{n,\tau}^2 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} +\infty \quad (8)$$

From the convergence above, we can derive a statistical test on the structure of correlation matrices.

3 Application on testing the covariance structure

Let us consider a random sample of size n from a $\mathcal{N}_p(\mu, \Sigma)$ distribution, arranged in a $n \times p$ matrix. We want to test the covariance structure of Σ . As already mentionned in [CJ11a] it can be of significant interest to test whether Σ is banded, i.e. to know if the longitudinal data, say X^k and X^j , are uncorrelated if $|k - j|$ is beyond a certain value. Hence we build our test as follows: for a given integer τ , and a fixed level $0 < \alpha < 1$ we wish to test

$$H_0 : r_{kj} = 0 \text{ when } |k - j| \geq \tau \quad (9)$$

$$\text{versus } H_1 : \exists r_{kj} = \epsilon_n \text{ where } |k - j| \geq \tau \text{ and } \epsilon_n \sim \sqrt{\frac{\log(p_n)}{n^h}} \text{ with } 0 < h < 1. \quad (10)$$

Proposition 3.1. Let us define the test

$$\Phi = \mathbb{1}\{L_{n,\tau}^2 \geq n^{-1}(4 \log p - \log \log p - \log(8\pi) - 2 \log \log(1 - \alpha)^{-1})\}$$

Under the conditions of Theorem 2.1 the test above has size α asymptotically. Moreover, under H_1 the power of the test is asymptotically 1.

Proof of Proposition 3.1: We have easily the expression of the quantile $q_{1-\alpha}$ of the Gumbel distribution F given in Theorem 2.1:

$$q_{1-\alpha} = -\log(8\pi) - 2 \log \log(1 - \alpha)^{-1}$$

which gives the rejection region for the asymptotic size α :

$$\{L_{n,\tau}^2 \geq n^{-1}(4 \log p - \log \log p - \log(8\pi) - 2 \log \log(1 - \alpha)^{-1})\}$$

Moreover, under H_1 we have from the Theorem 2.2 :

$$\frac{n}{\log(p)} L_{n,\tau}^2 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} +\infty \quad (11)$$

It means that the asymptotic power of our test will be 1. More precisely, under the H_1 alternative, we have :

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(n L_{n,\tau}^2 - 4 \log(p_n) + \log(\log(p_n)) \geq -2 \log \left(-\sqrt{8\pi} \log(1 - \alpha) \right) \right) = 1. \quad (12)$$

□

4 Numerical aspects

In this section, we provide some simulated examples to illustrate the behavior of our asymptotic result in practical simulations (n and p both large but finite). For this, we use the R Statistical Software [R C23]. A difficulty comes from the fact that in our context, we have to compute correlations of large matrices. We need Gaussian observation matrices of size $n \times p$ with $\log(p) = o\left(n^{\frac{1}{3}}\right)$. It means that for a large n , for example $n = 4000$, we will have $p \approx 45000$ taking $p = \left\lceil \exp\left(n^{\frac{1}{3.5}}\right) \right\rceil$ in our simulations (where $[x]$ is the integer part of x). For each $(n \times p)$ -observation matrix, we have to compute the $(p \times p)$ -correlation matrix to compute the τ -coherence. For the range of p that we consider, we can observe the evolution of the size of the $(p \times p)$ -matrix in Gb according to n in Figure 1. For example, with $n = 4000$ and $p = 44112$, we have, for correlation stored in double, a $14.5Gb$ $(p \times p)$ -matrix which is very large for a common computer. We must find a way to compute the τ -coherence without loading the entire $(p \times p)$ -correlation matrix in the computer memory (RAM).

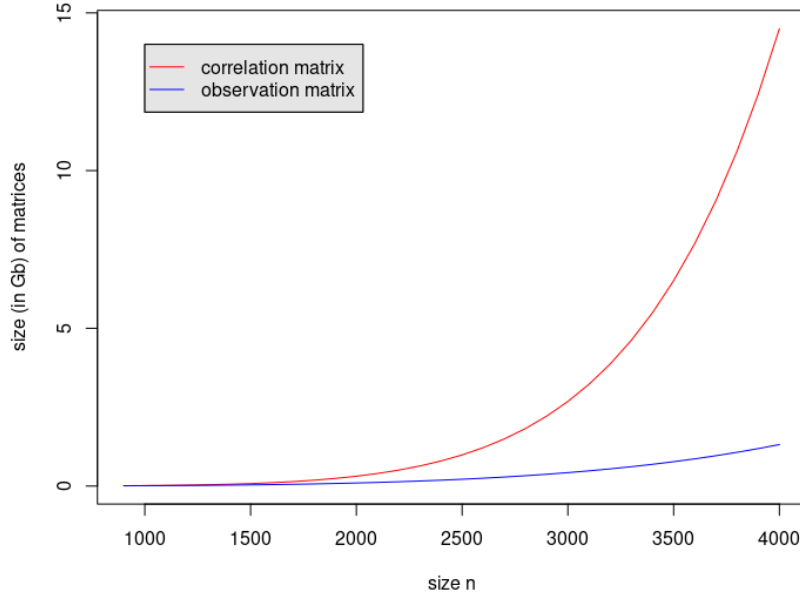


Figure 1: Size of $(p \times p)$ -correlation matrix and $(n \times p)$ -observation matrix according to n with $p = \left\lceil \exp(n^{1/3.5}) \right\rceil$ for real numbers stored as double precision numbers.

The idea is to generate the $(n \times p)$ -observation matrix by packets of columns. Each packet will have a size $(n \times Tb)$ where Tb is chosen depending on the available memory. With these packets of columns, we compute all correlation blocks of size $(Tb \times Tb)$ between each pair of packets of columns. In that way, we must choose a size Tb in order to have two blocks fitting simultaneously in the computer memory. Then, we can compute the τ -coherence by taking the largest coefficient in absolute value in our block paying attention to whether the block corresponds

to the central band (with bandwith τ) or not.

Using this strategy, we can generate correlation matrices for p as large as needed, so that we are able to study the limiting distribution of the τ -coherence. In that way, to illustrate our theorem, we consider the following parameters :

$$p = \left\lceil \exp \left(n^{1/3.5} \right) \right\rceil, \quad \tau = 5 * \lceil \log(p) \rceil, \quad K = 10 * \left\lceil n^{1/10} \log(p) \right\rceil, \quad \varepsilon_n = 0.1 * \sqrt{\frac{\log(p)}{n}}$$

Our purpose here is to simulate a sample of τ -coherence by a Monte-Carlo procedure in order to compare its empirical distribution with the asymptotic one. We thus run $R = 200$ replications of the following procedure, simulating R times matrices of observations and computing their correlations per blocks. For each replication, we generate an observation matrix \mathbf{X} of size $(n \times p)$ using the following numerical scheme :

$$\forall i \in \llbracket 1, n \rrbracket, \forall j \in \llbracket 1, p \rrbracket, X_i^j = \sum_{k=j}^{j+K-1} \varepsilon_n Y_i^k + \sum_{k=j+K}^{j+K+2\tau} r_k Y_i^k + \sum_{k=j+K+2\tau+1}^{j+2\tau+2K} \varepsilon_n Y_i^k \quad (13)$$

where all coefficients $(r_k)_{1 \leq k \leq 1+2\tau}$ are real numbers in $[-1, 1]$ (we take $r_1, \dots, r_{1+2\tau} \stackrel{i.i.d}{\sim} \mathcal{U}_{[-1,1]}$ in the simulation), X_i^j is the coefficient of \mathbf{X} on the i^{th} line and the j^{th} column and all the random variables $Y_i^k \stackrel{i.i.d}{\sim} \mathcal{N}(0, 1)$ arranged in a $(n \times (p + 2\tau + 2K))$ -matrix \mathbf{Y} . We highlight the fact that \mathbf{Y} is quite larger than \mathbf{X} . This numerical scheme is inspired by time series model.

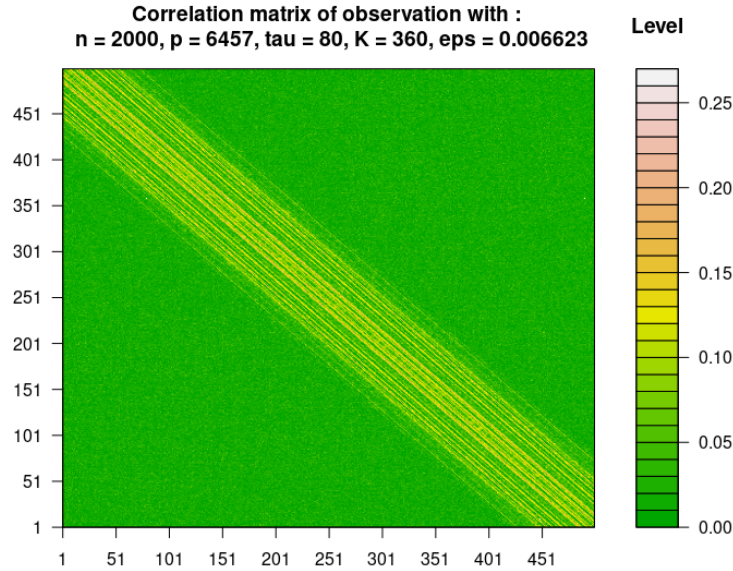


Figure 2: Level plot of the correlation's structure with observations : zoom on the square 1 : 500

We can observe that we generate data following our model and obtain an observation matrix associated to a correlation matrix with a band structure in Figure 2. We recognize a central band with non-null coefficients. Indeed, we also notice that the transition band with ε_n ' coefficients is not really recognizable but this is due to the fact that those coefficients are decreasing fastly to 0 when n goes to infinity (for instance, here, we have $\varepsilon_n \approx 0.007$ not different from 0 in the color scale).

With this observation matrix, we can use our procedure to compute the τ -coherence. After running R replications, we obtain a sample of τ -coherence. In Figure 3 and 4, we see that for n large enough, the sample distribution seems to approximate the limiting one.

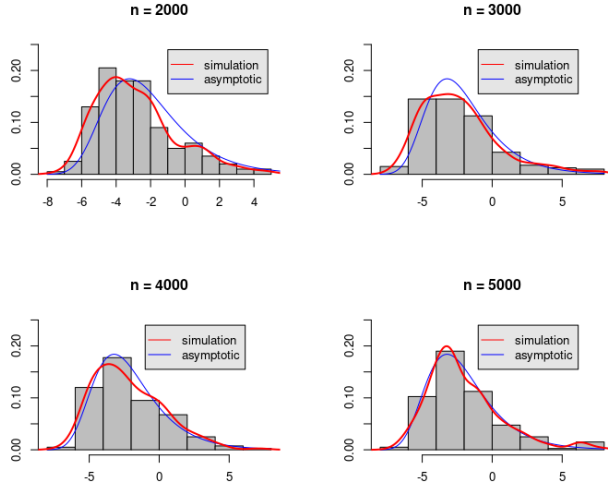


Figure 3: Histograms and Kernel density estimates for $n = 2000, 3000, 4000, 5000$ and for $R = 200$ replications

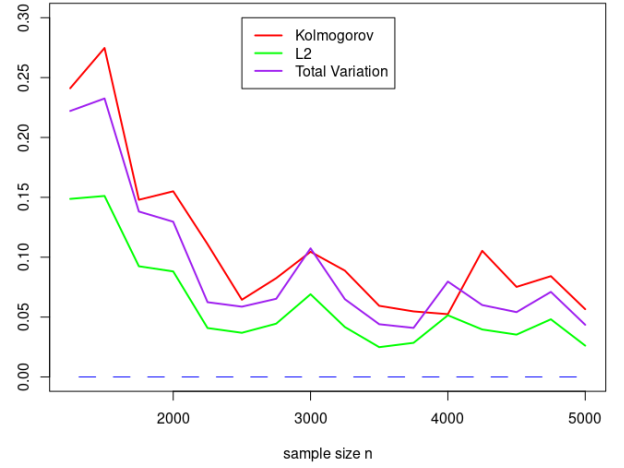


Figure 4: Evolution of Kolmogorov, \mathbb{L}^2 and Total Variation norm between simulated and asymptotic behavior

Precisely, we compare the estimated density of the sample (in red) with the asymptotic density (in blue) which is defined by $f(x) = \frac{1}{2\sqrt{8\pi}} \exp\left(-\frac{1}{2}y - \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{1}{2}y\right)\right)$ for all $x \in \mathbb{R}$. Also, in order to observe the convergence, we study numerically the distance between the sample and asymptotic distribution. We use the Kolmogorov, \mathbb{L}^2 and the Total Variation norms :

$$d_{KS}(\hat{f}, f) = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|, \quad d_2(\hat{f}, f) = \int \left| \hat{f}(x) - f(x) \right|^2 dx, \quad d_{TV}(\hat{f}, f) = \frac{1}{2} \int \left| \hat{f}(x) - f(x) \right| dx.$$

We observe in Figure 4, that the difference between both distributions decreases to 0 when n is increasing. These results provide numerical evidence that our limiting distribution is adequate. Moreover, we can illustrate the estimated power of our test. For this purpose, we keep the same regime for p .

We will consider $n = 1500, 2000, 2500, 3000$ and a size $\alpha = 0.05$. In addition, to illustrate the impact of the parameter h (and so the speed of (ε_n) coefficients), we compute esimated power for

Values for h	$\varepsilon_{\mathbf{n}} = \mathbf{0}$ (null)	$h = 2$	$h = 1$	$h = 0.9$	$h = 0.7$	$h = 0.5$	$h = 0.3$
$n = 1500$	0.045	0.045	0.050	0.055	0.055	0.080	0.835
$n = 2000$	0.035	0.033	0.030	0.035	0.035	0.055	0.950
$n = 2500$	0.03	0.03	0.03	0.03	0.04	0.07	1.00
$n = 3000$	0.065	0.065	0.055	0.055	0.055	0.070	0.995
$n = 3500$	0.03	0.035	0.045	0.04	0.07	1	1
$n = 4000$	0.055	0.055	0.06	0.07	0.075	1	1

Table 1: Estimated power computed on $R = 200$ repetitions for the level $\alpha = 0.05$. The first column corresponds to the null hypothesis while the other are for the alternative.

different values of n and for the values $h = 2, 1, 0.9, 0.7, 0.5, 0.3$. Also, we will consider the case when $\varepsilon_n = 0$ (i.e. simulation under the null hypothesis). We simulated $R = 200$ repetitions of τ -coherence for each value of n and p and for the different case of h . We sum up all the result in the following Table 1. We remark with this power estimations that we highlight (again) the fact that we faced a slow convergence. On the other way, for value of n until 3000, the sample size is not large enough to clearly identified the alternative for value of h close to 1 (it means close to the null hypothesis). It is no longer the case for a value of h such that we are far away from the null. For $n = 4000$, we observe an improvement for $h \leq 0.5$ while for $h = 0.9$ and $h = 0.7$ the sample size is not large enough to observe the empirical power equal to one.

We aware the reader about the high-dimensional framework impact on the simulation time. Indeed, we highlight the fact that the procedure we proposed here allows to compute τ -coherence corresponding to any large matrix \mathbb{X}_n arising in actual (big) data experiments. However, this procedure is not very efficient if it is done with a classical programming. For example, computing only one replication for $n = 4000$ (hence $p = 44112$), requires about 90 min to obtain the value of one realization of the τ -coherence. In order to obtain more usable (i.e fast) codes in perspective of real-size applications, we are currently exploring HPC (High Performance Computing) strategies to compute correlation blocks using GPGPU (General-purpose Processing on Graphics Processing Units) computation. We are very confident into the use of GPU to reduce simulation's time.

5 Discussion

In this paper, we consider the regime $\log(p_n) = o(n^{\frac{1}{3}})$ and broaden the framework of [CJ11a]. Indeed the convergence of the τ -coherence to the Gumbel law still holds when the correlation matrix presents a supplementary an additional band of entries ϵ_n sufficiently small. This additional band can be way much larger than τ , i.e. as large as $K(n) = O(p_n^\nu)$ for $\nu > 0$ and the entries $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$ with speed $O(\sqrt{\log(p)/n})$. The speed of convergence of ϵ_n to 0 is optimal and not surprisingly it fits the threshold of bandwidth tapering in [BL08a].

Numerous further theoretical questions arise. Among them, the question of ultra-high regimes, i.e. what is the behaviour of the τ -coherence if $\log(p_n)$ is of order n^α with $\alpha > 1/3$? Another natural generalization would be to consider the model with ϵ_n disseminated on the whole matrix and not only bandwise. The Chen-Stein method should be adapted then. Besides, Shao and co in [LLS08] improved the speed of convergence to the Gumbel law by choosing an intermediate

variable. They obtained a speed of $O((\log n)^{5/2}/\sqrt{n})$. To our knowledge this approach is still open for the τ -coherence.

In this work, we have shown that the ability of our test to detect the alternative hypothesis is linked to the value of the parameter h . As in [FJ19], the asymptotic power of this test is 1 or α according to the value h is smaller or greater than 1. It means that asymptotically, our test is unable to detect the alternative when $h \geq 1$ so that the limiting distribution is the same under the null and under the alternative. In this work, we focus on an alternative that can be seen as a perturbation of the null (with the same structure in asymptotic). Even if these kinds of alternatives are classic, the phenomenon happening here is not and it leads to the question of the power function : is there any renormalisation of the τ -coherence that could highlight a continuous power function? To the best of our knowledge, it is still an open question.

Our result also suggest perspectives for potential applications, and numerical considerations. Indeed, considering numerical study, optimization of computation of the τ -coherence, or even more, of Monte-Carlo simulations in an experimental framework can become necessary. In this paper we have avoided the issue of the matrix dimension in term of computer memory by splitting it in sub-matrices. One can think to use HPC (High Performance Computing) methods, such as parallel computing or GPGPU, in order to reduce the necessary time of computation which becomes untractable when very large matrices are involved.

6 Proof of the main result

In this section, we describe the proof of our main result. First, we would like to highlight the fact that, as we said, we apply the Chen-Stein method. But, we do not apply it directly to the τ -coherence. It is more efficient to use the Chen-Stein method to a new easier to handle random variable. First of all, we introduce many notation which will be used along these proofs. Moreover the very technical part are postponed to an Appendix.

6.1 Notations

- $I = \{(k, j) \in \llbracket 1, p \rrbracket^2 : 1 \leq k < j \leq p\}$
- $I_\tau = \{(k, j) \in I : |k - j| < \tau\}$
- $I_K = \{(k, j) \in I : \tau \leq |k - j| \leq \tau + K\}$
- $I_0 = \{(k, j) \in I : |k - j| > \tau + K\}$
- $E_\delta = \{(k, j) \in I : k \in \Gamma_{p, \delta} \text{ or } j \in \Gamma_{p, \delta}\}$
- $\Lambda_p^\tau = \{(k, j) \in I : |k - j| < \tau \text{ and } \max_{1 \leq k \neq q, j \neq q \leq p} (|r_{kq}|, |r_{jq}|) \leq 1 - \delta\}$
- $\Lambda_p^K = \{(k, j) \in I : \tau \leq |k - j| \leq \tau + K \text{ and } \max_{1 \leq k \neq q, j \neq q \leq p} (|r_{kq}|, |r_{jq}|) \leq 1 - \delta\}$
- $\Lambda_p^0 = \{(k, j) \in I : |k - j| > \tau + K \text{ and } \max_{1 \leq k \neq q, j \neq q \leq p} (|r_{kq}|, |r_{jq}|) \leq 1 - \delta\}$

With those definitions, E_δ corresponds to the set of lines or columns where there is at least one coefficient r_{ij} which magnitude is greater than the level $1 - \delta$; Λ_p^τ , Λ_p^K , Λ_p^0 are respectively the

set I_τ , I_K and I_0 without the lines or columns with r_{ij} 's coefficients greater than the level $1 - \delta$ (which means that r_{ij} 's are bounded in those three sets. With these different sets, we can write three different partitions of the set I :

1. $I = I_\tau \cup I_K \cup I_0$
2. $I = E_\delta \cup \Lambda_p^\tau \cup \Lambda_p^K \cup \Lambda_p^0$
3. $I_0 \cup I_K = \Lambda_p^K \cup \Lambda_p^0 \cup [E_\delta \cap \overline{I_\tau}]$

The following Lemma gives the sizes of these three sets:

Lemma 6.1. *With the previous notations*

$$|I_\tau| = (\tau - 1) \left(\frac{2p - \tau}{2} \right), \quad (14)$$

$$|I_K| = (K + 1) \left(\frac{2p - K - 2\tau}{2} \right), \quad (15)$$

$$|I_0| = \frac{(p - \tau - K - 1)(p - \tau - K)}{2}. \quad (16)$$

6.2 Auxiliary variables

Now we introduce an auxiliary random variable which will be more convenient to handle in the Chen–Stein method. Let

$$V_{n,\tau} = \max_{1 \leq k < j \leq p, |k-j| \geq \tau} |{}^t\mathbb{X}_n^k \mathbb{X}_n^j| = \max_{\alpha=(k,j) \in I_0 \cup I_K} \left| {}^t\mathbb{X}_n^k \mathbb{X}_n^j \right|. \quad (17)$$

In the sequel, we will use notation $\alpha = (k, j)$ to denote index into different sets.

Proposition 6.1. *Under the assumptions of Theorem 2.1, we have*

$$\frac{n^2 L_{n,\tau}^2 - V_{n,\tau}^2}{n} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0. \quad (18)$$

The proof is analogous to the proof of convergence (46) in [CJ11a] using Lemma 2.2 of [Jia04]

Hence to study the asymptotic behaviour of $L_{n,\tau}$, it is enough to study the limiting distribution of $V_{n,\tau}$. To do so, we use another slightly different random variable defined by:

$$V'_{n,\tau} = \max_{\alpha \in \Lambda_p^0 \cup \Lambda_p^K} (Z_\alpha) \quad (19)$$

where the index $\alpha = (k, j)$ and $Z_\alpha = Z_{kj} = \left| {}^t\mathbb{X}_n^k \mathbb{X}_n^j \right|$. The two variables $V_{n,\tau}$ and $V'_{n,\tau}$ are linked by the following inequalities:

Proposition 6.2. *Let*

$$a_n(y) = \sqrt{4n \log(p_n) - n \log \log(p_n) + ny} \text{ with } y \in \mathbb{R}. \quad (20)$$

We have :

$$\mathbb{P}(V'_{n,\tau} > a_n(y)) \leq \mathbb{P}(V_{n,\tau} > a_n(y)) \leq \mathbb{P}(V'_{n,\tau} > a_n(y)) + o(1) \quad (21)$$

Proof. (For seek of simplicity, we will denote $a_n(y)$ by a_n in the sequel).

To prove this result, we need the two following technical results whose proofs are postponed to Section 6.

Lemma 6.2. *Let a_n be as in formula (20). Then,*

$$\mathbb{P}_0 := \mathbb{P}(|{}^t\mathbb{X}_n^1 \mathbb{X}_n^{\tau+K+2}| > a_n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{p_n^2} (1 + o(1)) = O_{n \rightarrow +\infty} \left(\frac{1}{p_n^2} \right). \quad (22)$$

Lemma 6.3. *Let a_n be as in formula (20) and let us define $c_\gamma := \frac{1}{2}\gamma^2 - 2|\gamma| + 2$ with γ defined in Assumption 1. Then, for any $d \in [0; c_\gamma[$ and $n \rightarrow +\infty$:*

$$\mathbb{P}_K := \mathbb{P}(|{}^t\mathbb{X}_n^1 \mathbb{X}_n^{\tau+1}| > a_n) = o(p_n^{-d}). \quad (23)$$

According to the partition $I_0 \cup I_K = \Lambda_p^0 \cup \Lambda_p^K \cup (E_\delta \cap \overline{I_\tau})$,

$$\begin{aligned} \mathbb{P}(V_{n,\tau} > a_n) &= \mathbb{P}\left(\max_{\alpha=(k,j) \in I_0 \cup I_K} |{}^t\mathbb{X}_n^k \mathbb{X}_n^j| > a_n\right) \\ &\leq \mathbb{P}(V'_{n,\tau} > a_n) + \mathbb{P}\left(\max_{\alpha=(k,j) \in E_\delta \cap \overline{I_\tau}} |{}^t\mathbb{X}_n^k \mathbb{X}_n^j| > a_n\right) \\ &\leq \mathbb{P}(V'_{n,\tau} > a_n) + \sum_{\alpha=(k,j) \in E_\delta \cap \overline{I_\tau}} \mathbb{P}\left(|{}^t\mathbb{X}_n^k \mathbb{X}_n^j| > a_n\right) \\ &\leq \mathbb{P}(V'_{n,\tau} > a_n) + \sum_{\alpha \in [E_\delta \cap \overline{I_\tau}] \cap I_K} \mathbb{P}(Z_\alpha > a_n) + \sum_{\alpha \in [E_\delta \cap \overline{I_\tau}] \cap I_0} \mathbb{P}(Z_\alpha > a_n). \end{aligned}$$

All variables Z_α having same distributions in the different sets above, we have

$$\begin{aligned} \mathbb{P}(V_{n,\tau} > a_n) &\leq \mathbb{P}(V'_{n,\tau} > a_n) + |[E_\delta \cap \overline{I_\tau}] \cap I_K| \mathbb{P}(Z_{1,\tau+1} > a_n) + |[E_\delta \cap \overline{I_\tau}] \cap I_0| \mathbb{P}(Z_{1,\tau+K+2} > a_n) \\ &\leq \mathbb{P}(V'_{n,\tau} > a_n) + |I_K| \mathbb{P}_K + |E_\delta| \mathbb{P}_0. \end{aligned}$$

We can use the following straightforward result :

Lemma 6.4.

$$|E_\delta| \leq 2p_n |\Gamma_{p,\delta}|$$

Hence from assumption 3 of theorem 2.1, we have :

$$|E_\delta| = o(p_n^2) \quad (24)$$

Now, we need to prove that $|I_K| \mathbb{P}_K + |E_\delta| \mathbb{P}_0 \xrightarrow{n \rightarrow +\infty} 0$. First, from lemma 6.2 and (24),

$$\begin{aligned} |E_\delta| \mathbb{P}(Z_{1,\tau+K+2} > a_n) &\underset{n \rightarrow +\infty}{\sim} |E_\delta| \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{p_n^2} \\ &\underset{n \rightarrow +\infty}{=} o(p_n^2) \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{p_n^2} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} o(1) \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

Secondly, using lemma 6.1 (more precisely Equation (15)) and lemma 6.3 we have :

$$|I_K| \mathbb{P}_K \xrightarrow{n \rightarrow +\infty} 0 \Leftrightarrow \nu < c_\gamma - 1 \quad (25)$$

$$(26)$$

and this is fulfilled from assumptions on theorem 1. Finally, we obtain :

$$|I_K| \mathbb{P}_K + |E_\delta| \mathbb{P}_0 \xrightarrow{n \rightarrow +\infty} 0. \quad (27)$$

Then,

$$\mathbb{P}(V_{n,\tau} > a_n) \leq \mathbb{P}(V'_{n,\tau} > a_n) + o(1). \quad (28)$$

Also, it is easy to see that :

$$\mathbb{P}(V'_{n,\tau} > a_n) \leq \mathbb{P}(V_{n,\tau} > a_n). \quad (29)$$

□

Remark 2. The main constraint so far is $p_n K \mathbb{P}_K \rightarrow 0$ when $n \rightarrow \infty$ which leads to

$$\nu < \frac{1}{2}\gamma^2 - 2|\gamma| + 1.$$

Moreover, it also implies the following condition:

$$\gamma \in [-2, 2] \text{ is such that } \frac{1}{2}\gamma^2 - 2|\gamma| + 1 > 0 \iff \gamma \in]2 + \sqrt{2}; 2 - \sqrt{2}[$$

6.3 Chen–Stein method for $V'_{n,\tau}$

We focus now on the asymptotic behaviour of $V'_{n,\tau}$. For that purpose, we apply the Chen-Stein method. We recall here this result, which can be found in [AGG89].

Lemma 6.5. *The Chen-Stein Method*

Let Λ be a set of indices. Let $\alpha \in \Lambda$ and B_α a set of subset of Λ (i.e. for all α , $B_\alpha \subset \Lambda$). Let η_α be random variables. For a given $a \in \mathbb{R}$, we define $\lambda := \sum_{\alpha \in \Lambda} \mathbb{P}(\eta_\alpha > a)$. Then,

$$\left| \mathbb{P}\left(\max_{\alpha \in \Lambda} (\eta_\alpha) \leq a\right) - e^{-\lambda} \right| \leq \min\left(1, \frac{1}{\lambda}\right) \cdot (b_1 + b_2 + b_3), \quad (30)$$

where

- $b_1 = \sum_{\alpha \in \Lambda} \sum_{\beta \in B_\alpha} \mathbb{P}(\eta_\alpha > a) \mathbb{P}(\eta_\beta > a)$
- $b_2 = \sum_{\alpha \in \Lambda} \sum_{\alpha \neq \beta \in B_\alpha} \mathbb{P}(\eta_\alpha > a, \eta_\beta > a)$
- $b_3 = \sum_{\alpha \in \Lambda} \mathbb{E}[|\mathbb{E}[\mathbf{1}_{\eta_\alpha > a} | \sigma(\eta_\beta, \beta \in \mathcal{I} \setminus B_\alpha)] - \mathbb{E}[\mathbf{1}_{\eta_\alpha > a}]|].$

As we said, this method is an approximation of weakly dependent events by a Poisson law which is represented by the quantity $e^{-\lambda}$ (corresponding to $\mathbb{P}(Z = 0)$, Z having a Poisson law $\mathcal{P}(\lambda)$). We need to find weakly dependent events to have b_1 , b_2 and b_3 small (even null or asymptotically null). In our case, notations are :

- $\Lambda = \Lambda_p^0 \cup \Lambda_p^K$.
- $\alpha = (k, j) \in \Lambda$.
- $B_\alpha = B_{kj} = \{(u, v) \in \Lambda : |k - u| < \tau + K, |j - v| < \tau + K \text{ and } (k, j) \neq (u, v)\}$.
- $\eta_\alpha = Z_\alpha = Z_{kj} = \left| \sum_{i=1}^n X_i^k X_i^j \right|$.
- $\lambda_n = \sum_{\alpha \in \Lambda} \mathbb{P}(Z_\alpha > a_n)$.
- $b_{1,n} = \sum_{\alpha \in \Lambda} \sum_{\beta \in B_\alpha} \mathbb{P}(Z_\alpha > a_n) \mathbb{P}(Z_\beta > a_n)$.
- $b_{2,n} = \sum_{\alpha \in \Lambda} \sum_{\alpha \neq \beta \in B_\alpha} \mathbb{P}(Z_\alpha > a_n, Z_\beta > a_n)$.
- $b_{3,n} = \sum_{\alpha \in \Lambda} \mathbb{E} [|\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} | \sigma(Z_\beta, \beta \in \Lambda \setminus B_\alpha)] - \mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}]|]$.

The following Lemmas present the asymptotics of $\lambda_n, b_{i,n}, i = 1$ to 3 .

Lemma 6.6. *Considering the previous notations, with straightforward computations we obtain the following results :*

- $|\Lambda_p^0| \sim p_n^2/2$ as $n \rightarrow +\infty$
- $|B_{ij}| \leq 8(\tau + K)p_n \sim 8Kp_n$ as $n \rightarrow +\infty$
- $|\Lambda_p^K| \leq |I_K|$

Lemma 6.7. *With the previous notations we have*

$$\lim_{n \rightarrow +\infty} (\lambda_n) = \frac{1}{\sqrt{8\pi}} e^{y/2}. \quad (31)$$

Proof: According to the Chen-Stein method and using the fact that random variables have the same law when indices are in the same set, we have

$$\lambda_n = \sum_{\alpha \in \Lambda_p^0 \cup \Lambda_p^K} \mathbb{P}(Z_\alpha > a_n) = \sum_{\alpha \in \Lambda_p^0} \mathbb{P}(Z_\alpha > a_n) + \sum_{\alpha \in \Lambda_p^K} \mathbb{P}(Z_\alpha > a_n) = |\Lambda_p^0| \cdot \mathbb{P}_0 + |\Lambda_p^K| \cdot \mathbb{P}_K$$

From theorem 1, lemma 6.3 and lemma 6.6, we have

$$\lim_{n \rightarrow +\infty} |\Lambda_p^K| \cdot \mathbb{P}_K = 0$$

while, according to lemma 6.2 and lemma 6.6,

$$|\Lambda_p^0| \cdot \mathbb{P}_0 \underset{+\infty}{\sim} p_n^2 \frac{1}{p_n^2} \frac{1}{\sqrt{8\pi}} e^{y/2} = \frac{1}{\sqrt{8\pi}} e^{y/2}.$$

Hence we obtain the desired result. We remark that this asymptotic will give the distribution function of the asymptotic Gumbel random variable.

Lemma 6.8. *In the previous framework, we have*

$$\lim_{n \rightarrow \infty} b_{i,n} = 0, \text{ for } i = 1, 2, 3.$$

The proof of this Lemma is postponed to the Appendix.

6.4 Proof of Theorem 2.1

We showed in Section 6.3 that

$$\mathbb{P}(V'_{n,\tau} \leq a_n) = \exp(-\lambda_n) + o(1) \text{ as } n \rightarrow +\infty \quad (32)$$

Thanks to lemma 6.2, we have, for n large enough :

$$\mathbb{P}(V_{n,\tau} \leq a_n) = \exp(-\lambda_n) + o(1) \text{ as } n \rightarrow +\infty \quad (33)$$

From the expressions of a_n and λ_n , this leads us to the asymptotic behaviour :

$$\frac{1}{n} V_{n,\tau}^2 - 4 \log(p_n) + \log \log(p_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Z \quad (34)$$

where Z has the Gumbel cdf defined in theorem 2.1. Then, we can write :

$$\frac{1}{n \log(p_n)} V_{n,\tau}^2 - 4 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0 \quad (35)$$

Then, from Proposition 6.1 we have (5). □

6.5 Proof of Theorem theorem 2.2

In order to prove this theorem, we first notice that according to the definition of $L_{n,\tau}$,

$$|\rho_{ij}| \leq L_{n,\tau} \text{ for any } (i, j) \in I_0 \cup I_K$$

Therefore to prove the convergence (8), it suffices to prove that it exists $(i, j) \in I_0 \cup I_K$ such that

$$\frac{n}{\log(p)} \rho_{ij}^2 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} +\infty$$

Then we consider $(i, j) \in I_K$ and ρ_{ij} the empirical correlation coefficient from the bivariate Gaussian distribution $\mathcal{N}_2 \left(0, \begin{pmatrix} 1 & \varepsilon_n \\ \varepsilon_n & 1 \end{pmatrix} \right)$. The density f_{ij} of ρ_{ij} is given by Muirhead [Mui82]: for any $x \in]-1, 1[$,

$$f_{ij}(x) = \frac{(n-2)\Gamma(n-1)}{\sqrt{2\pi}\Gamma(n-\frac{1}{2})} (1-\varepsilon_n^2)^{\frac{n-1}{2}} (1-x^2)^{\frac{n-4}{2}} (1-x\varepsilon_n)^{\frac{3}{2}-n} {}_2F_1 \left(\frac{1}{2}, \frac{1}{2}; n + \frac{1}{2}; \frac{1+x\varepsilon_n}{2} \right)$$

It leads to the density of $\sqrt{\frac{n}{\log(p)}} \rho_{ij}$, denoted by f_{ij}^* and defined for any $x \in]-\sqrt{\frac{n}{\log(p)}}, \sqrt{\frac{n}{\log(p)}}[$

$$f_{ij}^*(x) = C_n \left(1 - \frac{\log(p)}{n} x^2 \right)^{\frac{n-4}{2}} (1 - x\varepsilon_n \sqrt{\frac{\log(p)}{n}})^{\frac{3}{2}-n} {}_2F_1 \left(\frac{1}{2}, \frac{1}{2}; n + \frac{1}{2}; \frac{1 + x \sqrt{\frac{\log(p)}{n}} \varepsilon_n}{2} \right)$$

where $C_n := \sqrt{\frac{\log(p)}{n} \frac{(n-2)\Gamma(n-1)}{\sqrt{2\pi}\Gamma(n-\frac{1}{2})}} (1 - \varepsilon_n^2)^{\frac{n-1}{2}}$ is an explicit constant which does not depend on x . It remains to show that, for any $\delta > 0$:

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{n}{\log(p)} \rho_{ij}^2 < \delta\right) = \lim_{n \rightarrow +\infty} \int_{-\sqrt{\delta}}^{\sqrt{\delta}} f_{ij}^*(x) dx = 0$$

From a straightforward bound on ${}_2F_1$ we get

$$\begin{aligned} C_n(1 + |\varepsilon_n| \sqrt{\delta} \sqrt{\frac{\log(p)}{n}})^{3/2-n} \int_{-\sqrt{\delta}}^{\sqrt{\delta}} \left(1 - x^2 \frac{\log(p)}{n}\right)^{\frac{n-4}{2}} dx &\leq \mathbb{P}\left(\frac{n}{\log(p)} \rho_{ij}^2 < \delta\right) \\ &\leq C_n \sqrt{2} \left(1 + |\varepsilon_n| \sqrt{\delta} \sqrt{\frac{\log(p)}{n}}\right)^{1-n} \int_{-\sqrt{\delta}}^{\sqrt{\delta}} \left(1 - x^2 \frac{\log(p)}{n}\right)^{\frac{n-4}{2}} dx \end{aligned}$$

Moreover from Sirtling equivalent, we show that, as $n \rightarrow +\infty$

$$C_n \sim \frac{1}{\sqrt{2\pi n}} (1 - \varepsilon_n^2)^{\frac{n-1}{2}}$$

which leads to the desired limit. \square

Acknowledgments: the authors wish to thank Laurent Delsol for many fruitful discussions.

References

- [AGG89] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. Ann. Probab., 17(1):9–25, 1989.
- [AGZ10] G.W. Anderson, A. Guionnet, and O. Zeitouni. An introduction to random matrices. In Cambridge Studies in Advanced Mathematics, volume 118. Cambridge University Press, Cambridge, 2010.
- [And91] Donald W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica, 59(3):817–858, 1991.
- [BC12] Charles Bordenave and Djalil Chafaï. Around the circular law. Probab. Surv., 9:1–89, 2012.
- [BG16] Cristina Butucea and Ghislaine Gayraud. Sharp detection of smooth signals in a high-dimensional sparse matrix with indirect observations. Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, 52(4):1564 – 1591, 2016.
- [BJYZ09] Zhidong Bai, Dandan Jiang, Jian-Feng Yao, and Shurong Zheng. Corrections to LRT on large-dimensional covariance matrix by RMT. Ann. Stat., 37(6B):3822–3840, 2009.
- [BL08a] Peter J. Bickel and Elizaveta Levina. Covariance regularization by thresholding. Ann. Stat., 36(6):2577–2604, 2008.

- [BL08b] Peter J. Bickel and Elizaveta Levina. Regularized estimation of large covariance matrices. Ann. Stat., 36(1):199–227, 2008.
- [BRT09] Peter J. Bickel, Ya’acov Ritov, and Alexandre B. Tsybakov. Simultaneous analysis of Lasso and Dantzig selector. Ann. Stat., 37(4):1705–1732, 2009.
- [BS96] Zhidong Bai and Hewa Saranadasa. Effect of high dimension: by an example of a two sample problem. Stat. Sin., 6(2):311–329, 1996.
- [BS10] Zhidong Bai and Jack W. Silverstein. Spectral analysis of large dimensional random matrices. 2nd ed. Dordrecht: Springer, 2nd ed. edition, 2010.
- [CJ11a] T. Tony Cai and Tiefeng Jiang. Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. Ann. Statist., 39(3):1496–1525, 2011.
- [CJ11b] T. Tony Cai and Tiefeng Jiang. Supplement to "limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices". DOI:10.1214/11-AOS879SUPP, 2011.
- [CJ12] T. Tony Cai and Tiefeng Jiang. Phase transition in limiting distributions of coherence of high-dimensional random matrices. J. Multivariate Anal., 107:24–39, 2012.
- [CRT06a] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory, 52(2):489–509, 2006.
- [CRT06b] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math., 59(8):1207–1223, 2006.
- [CT05] Emmanuel J. Candès and Terence Tao. Decoding by linear programming. IEEE Trans. Inf. Theory, 51(12):4203–4215, 2005.
- [CT07] Emmanuel Candès and Terence Tao. The Dantzig selector: statistical estimation when p is much larger than n . (With discussions and rejoinder). Ann. Stat., 35(6):2313–2404, 2007.
- [CWX10a] T. Tony Cai, Lie Wang, and Guangwu Xu. Shifting inequality and recovery of sparse signals. IEEE Trans. Signal Process., 58(3):1300–1308, 2010.
- [CWX10b] Tony Tony Cai, Lie Wang, and Guangwu Xu. Stable recovery of sparse signals and an oracle inequality. IEEE Trans. Inf. Theory, 56(7):3516–3522, 2010.
- [CZ16] T. Tony Cai and Anru Zhang. Inference for high-dimensional differential correlation matrices. J. Multivariate Anal., 143:107–126, 2016.
- [CZZ10] T. Tony Cai, Cun-Hui Zhang, and Harrison H. Zhou. Optimal rates of convergence for covariance matrix estimation. Ann. Stat., 38(4):2118–2144, 2010.
- [DH01] David L. Donoho and Xiaoming Huo. Uncertainty principles and ideal atomic decomposition. IEEE Trans. Inf. Theory, 47(7):2845–2862, 2001.

- [Don00] David Donoho. High-dimensional data analysis: The curses and blessings of dimensionality. AMS Math Challenges Lecture, pages 1–32, 01 2000.
- [Don06] D.L. Donoho. Compressed sensing. IEEE Trans. Inf. Theory, 52:1289–1306, 2006.
- [EK18] Nouredine El Karoui. Random matrices and high-dimensional statistics: beyond covariance matrices. In Proceedings of the international congress of mathematicians 2018, ICM 2018, Rio de Janeiro, Brazil, August 1–9, 2018. Volume IV. Invited lectures, pages 2857–2876. Hackensack, NJ: World Scientific; Rio de Janeiro: Sociedade Brasileira de Matemática (SBM), 2018.
- [FJ19] Jianqing Fan and Tiefeng Jiang. Largest entries of sample correlation matrices from equi-correlated normal populations. Ann. Probab., 47(5):3321–3374, 2019.
- [For10] P.J. Forrester. Log-gases and random matrices. In London Mathematical Society Monographs Series, volume 34. Princeton University Press, Princeton, 2010.
- [Fuc04] J.-J. Fuchs. On sparse representations in arbitrary redundant bases. IEEE Trans. Inf. Theory, 50(6):1341–1344, 2004.
- [HMY21] Johannes Heiny, Thomas Mikosch, and Jorge Yslas. Point process convergence for the off-diagonal entries of sample covariance matrices. Ann. Appl. Probab., 31(2):538–560, 2021.
- [Jia04] Tiefeng Jiang. The asymptotic distributions of the largest entries of sample correlation matrices. Ann. Appl. Probab., 14(2):865–880, 2004.
- [Joh01] Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. Ann. Stat., 29(2):295–327, 2001.
- [Joh08] Iain M. Johnstone. Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy-Widom limits and rates of convergence. Ann. Stat., 36(6):2638–2716, 2008.
- [JT09] Iain M. Johnstone and D. Michael Titterton. Statistical challenges of high-dimensional data. Philos. Trans. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci., 367(1906):4237–4253, 2009.
- [LB95] Antonio V. Ligerde and Bryan W. Brown. Band covariance matrix estimation using restricted residuals: A Monte Carlo analysis. Int. Econ. Rev., 36(3):751–767, 1995.
- [LLR10] Deli Li, Wei-Dong Liu, and Andrew Rosalsky. Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. Probab. Theory Relat. Fields, 148(1-2):5–35, 2010.
- [LLS08] Wei-Dong Liu, Zhengyan Lin, and Qi-Man Shao. The asymptotic distribution and Berry-Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. Ann. Appl. Probab., 18(6):2337–2366, 2008.
- [LQR12] Deli Li, Yongcheng Qi, and Andrew Rosalsky. On Jiang’s asymptotic distribution of the largest entry of a sample correlation matrix. J. Multivariate Anal., 111:256–270, 2012.

- [LR06] Deli Li and Andrew Rosalsky. Some strong limit theorems for the largest entries of sample correlation matrices. Ann. Appl. Probab., 16(1):423–447, 2006.
- [Meh04] M.L. Mehta. Random matrices. In Pure and Applied Mathematics (Amsterdam), volume 142. Elsevier/Academic Press, Amsterdam, 2004.
- [Mui82] Robb J. Muirhead. Aspects of multivariate statistical theory. John Wiley & Sons, Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.
- [Pec12] Giovanni Peccati. The Chen-Stein method for Poisson functionals, <https://hal.archives-ouvertes.fr/hal-00654235>. April 2012.
- [R C23] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2023.
- [SZ14] Qi-Man Shao and Wen-Xin Zhou. Necessary and sufficient conditions for the asymptotic distributions of coherence of ultra-high dimensional random matrices. Ann. Probab., 42(2):623–648, 03 2014.
- [WD14] Y. Wang and M. J. Daniels. Computationally efficient banding of large covariance matrices for ordered data and connections to banding the inverse Cholesky factor. J. Multivariate Anal., 130:21–26, 2014.
- [Wig58] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. Ann. Math. (2), 67:325–327, 1958.
- [Zho07] Wang Zhou. Asymptotic distribution of the largest off-diagonal entry of correlation matrices. Trans. Am. Math. Soc., 359(11):5345–5363, 2007.

7 Appendix: Proofs of technical results

7.1 Proof of Lemma 6.2

For the proof of this Lemma, we need the following technical result which is presented in [CJ11a] (see Lemma 6.8) and proved in the supplementary paper.

Lemma 7.1. *We consider the following hypotheses :*

1. ξ_1, \dots, ξ_n i.i.d random variables such that $\mathbb{E}[\xi_1] = 0$ and $\mathbb{E}[\xi_1^2] = 1$.
2. $\exists t_0 > 0, \exists \alpha \in]0, 1]$ such that $\mathbb{E}[e^{t_0|\xi_1|^\alpha}] < +\infty$.
3. $(p_n)_{n \in \mathbb{N}^*}$ such that $p_n \xrightarrow{n \rightarrow +\infty} +\infty$ and $\log(p_n) = o\left(n^{\frac{\alpha}{2+\alpha}}\right)$ as $n \rightarrow +\infty$
4. $(y_n)_{n \geq 1}$ such that $y_n \xrightarrow{n \rightarrow +\infty} y > 0$

Then,

$$\mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} \sum_{k=1}^n \xi_k \geq y_n\right) \underset{n \rightarrow +\infty}{\sim} \frac{1}{y\sqrt{2\pi}} p_n^{-\frac{1}{2}y_n^2} \sqrt{\log(p_n)}^{-1} \quad (36)$$

Let us check all the hypotheses of this lemma above. First we write :

$$\mathbb{P}(|X^1 X^{\tau+K+2}| > a_n) = \mathbb{P}\left(\left|\sum_{i=1}^n X_i^1 X_i^{\tau+K+2}\right| > a_n\right) \quad (37)$$

Now, if we define $\xi_i = X_i^1 X_i^{\tau+K+2}$, we have :

1. $\mathbb{E}[\xi_i] \stackrel{\parallel}{=} \mathbb{E}[X_i^1] \mathbb{E}[X_i^{\tau+K+2}] = 0 \times 0 = 0$ where the independence come from the sample.
2. $\mathbb{E}[\xi_i^2] \stackrel{\parallel}{=} \mathbb{E}[(X_i^1)^2] \mathbb{E}[(X_i^{\tau+K+2})^2] = 1 \times 1 = 1$
3. For $t_0 = \frac{1}{2}$ and $\alpha = 1$, we have :

$$\mathbb{E}[e^{t_0 |\xi_1|^\alpha}] = \mathbb{E}[e^{\frac{|X_1^1 X_1^{\tau+K+2}|}{2}}] \leq \mathbb{E}[e^{\frac{1}{2}(X_1^1)^2}] \mathbb{E}[e^{\frac{1}{2}(X_1^{\tau+K+2})^2}] < +\infty$$

4. We have $w_n := \frac{a_n}{\sqrt{n \log(p)}} \xrightarrow{n \rightarrow +\infty} \sqrt{4} = 2 > 0$

5. According to the hypothesis 1 from theorem 2.1 : $\log(p_n) = o(n^{\frac{1}{3}})$ as $n \xrightarrow{n \rightarrow +\infty} +\infty$

So we have all hypothesis needed to apply the lemma 7.1, and then :

$$\begin{aligned} \mathbb{P}(|X^1 X^{\tau+K+2}| > a_n) &= \mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} |X^1 X^{\tau+K+2}| > \frac{a_n}{\sqrt{n \log(p_n)}}\right) \\ &= \mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} \left|\sum_{i=1}^n X_i^1 X_i^{\tau+K+2}\right| > \frac{a_n}{\sqrt{n \log(p_n)}}\right) \\ &= \mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} \sum_{i=1}^n X_i^1 X_i^{\tau+K+2} > \frac{a_n}{\sqrt{n \log(p_n)}}\right) + \mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} \sum_{i=1}^n X_i^1 X_i^{\tau+K+2} < -\frac{a_n}{\sqrt{n \log(p_n)}}\right) \\ &= \mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} \sum_{i=1}^n X_i^1 X_i^{\tau+K+2} > w_n\right) + \mathbb{P}\left(-\frac{1}{\sqrt{n \log(p_n)}} \sum_{i=1}^n X_i^1 X_i^{\tau+K+2} > w_n\right) \quad (38) \end{aligned}$$

From lemma 7.1

$$\begin{aligned} \mathbb{P}(|X^1 X^{\tau+K+2}| > a_n) &= \frac{1}{2\sqrt{2\pi}} p_n^{-\frac{1}{2}w_n^2} \frac{1}{\sqrt{\log(p_n)}} (1 + o(1)) + \frac{1}{2\sqrt{2\pi}} p_n^{-\frac{1}{2}w_n^2} \frac{1}{\sqrt{\log(p_n)}} (1 + o(1)) \\ &= \frac{1}{\sqrt{2\pi}} p_n^{-\frac{1}{2}w_n^2} \frac{1}{\sqrt{\log(p_n)}} (1 + o(1)) \\ &= \frac{1}{2\pi \log(p_n)} e^{-\frac{1}{2} \frac{a_n^2}{n \log(p_n)} - \frac{1}{2} \log \log(p_n)} (1 + o(1)) = \frac{1}{p_n^2 \sqrt{2\pi}} e^{-\frac{1}{2}y} (1 + o(1)) \end{aligned}$$

7.2 Proof of Lemma lemma 6.3

We remind that $\mathbb{P}_K := \mathbb{P}(|^t\mathbb{X}_n^1\mathbb{X}_n^{\tau+1}| > a_n)$. We will apply once again lemma 7.1, with new quantities ξ and ω :

- $\mathbb{P}_K^+ := \mathbb{P}({}^t\mathbb{X}_n^1\mathbb{X}_n^{\tau+1} > a_n)$
- $\mathbb{P}_K^- := \mathbb{P}({}^t\mathbb{X}_n^1\mathbb{X}_n^{\tau+1} < -a_n)$
- $\xi_k := X_k^1 X_k^{\tau+1}$
- $w_k := \frac{(\xi_k - \varepsilon_n)}{\sqrt{1 + \varepsilon_n^2}}$.

Notice that $(\xi_k)_{k \geq 1}$ are independent due to the independence between each line of \mathbf{X}_n . First we compute $\mathbb{E}[\xi_k] = \varepsilon_n$ and $\text{var}(\xi_k) = 1 + \varepsilon_n^2$. So, $\mathbb{E}[w_k] = 0$ and $\text{var}(w_k) = 1$. We will apply the lemma 7.1 with w_k . Then,

$$\mathbb{P}_K^+ = \mathbb{P}\left(\frac{1}{\sqrt{n \log(p_n)}} \sum_{k=1}^n w_k > \underbrace{\frac{a_n - n\varepsilon_n}{\sqrt{(1 + \varepsilon_n^2)n \log(p_n)}}}_{:= z_n}\right) \quad (39)$$

From hypotheses of theorem 2.1, we have $\lim_{n \rightarrow +\infty} [z_n] := z = 2 - \gamma > 0$. Then,

$$\begin{aligned} \mathbb{P}_K^+ &\sim \frac{1}{z\sqrt{2\pi}} p_n^{-\frac{1}{2}z_n^2} \sqrt{\log(p_n)}^{-1} \\ &\sim \frac{1}{z\sqrt{2\pi}} \exp\left[-\frac{1}{2}z_n^2 \log(p_n) - \frac{1}{2} \log \log(p_n)\right] \\ &\sim \frac{1}{z\sqrt{2\pi}} \exp\left[\frac{-2}{1 + \varepsilon_n^2} \log(p_n) \left(1 + \frac{\varepsilon_n^2}{4} \frac{\log \log(p_n)}{\log(p_n)} + \frac{y}{4} \frac{1}{\log(p_n)} + \frac{n\varepsilon_n^2}{4 \log(p_n)} - \frac{\varepsilon_n}{2} \frac{a_n}{\log(p_n)}\right)\right] \end{aligned}$$

With our hypotheses on ε_n , we have :

- $\frac{-2}{1 + \varepsilon_n^2} \log(p_n) \xrightarrow{n \rightarrow +\infty} -\infty$
- $\frac{\varepsilon_n^2 \log \log(p_n)}{4 \log(p_n)} \xrightarrow{n \rightarrow +\infty} 0$
- $\frac{y}{4} \frac{1}{\log(p_n)} \xrightarrow{n \rightarrow +\infty} 0$
- $\frac{n\varepsilon_n^2}{4 \log(p_n)} \xrightarrow{n \rightarrow +\infty} \frac{1}{4} \gamma^2$ from $\varepsilon_n \sim \gamma \sqrt{\frac{\log(p_n)}{n}}$.
- $\frac{\varepsilon_n a_n}{2 \log(p_n)} \xrightarrow{n \rightarrow +\infty} \gamma$ from $a_n \sim 2\sqrt{n \log(p_n)}$

$$p_n^a \mathbb{P}_K^+ \sim \frac{1}{z\sqrt{2\pi}} \exp \left[\left(a - 2 - \frac{1}{2}\gamma^2 + 2\gamma + o(1) \right) \log(p_n) \right] \quad (40)$$

Finally, for $\gamma \in]-2, 2[$:

$$\lim_{n \rightarrow +\infty} [p_n^a \mathbb{P}_K^+] = 0 \Leftrightarrow a < 2 + \frac{1}{2}\gamma^2 - 2\gamma \quad (41)$$

Analogously, we have :

$$\mathbb{P}_K^- = \mathbb{P} \left(\frac{1}{\sqrt{n \log(p_n)}} \sum_{k=1}^n w_k < -\frac{a_n + n\varepsilon_n}{\sqrt{(1 + \varepsilon_n^2)n \log(p_n)}} \right) \quad (42)$$

$$= \mathbb{P} \left(\frac{-1}{\sqrt{n \log(p_n)}} \sum_{k=1}^n w_k > \underbrace{\frac{a_n + n\varepsilon_n}{\sqrt{(1 + \varepsilon_n^2)n \log(p_n)}}}_{:= \tilde{z}_n} \right) \quad (43)$$

Thanks to lemma 7.1 and because $\lim_{n \rightarrow +\infty} [\tilde{z}_n] := \tilde{z} = 2 + \gamma > 0$, we have :

$$\mathbb{P}_K^- \sim \frac{1}{\tilde{z}\sqrt{2\pi}} \exp \left[\frac{-2}{1 + \varepsilon_n^2} \log(p_n) \left(1 + \frac{\varepsilon_n^2 \log \log(p_n)}{4 \log(p_n)} + \frac{y}{4} \frac{1}{\log(p_n)} + \frac{n\varepsilon_n^2}{4 \log(p_n)} + \frac{\varepsilon_n}{2} \frac{a_n}{\log(p_n)} \right) \right] \quad (44)$$

$$p_n^b \mathbb{P}_K^- \sim \frac{1}{\tilde{z}\sqrt{2\pi}} \exp \left[\left(b - 2 - \frac{1}{2}\gamma^2 - 2\gamma + o(1) \right) \log(p_n) \right] \quad (45)$$

And finally, for $\gamma \in]-2, 2[$:

$$\lim_{n \rightarrow +\infty} [p_n^b \mathbb{P}_K^-] = 0 \Leftrightarrow b < 2 + \frac{1}{2}\gamma^2 + 2\gamma \quad (46)$$

To conclude, observing that

$$\min \left(2 + \frac{1}{2}\gamma^2 + 2\gamma, 2 + \frac{1}{2}\gamma^2 - 2\gamma \right) = \frac{1}{2}\gamma^2 - 2|\gamma| + 2 := c_\gamma,$$

combining eq. (46) and eq. (41), we obtain, for all $d \in [0; c_\gamma[$ and as $n \rightarrow +\infty$:

$$\mathbb{P}_K := \mathbb{P} \left(|{}^t \mathbb{X}_n^1 \mathbb{X}_n^{\tau+1}| > a_n \right) = o \left(p_n^{-d} \right). \quad (47)$$

7.3 Proof of Lemma 6.8

7.3.1 Computation of $b_{1,n}$

We add some notations :

- $B_\alpha^0 := B_\alpha \cap \Lambda_p^0$ and $|B_\alpha^0| \leq |B_\alpha| \leq 8(\tau + K)p_n$

- $B_\alpha^K := B_\alpha \cap \Lambda_p^K$ and $|B_\alpha^K| \leq K^2$
- $\mathbb{P}_\alpha := \mathbb{P}(Z_\alpha > a_n)$

As used above, Z_{α_1} and Z_{α_2} will have the same law as long as α_1 and α_2 belong to the same set. Then, we have :

$$\begin{aligned}
b_{1,n} &= \sum_{\alpha \in \Lambda_p^0 \cup \Lambda_p^K} \sum_{\beta \in B_\alpha} \mathbb{P}_\alpha \mathbb{P}_\beta \\
&= \sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^0} \mathbb{P}_\alpha \mathbb{P}_\beta + \sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_\alpha \mathbb{P}_\beta + \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^0} \mathbb{P}_\alpha \mathbb{P}_\beta + \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^K} \mathbb{P}_\alpha \mathbb{P}_\beta \\
&= \sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^0} (\mathbb{P}_0)^2 + \sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_0 \mathbb{P}_K + \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^0} \mathbb{P}_K \mathbb{P}_0 + \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^K} (\mathbb{P}_K)^2 \\
&= |\Lambda_p^0| \cdot |B_\alpha^0| \cdot (\mathbb{P}_0)^2 + |\Lambda_p^0| \cdot |B_\alpha^K| \mathbb{P}_0 \mathbb{P}_K + |\Lambda_p^K| \cdot |B_\alpha^0| \mathbb{P}_K \mathbb{P}_0 + |\Lambda_p^K| \cdot |B_\alpha^K| \cdot (\mathbb{P}_K)^2.
\end{aligned}$$

At this point, we need to check that $\lim_{n \rightarrow +\infty} (b_{1,n}) = 0$, so we focus particularly on :

1. $|\Lambda_p^0| \cdot |B_\alpha^0| \cdot (\mathbb{P}_0)^2 :$

$$|\Lambda_p^0| \cdot |B_\alpha^0| \cdot (\mathbb{P}_0)^2 \sim \frac{1}{2} p_n^2 \cdot |B_\alpha^0| \cdot (\mathbb{P}_0)^2 \leq 4(\tau + K) p_n^3 \cdot O\left(\frac{1}{p_n^4}\right) = O(p_n^{\nu-1}) \quad (48)$$

From assumptions on ν we have $\lim_{n \rightarrow +\infty} [|\Lambda_p^0| \cdot |B_\alpha^0| \cdot (\mathbb{P}_0)^2] = 0$

2. $|\Lambda_p^0| \cdot |B_\alpha^K| \mathbb{P}_0 \mathbb{P}_K :$

$$|\Lambda_p^0| \cdot |B_\alpha^K| \mathbb{P}_0 \mathbb{P}_K \sim \frac{1}{2} p_n^2 |B_\alpha^K| \mathbb{P}_0 \mathbb{P}_K \leq \frac{1}{2} K^2 p_n^2 \mathbb{P}_0 \mathbb{P}_K \leq O(p_n^{2+2\nu}) O(p_n^{-2}) \mathbb{P}_K = O(p_n^{2\nu} \mathbb{P}_K)$$

According to lemma 6.3, we will have $\lim_{n \rightarrow +\infty} [p_n^{2\nu} \mathbb{P}_K] = 0$ iff $2\nu < c_\gamma$ which is true from hypothesis 5 in theorem 2.1. Then, we obtain :

$$\lim_{n \rightarrow +\infty} [|\Lambda_p^0| \cdot |B_\alpha^K| \mathbb{P}_0 \mathbb{P}_K] = 0$$

3. $|\Lambda_p^K| \cdot |B_\alpha^0| \mathbb{P}_K \mathbb{P}_0 :$ We use the same principle of computation than previously :

$$|\Lambda_p^K| \cdot |B_\alpha^0| \mathbb{P}_K \mathbb{P}_0 \leq p_n K |B_\alpha^0| \mathbb{P}_K \mathbb{P}_0 \leq 8 p_n K (\tau + K) p_n \mathbb{P}_K \mathbb{P}_0 = O(p_n^{2\nu} \mathbb{P}_K) \quad (49)$$

So, from previous assumptions on ν , we have :

$$\lim_{n \rightarrow +\infty} [|\Lambda_p^K| \cdot |B_\alpha^0| \mathbb{P}_K \mathbb{P}_0] = 0.$$

4. $|\Lambda_p^K| \cdot |B_\alpha^K| \cdot (\mathbb{P}_K)^2 :$ We have :

$$|\Lambda_p^K| \cdot |B_\alpha^K| \cdot (\mathbb{P}_K)^2 \leq p K^3 (\mathbb{P}_K)^2 = O(p_n^{1+3\nu}) (\mathbb{P}_K)^2 \quad (50)$$

According to lemma 6.3, if $1 + 3\nu < 2c_\gamma$, we have

$$\lim_{n \rightarrow +\infty} [|\Lambda_p^K| \cdot |B_\alpha^K| \cdot (\mathbb{P}_K)^2] = 0.$$

To conclude, we finally obtain :

$$\lim_{n \rightarrow +\infty} [b_{1,n}] = 0.$$

Remark 3. The main constraint here is $p_n K^3 (\mathbb{P}_K)^2 \rightarrow 0$ which is true from condition $p_n K \mathbb{P}_K \rightarrow 0$ of remark 2.

7.3.2 Computation of $b_{2,n}$

The computation of $b_{2,n}$ is the most technical part. As we did for the computation of $b_{1,n}$, we will divide this computation into four parts (according on which set we are). We remind the definition of $b_{2,n}$:

$$b_{2,n} = \sum_{\alpha \in \Lambda_p^0 \cup \Lambda_p^K} \sum_{\beta \in B_\alpha} \mathbb{P}(Z_\alpha > a_n, Z_\beta > a_n).$$

Here we introduce some new notations :

- $\mathbb{P}_{\alpha\beta} := \mathbb{P}(Z_\alpha > a_n, Z_\beta > a_n)$
- $\mathbb{P}_{0i} := \mathbb{P}_{\alpha\beta} \mathbf{1}_{\alpha \in \Lambda_p^0} \mathbf{1}_{\beta \in \Omega_i}$ where Ω_i will be a subset of indices and i an integer.
- $\mathbb{P}_{Ki} := \mathbb{P}_{\alpha\beta} \mathbf{1}_{\alpha \in \Lambda_p^K} \mathbf{1}_{\beta \in \Omega_i}$ where Ω_i will be a subset of indices and i an integer.

To show that $\lim_{n \rightarrow +\infty} [b_{2,n}] = 0$, we will divide it into four sums, each one being the sum of the same probability on a given set of indices. Then, we have :

$$b_{2,n} = \underbrace{\sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^0} \mathbb{P}_{\alpha\beta}}_{:=Q_1} + \underbrace{\sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta}}_{:=Q_2} + \underbrace{\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^0} \mathbb{P}_{\alpha\beta}}_{:=Q_3} + \underbrace{\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta}}_{:=Q_4} \quad (51)$$

Computation of Q_1 :

First, we define some additional subsets of indices. In particular, we have :

1. $\Omega_1 := \{(u, v) \in \Lambda_p^0 : i - u < \tau \text{ and } j - v < \tau\}$ and $|\Omega_1| \leq \tau^2$
2. $\Omega_2 := \{(u, v) \in \Lambda_p^0 : i - u < \tau \text{ and } \tau < j - v < \tau + K\}$ and $|\Omega_2| \leq \tau K$
3. $\Omega_3 := \{(u, v) \in \Lambda_p^0 : \tau < i - u < \tau + K \text{ and } j - v < \tau\}$ and $|\Omega_3| \leq \tau K$
4. $\Omega_4 := \{(u, v) \in \Lambda_p^0 : i - u < \tau \text{ and } \tau + K \leq j - v\}$ and $|\Omega_4| \leq \tau(p_n - \tau - K) \leq \tau p_n$
5. $\Omega_5 := \{(u, v) \in \Lambda_p^0 : \tau + K \leq i - u \text{ and } j - v < \tau\}$ and $|\Omega_5| \leq \tau(p_n - \tau - K) \leq \tau p_n$
6. $\Omega_6 := \{(u, v) \in \Lambda_p^0 : \tau < i - u < \tau + K \text{ and } \tau < j - v < \tau + K\}$ and $|\Omega_6| \leq K^2$
7. $\Omega_7 := \{(u, v) \in \Lambda_p^0 : \tau < i - u < \tau + K < \text{ and } \tau + K \leq j - v\}$
and $|\Omega_7| \leq K(p_n - \tau - K) \leq K p_n$
8. $\Omega_8 := \{(u, v) \in \Lambda_p^0 : \tau + K \leq i - u \text{ and } j - v < \tau\}$ and $|\Omega_8| \leq K(p_n - \tau - K) \leq K p_n$

We have :

$$Q_1 \leq 4 \sum_{i=1}^8 \sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in \Omega_i} \mathbb{P}_{\alpha\beta} \quad (52)$$

Then, using the fact that on each given subset the random variables have the same law :

$$Q_1 \leq |\Lambda_p^0| \cdot |\Omega_1| \mathbb{P}_{01} + |\Lambda_p^0| \cdot |\Omega_2| \mathbb{P}_{02} + |\Lambda_p^0| \cdot |\Omega_3| \mathbb{P}_{03} + |\Lambda_p^0| \cdot |\Omega_4| \mathbb{P}_{04} \quad (53)$$

$$+ |\Lambda_p^0| \cdot |\Omega_5| \mathbb{P}_{05} + |\Lambda_p^0| \cdot |\Omega_6| \mathbb{P}_{06} + |\Lambda_p^0| \cdot |\Omega_7| \mathbb{P}_{07} + |\Lambda_p^0| \cdot |\Omega_8| \mathbb{P}_{08} \quad (54)$$

So, we just have to show that each part will have a null limit when n is going to infinity.

Lemma 7.2. *Using the previous notations, we have, as $n \rightarrow +\infty$:*

$$|\Lambda_p^0| \cdot |\Omega_1| \mathbb{P}_{01} \rightarrow 0 \quad (55)$$

Proof. We have :

$$|\Lambda_p^0| \cdot |\Omega_1| \mathbb{P}_{01} \leq |\Lambda_p^0| \tau^2 \mathbb{P}_{01} \sim \frac{1}{2} p_n^2 \tau^2 \mathbb{P}_{01} = o(p_n^{2+2t} \mathbb{P}_{01}) \text{ for any } t > 0$$

where we use lemma 6.6 for the equivalent. We can write :

$$\mathbb{P}_{01} = \mathbb{P} \left(\left| \sum_{k=1}^n u_k^1 u_k^2 \right| > a_n, \left| \sum_{k=1}^n u_k^3 u_k^4 \right| > a_n \right) \quad (56)$$

$$(57)$$

where $(u_k^1, u_k^2, u_k^3, u_k^4)_{1 \leq k \leq n} \stackrel{i.i.d}{\sim} \mathcal{N}_4(0, \Sigma_4)$ and

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & r_1 & 0 \\ 0 & 1 & 0 & r_2 \\ r_1 & 0 & 1 & 0 \\ 0 & r_2 & 0 & 1 \end{pmatrix},$$

where coefficients r_1, r_2 are from the correlation matrix (r_{kj}) . From Lemma 6.11 of [CJ11a], focusing on equation (131), we know that

$$\mathbb{P}_{01} \leq O(p_n^{-2b^2+\varepsilon_1}) + O(p_n^{-2-2c^2+\varepsilon_2}) \text{ as } n \rightarrow +\infty \quad (58)$$

for any $\varepsilon_1, \varepsilon_2 > 0$ and where $a = \frac{1+(1-\delta)^2}{2}$, $b = \frac{a}{(1-\delta)^2}$ and $c = \frac{1-a}{3}$ for $\delta \in]0, 1[$. By construction $b^2 - 1 > 0$, hence for a well-chosen t such that $t < b^2 - 1$, there exists $\varepsilon_1(\delta) > 0$ such that we have :

$$\varepsilon_1 < 2b^2 - 2 - 2t. \quad (59)$$

Analogously, we can find $\varepsilon_2(\delta)$ such that:

$$\varepsilon_2 < 2(c^2 - t). \quad (60)$$

Since $\tau = o(p^t)$ for any $t > 0$, we have

$$|\Lambda_p^0| \cdot |\Omega_1| \mathbb{P}_{01} \rightarrow 0 \text{ as } n \rightarrow +\infty. \quad (61)$$

□

Lemma 7.3. *Using previous notations, we have :*

$$|\Lambda_p^0| \cdot |\Omega_2| \mathbb{P}_{02} \rightarrow 0 \quad (62)$$

Proof. We have :

$$|\Lambda_p^0| \cdot |\Omega_2| \mathbb{P}_{02} \leq \tau K |\Lambda_p^0| \mathbb{P}_{02} \sim \frac{1}{2} p_n^2 \tau K \mathbb{P}_{02} = O(\tau p_n^{2+\nu} \mathbb{P}_{02})$$

where we use the lemma 6.6 for the equivalence above. In this proof, we almost have the same case than in the proof of lemma 7.2. In fact, the only difference is the matrix Σ_4 which is now

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & r & 0 \\ 0 & 1 & 0 & \varepsilon_n \\ r & 0 & 1 & 0 \\ 0 & \varepsilon_n & 0 & 1 \end{pmatrix},$$

where r is a coefficient from the matrix (r_{kj}) . So, by the same method we have

$$p_n^{2+\nu} \mathbb{P}_{02} \rightarrow 0 \quad (63)$$

iff $\varepsilon_1 < 2b^2 - 2 - \nu$ and $\varepsilon_2 < 2c^2 - \nu$ where we still have $b = \frac{1+(1-\delta)^2}{2(1-\delta)^2}$ and $c = \frac{1-(1-\delta)^2}{6}$. Moreover we can show that $b^2 - 1 > c^2$. Then, if $\nu < 2c^2$ (fullfilled by assumptions in theorem 2.1), and from $\tau = o(p_n^t)$ for any $t > 0$, we have :

$$|\Lambda_p^0| \cdot |\Omega_2| \mathbb{P}_{02} \rightarrow 0. \quad (64)$$

□

Lemma 7.4. *Using notations previously introduced, we have :*

$$|\Lambda_p^0| \cdot |\Omega_3| \mathbb{P}_{03} \rightarrow 0 \quad (65)$$

Proof. This proof is exactly the same than for lemma 7.3 except that the matrix becomes

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & \varepsilon_n & 0 \\ 0 & 1 & 0 & r \\ \varepsilon_n & 0 & 1 & 0 \\ 0 & r & 0 & 1 \end{pmatrix}.$$

In particular, we obtain the same condition on ν .

□

Lemma 7.5. *Using notations previously introduced, we have :*

$$|\Lambda_p^0| \cdot |\Omega_4| \mathbb{P}_{04} \rightarrow 0 \quad (66)$$

Proof. We have :

$$|\Lambda_p^0| \cdot |\Omega_4| \mathbb{P}_{04} \leq \tau p_n |\Lambda_p^0| \mathbb{P}_{04} \sim \tau p_n^3 \mathbb{P}_{04} \quad (67)$$

Now, the correlation matrix is $\Sigma_4 = \begin{pmatrix} 1 & 0 & r & 0 \\ 0 & 1 & 0 & 0 \\ r & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Thanks to the lemma 6.9 in [CJ11a], proved in the supplementary paper, we obtain $\mathbb{P}_{04} = O(p_n^{-4+\varepsilon})$ for any $\varepsilon > 0$. Then, we have $p_n^3 \tau \mathbb{P}_{04} = O\left(\frac{\tau}{p_n^{1-\varepsilon}}\right)$ which tends to 0 as $n \rightarrow \infty$ since $\tau = o(p_n^t)$ for any $t > 0$. \square

Lemma 7.6. *Using notations previously introduced, we have :*

$$|\Lambda_p^0| \cdot |\Omega_5| \mathbb{P}_{05} \rightarrow 0. \quad (68)$$

Proof. This proof is exactly the same than for lemma 7.5 considering the correlation matrix

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & r \\ 0 & 0 & 1 & 0 \\ 0 & r & 0 & 1 \end{pmatrix}$$

\square

Lemma 7.7. *Using notations previously introduced, we have :*

$$|\Lambda_p^0| \cdot |\Omega_6| \mathbb{P}_{06} \rightarrow 0 \quad (69)$$

Proof. This proof is exactly the same than for lemma 7.3 except that the matrix become

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & \varepsilon_n & 0 \\ 0 & 1 & 0 & \varepsilon_n \\ \varepsilon_n & 0 & 1 & 0 \\ 0 & \varepsilon_n & 0 & 1 \end{pmatrix}$$

In particular, we have :

$$|\Lambda_p^0| \cdot |\Omega_6| \mathbb{P}_{06} \leq K^2 |\Lambda_p^0| \mathbb{P}_{06} \sim O(p_n^{2+2\nu} \mathbb{P}_{06}) \quad (70)$$

with Σ_4 as correlation matrix for the 4-uplet in \mathbb{P}_{06} . As for lemma 7.3, we have the following conditions

$$\varepsilon_1 < 2b^2 - 2 - 2\nu \text{ and } \varepsilon_2 < 2(c^2 - \nu) \quad (71)$$

which is summarized in $\nu < c^2$, and which is true considering theorem 2.1. Then, we obtain the desired result :

$$|\Lambda_p^0| \cdot |\Omega_6| \mathbb{P}_{06} \rightarrow 0 \quad (72)$$

\square

Lemma 7.8. *Using notations previously introduced, we have :*

$$|\Lambda_p^0| \cdot |\Omega_7| \mathbb{P}_{07} \rightarrow 0 \quad (73)$$

Proof. This proof is exactly the same than for lemma 7.5 considering the correlation matrix

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & \varepsilon_n & 0 \\ 0 & 1 & 0 & 0 \\ \varepsilon_n & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

□

Lemma 7.9. *Using notations previously introduced, we have :*

$$|\Lambda_p^0| \cdot |\Omega_8| \mathbb{P}_{08} \rightarrow 0 \quad (74)$$

Proof. This proof is exactly the same than for lemma 7.5 considering the correlation matrix

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \varepsilon_n \\ 0 & 0 & 1 & 0 \\ 0 & \varepsilon_n & 0 & 1 \end{pmatrix}.$$

□

Remark 4. *The main constraint here for Q_1 is $\nu < c^2$.*

Computation of Q_2 :

For this case, we will divide the computation into two parts. Indeed, we consider two cases : when α is close to the set Λ_p^K and when it is not. For that purpose, we introduce the following sets :

$$I_{0,I} = \{(i, j) \in \llbracket 1, p \rrbracket, i < j \text{ and } \tau + K < j - i < \tau + 4K\} \text{ and } \Lambda_{p,I}^0 = \Lambda_p^0 \cap I_{0,I}$$

and

$$I_{0,II} = I_0 \setminus I_{0,I} \text{ and } \Lambda_{p,II}^0 = \Lambda_p^0 \cap I_{0,II}.$$

We can write :

$$Q_2 := \sum_{\alpha \in \Lambda_p^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta} = \sum_{\alpha \in \Lambda_{p,I}^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta} + \sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta}.$$

Now, we look at the sum on $\Lambda_{p,I}^0$. We notice that on this set, the probability $\mathbb{P}_{\alpha\beta}$ is issued from a Gaussian vector with correlation matrix

$$\Sigma_4 = \begin{pmatrix} 1 & 0 & r_1 & r_2 \\ 0 & 1 & r_3 & r_4 \\ r_1 & r_3 & 1 & \varepsilon_n \\ r_2 & r_4 & \varepsilon_n & 1 \end{pmatrix}$$

where $|r_i| \leq 1 - \delta$ for all $i \in \{1, 2, 3, 4\}$. Moreover, coefficients $(r_i)_i$ may be replaced here by ε_n according to the position of the indice in both sets $\Lambda_{p,I}^0$ and B_α^K . But we know that $\lim_{n \rightarrow +\infty} (\varepsilon_n) = 0$ then, for n large enough, we still have $|r_i| \leq 1 - \delta$. Using Cauchy-Schwarz inequality, we can bound :

$$\mathbb{P}_{\alpha\beta} = \mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} \mathbf{1}_{Z_\beta > a_n}] \leq \sqrt{\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}^2] \mathbb{E} [\mathbf{1}_{Z_\beta > a_n}^2]} \leq \sqrt{\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}] \mathbb{E} [\mathbf{1}_{Z_\beta > a_n}]} = \sqrt{\mathbb{P}_\alpha \mathbb{P}_\beta}$$

Now, we use the fact that $\alpha \in \Lambda_{p,I}^0 \subset \Lambda_p^0$ and $\beta \in B_\alpha^K \subset I_K$, then :

$$\mathbb{P}_{\alpha\beta} \leq \sqrt{\mathbb{P}_0 \mathbb{P}_K}$$

which is true for any $|r_i| \leq 1$ then, $\sup_{|r_i| \leq 1, i=1, \dots, 4} \mathbb{P}_{\alpha\beta} \leq \sqrt{\mathbb{P}_0 \mathbb{P}_K}$. At this point, using $|\Lambda_{p,I}^0| \leq 3Kp$, $|B_\alpha^K| \leq K^2$, and $\mathbb{P}_0 = O(p^{-2})$, we get:

$$\sum_{\alpha \in \Lambda_{p,I}^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta} \leq 3K^3 p \sqrt{\mathbb{P}_0 \mathbb{P}_K} \leq 3K^3 \mathbb{P}_K^{1/2} O(1) = O(p^{3\nu} \mathbb{P}_K^{1/2}) \text{ as } n \rightarrow +\infty.$$

Now, using lemma 6.3:

$$p^{3\nu} \mathbb{P}_K^{1/2} \xrightarrow{n \rightarrow +\infty} 0 \Leftrightarrow 3\nu < \frac{1}{2} \left(\frac{1}{2} \gamma^2 - 2|\gamma| + 2 \right) \Leftrightarrow \nu < \frac{1}{6} c_\gamma \quad (75)$$

which is true according to assumptions of theorem 2.1.

Now, let us focus on the computation of $\Lambda_{p,II}^0$. For that purpose, we introduce four subsets:

- $\Omega_1^2 := \{(u, v) \in B_\alpha^K : u - i < \tau \text{ and } j - v > \tau + K\} \text{ and } |\Omega_1^2| \leq K\tau$
- $\Omega_2^2 := \{(u, v) \in B_\alpha^K : \tau + K < u - i \text{ and } j - v < \tau\} \text{ and } |\Omega_2^2| \leq K\tau$
- $\Omega_3^2 := \{(u, v) \in B_\alpha^K : \tau \leq u - i \leq \tau + K \text{ and } j - v > \tau + K\} \text{ and } |\Omega_3^2| \leq K^2$
- $\Omega_4^2 := \{(u, v) \in B_\alpha^K : \tau + K < u - i \text{ and } \tau \leq j - v \leq \tau + K\} \text{ and } |\Omega_4^2| \leq K^2$

We have :

$$\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta} \leq 4 \sum_{i=1}^4 \sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_i^2} \mathbb{P}_{\alpha\beta}$$

In order to consider all these subset, we have the four next lemmas :

Lemma 7.10. *Considering the same notations as previously : $(u_k^1, u_k^2, u_k^3, u_k^4)_{1 \leq k \leq n} \stackrel{i.i.d}{\sim} \mathcal{N}_4(0, \Sigma_4)$.*

If the probability $\mathbb{P}_{\alpha\beta}$ is issued from a Gaussian vector with covariance matrix $\Sigma_4 = \begin{pmatrix} 1 & 0 & r_1 & x \\ 0 & 1 & 0 & 0 \\ r_1 & 0 & 1 & \varepsilon_n \\ x & 0 & \varepsilon_n & 1 \end{pmatrix}$

where $x \in \{\varepsilon_n, 0\}$, then :

$$\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_1^2} \mathbb{P}_{\alpha\beta} \leq O\left(p^{t+\nu+\epsilon} \mathbb{P}_K^{1/2}\right) \quad (76)$$

for any $t > 0$ and any $\epsilon > 0$.

Proof. In order to prove this result, we observe that in this case, for all $k \geq 1$, u_k^2 is independent of $\{u_k^1, u_k^3, u_k^4\}$. It means that conditionally on u_k^1 , we have independence between Z_{12} and Z_{34} . By consequence, using Cauchy-Schwarz, we obtain :

$$\mathbb{P}_{\alpha\beta} = \mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Z_{12} > a_n} \mathbf{1}_{Z_{34} > a_n} | u_k^1, k = 1, \dots, n \right] \right] \quad (77)$$

$$= \mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Z_{12} > a_n} | u_k^1, k = 1, \dots, n \right] \mathbb{E} \left[\mathbf{1}_{Z_{34} > a_n} | u_k^1, k = 1, \dots, n \right] \right] \quad (78)$$

$$\leq \sqrt{\mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Z_{12} > a_n} | u_k^1, k = 1, \dots, n \right]^2 \right] \mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Z_{34} > a_n} | u_k^1, k = 1, \dots, n \right]^2 \right]} \quad (79)$$

Now, because u_k^1 is independent of u_k^2 , we can use lemma 6.7 from [CJ11a] and we have :

$$\mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Z_{12} > a_n} | u_k^1, k = 1, \dots, n \right]^2 \right] = O\left(p^{-4+\epsilon}\right)$$

for any $\epsilon > 0$. And on the other side, we have :

$$\mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Z_{34} > a_n} | u_k^1, k = 1, \dots, n \right]^2 \right] \leq \mathbb{P}_K$$

Finally, using $|\Lambda_{p,II}^0| \leq p^2$ and $|\Omega_1^2| \leq K\tau$, and writing $K = O(p^\nu)$, we have the desired result:

$$\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_1^2} \mathbb{P}_{\alpha\beta} \leq O\left(p^{t+\nu+\epsilon} \mathbb{P}_K^{1/2}\right) \quad (80)$$

□

Now, from lemma 7.10 we have the condition :

$$t + \nu + \epsilon < \frac{1}{2} \left(\frac{1}{2} \gamma^2 - 2|\gamma| + 2 \right),$$

which can be fulfilled from condition in eq. (75) and for well-chosen $t > 0$ and $\epsilon > 0$. Finally we obtain, with our condition on ν that :

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_1^2} \mathbb{P}_{\alpha\beta} \right] = 0.$$

For the other subsets Ω_i^2 , $i = 2, 3, 4$, we will use the same method. Indeed we notice that respectively for Ω_2^2 , Ω_3^2 and Ω_4^2 , the covariance matrices involved are respectively :

$$\Sigma_4^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & r \\ 0 & x & 1 & \varepsilon_n \\ 0 & r & \varepsilon_n & 1 \end{pmatrix}, \quad \Sigma_4^3 = \begin{pmatrix} 1 & 0 & \varepsilon_n & x \\ 0 & 1 & 0 & 0 \\ \varepsilon_n & 0 & 1 & \varepsilon_n \\ x & 0 & \varepsilon_n & 1 \end{pmatrix}, \quad \Sigma_4^4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & \varepsilon_n \\ 0 & x & 1 & \varepsilon_n \\ 0 & \varepsilon_n & \varepsilon_n & 1 \end{pmatrix} \quad (81)$$

For each case, we use the fact that we always have u_k^2 (or u_k^1) independent of the other three random variables. Also, in order to use the lemma 7.10, we notice that by construction $Z_{12} = Z_{21}$. Then, for cases Ω_3^2 and Ω_4^2 , we just have to consider the Gaussian vector $(u_k^2, u_k^1, u_k^3, u_k^4)$ instead of $(u_k^1, u_k^2, u_k^3, u_k^4)$. In that way, all matrices have the same form than in lemma 7.10 and similar upper-bound for the probability. Now, we use the upper-bound of subsets Ω_i^2 . More precisely, for Ω_3^2 and Ω_4^2 :

$$\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_3^2} \mathbb{P}_{\alpha\beta} \leq |\Lambda_{p,II}^0| \cdot |\Omega_3^2| \mathbb{P}_K^{1/2} O(p^{-2+\epsilon}) = O(p^{2\nu+\epsilon} \mathbb{P}_K^{1/2}) \text{ for any } \epsilon > 0$$

It means that we need to have, according to lemma 6.3 :

$$2\nu < \frac{1}{2} \left(\frac{1}{2} \gamma^2 - 2|\gamma| + 2 \right) \Leftrightarrow \nu < \frac{1}{4} \left(\frac{1}{2} \gamma^2 - 2|\gamma| + 2 \right). \quad (82)$$

With this condition

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_3^2} \mathbb{P}_{\alpha\beta} \right] = 0$$

and

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in \Omega_4^2} \mathbb{P}_{\alpha\beta} \right] = 0$$

Finally, the case Ω_2^2 leads to the exactly same result than for Ω_1^2 because of the upper-bound on $|\Omega_2^2|$ which is the same than for $|\Omega_1^2|$. Then, we have :

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_{p,II}^0} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta} \right] = 0$$

and then

$$\lim_{n \rightarrow +\infty} [Q_2] = 0$$

Remark 5. The main constraint here is $\nu < \frac{1}{6} c_\gamma$.

Computation of Q_3 :

We focus here on $Q_3 = \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^0} \mathbb{P}_{\alpha\beta}$. Once more, we will consider different subsets for β according to its place into B_α^0 . More precisely, let us define :

- $\Omega_1^3 = \{(u, v) \in \Lambda_p^0 : i - u < \tau \text{ and } v - j < \tau\} \text{ and } |\Omega_1^3| \leq \tau^2.$
- $\Omega_2^3 = \{(u, v) \in \Lambda_p^0 : i - u < \tau \text{ and } \tau \leq v - j \leq \tau + K\} \text{ and } |\Omega_2^3| \leq \tau K.$
- $\Omega_3^3 = \{(u, v) \in \Lambda_p^0 : i - u < \tau \text{ and } \tau + K < v - j\} \text{ and } |\Omega_3^3| \leq \tau p.$

- $\Omega_4^3 = \{(u, v) \in \Lambda_p^0 : \tau \leq i - u \leq \tau + K \text{ and } v - j < \tau\} \text{ and } |\Omega_4^3| \leq \tau K.$
- $\Omega_5^3 = \{(u, v) \in \Lambda_p^0 : \tau + K < i - u \text{ and } v - j < \tau\} \text{ and } |\Omega_5^3| \leq \tau p.$
- $\Omega_6^3 = \{(u, v) \in \Lambda_p^0 : \tau \leq i - u \leq \tau + K \text{ and } \tau \leq v - j \leq \tau + K\} \text{ and } |\Omega_6^3| \leq K^2.$
- $\Omega_7^3 = \{(u, v) \in \Lambda_p^0 : \tau \leq i - u \leq \tau + K \text{ and } \tau + K < v - j\} \text{ and } |\Omega_7^3| \leq Kp.$
- $\Omega_8^3 = \{(u, v) \in \Lambda_p^0 : \tau + K < i - u \text{ and } \tau \leq v - j \leq \tau + K\} \text{ and } |\Omega_8^3| \leq Kp.$

Then, we have :

$$Q_3 \leq \sum_{i=1}^8 \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_i^3} \mathbb{P}_{\alpha\beta}$$

For Q_3 , we use the computation of Q_2 . Indeed, covariance matrices which are involved in the computation of Q_3 are similar than for Q_2 . The similarity comes from the fact that we exchange the role between α and β . More precisely, for Q_2 we had $\alpha \in \Lambda_p^0$ and $\beta \in B_\alpha^K$ and now we have $\alpha \in \Lambda_p^K$ and $\beta \in B_\alpha^0$. So, covariance matrices here will have the same structure than in Q_2 exchanging columns $\{1, 2\}$ and columns $\{3, 4\}$.

We notice that :

$$|\Omega_1^3|, |\Omega_2^3|, |\Omega_4^3| \leq |\Omega_6^3| \leq K^2 \quad (83)$$

Using the fact that $B_\alpha^0 \subset \Lambda_p^0$, we have :

$$\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_6^3} \mathbb{P}_{\alpha\beta} \leq \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_6^3} \mathbb{P}_0 \leq pK \cdot K^2 \mathbb{P}_0 = O(p^{3\nu-1}) \text{ as } n \rightarrow +\infty$$

However, from condition in eq. (75), we have $\nu < \frac{1}{3} (\frac{1}{4}\gamma^2 - |\gamma| + 1)$. But, $\gamma \in]-2+\sqrt{2}, 2-\sqrt{2}[$. Then, $\frac{1}{3} (\frac{1}{4}\gamma^2 - |\gamma| + 1) \in]\frac{1}{6}, \frac{1}{3}[$. It leads, in particular, to $\nu < \frac{1}{3}$ and then :

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_6^3} \mathbb{P}_{\alpha\beta} \right] = 0$$

And so, from equation eq. (83) :

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_1^3} \mathbb{P}_{\alpha\beta} \right] = \lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_2^3} \mathbb{P}_{\alpha\beta} \right] = \lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_4^3} \mathbb{P}_{\alpha\beta} \right] = 0 \quad (84)$$

Now, we look at the case when β belongs to $\Omega_3^3, \Omega_5^3, \Omega_7^3$ and Ω_8^3 . We start by describing each covariance matrix involved. We note $x \in \{\varepsilon_n, 0\}$ and r a correlation coefficient such that $|r| \leq 1 - \delta$. We have :

- for $\alpha \in \Lambda_p^K, \beta \in \Omega_3^3, \Sigma_4^3 = \begin{pmatrix} 1 & \varepsilon_n & r & 0 \\ \varepsilon_n & 1 & x & 0 \\ r & x & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

- for $\alpha \in \Lambda_p^K, \beta \in \Omega_5^3, \Sigma_4^3 = \begin{pmatrix} 1 & \varepsilon_n & 0 & x \\ \varepsilon_n & 1 & 0 & r \\ 0 & 0 & 1 & 0 \\ x & r & 0 & 1 \end{pmatrix}$
- for $\alpha \in \Lambda_p^K, \beta \in \Omega_7^3, \Sigma_4^3 = \begin{pmatrix} 1 & \varepsilon_n & \varepsilon_n & 0 \\ \varepsilon_n & 1 & x & 0 \\ \varepsilon_n & x & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
- for $\alpha \in \Lambda_p^K, \beta \in \Omega_8^3, \Sigma_4^3 = \begin{pmatrix} 1 & \varepsilon_n & 0 & x \\ \varepsilon_n & 1 & 0 & \varepsilon_n \\ 0 & 0 & 1 & 0 \\ x & \varepsilon_n & 0 & 1 \end{pmatrix}.$

We observe that for each case above one variable u_k^3 or u_k^4 is independent from the three other ones. Then we can use the same method than for Q_2 (conditioning on u_k^3 when u_k^4 is independent of the other ones or reversely on u_k^4). Then, by Cauch-Schwarz, we obtain the same upper-bound for $\mathbb{P}_{\alpha\beta}$. To show that we obtain the desired convergence, we study here the worst case. That is to say, using the fact that $|\Omega_i^3| \leq Kp$ for $i = 3, 5, 7, 8$, we can write :

$$\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_7^3} \mathbb{P}_{\alpha\beta} \leq |\Lambda_p^K| \cdot |\Omega_7^3| \mathbb{P}_K^{1/2} O(p^{-2+\epsilon}) \leq O(p^{2\nu+\epsilon} \mathbb{P}_K) \text{ as } n \rightarrow +\infty$$

Then, we have exactly the same condition on ν that for equation eq. (82). It means that

$$\lim_{n \rightarrow +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_7^3} \mathbb{P}_{\alpha\beta} \right] = 0,$$

and it induces that

$$\lim_{n \rightarrow +\infty} [Q_3] = 0.$$

Computation of Q_4 :

This last quantity is simpler because we can write :

$$Q_4 = \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^K} \mathbb{P}_{\alpha\beta} \leq \sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in B_\alpha^K} \mathbb{P}_K = |\Lambda_p^K| \cdot |B_\alpha^K| \mathbb{P}_K \leq pK^3 \mathbb{P}_K = O(p^{1+3\nu} \mathbb{P}_K) \text{ as } n \rightarrow +\infty$$

So, to have $\lim_{n \rightarrow +\infty} (Q_4) = 0$ we must have, according to lemma 6.3:

$$1 + 3\nu < \frac{1}{2}\gamma^2 - 2|\gamma| + 2 \Leftrightarrow \nu < \frac{1}{3}(c_\gamma - 1)$$

Under this assumption on ν we have

$$\lim_{n \rightarrow +\infty} [Q_4] = 0$$

Finally, gathering the results from Q_1 to Q_4 and under sufficient assumptions, we have :

$$\lim_{n \rightarrow +\infty} [b_{2,n}] = 0.$$

Remark 6. *The constraint on ν is $\nu < \frac{1}{3}(c_\gamma - 1)$.*

7.3.3 Computation of $b_{3,n}$

We have :

$$\begin{aligned} b_{3,n} &= \sum_{\alpha \in \Lambda_p^0 \cup \Lambda_p^K} \mathbb{E} [|\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} | \sigma(Z_\beta, \beta \in (\Lambda_p^0 \cup \Lambda_p^K) \setminus B_\alpha)] - \mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}]]|] \\ &= \sum_{\alpha \in \Lambda_p^0} \mathbb{E} [|\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} | \sigma(Z_\beta, \beta \in (\Lambda_p^0 \cup \Lambda_p^K) \setminus B_\alpha)] - \mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}]]|] \\ &\quad + \sum_{\alpha \in \Lambda_p^K} \mathbb{E} [|\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} | \sigma(Z_\beta, \beta \in (\Lambda_p^0 \cup \Lambda_p^K) \setminus B_\alpha)] - \mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}]]|] \end{aligned} \tag{85}$$

The first term of the RHS above is 0 from the choice of B_α . Hence

$$b_{3,n} = \sum_{\alpha \in \Lambda_p^K} \mathbb{E} [|\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} | \sigma(Z_\beta, \beta \in (\Lambda_p^0 \cup \Lambda_p^K) \setminus B_\alpha)] - \mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}]]|] \tag{86}$$

$$\leq \sum_{\alpha \in \Lambda_p^K} \mathbb{E} [\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n} | \sigma(Z_\beta, \beta \in (\Lambda_p^0 \cup \Lambda_p^K) \setminus B_\alpha)]] + \mathbb{E} [\mathbb{E} [\mathbf{1}_{Z_\alpha > a_n}]] \tag{87}$$

$$\leq 2 |\Lambda_p^K| \mathbb{P}_K \tag{88}$$

$$\tag{89}$$

According to the hypotheses in theorem 1, we have

$$\lim_{n \rightarrow +\infty} [|\Lambda_p^K| \mathbb{P}_K] = 0$$

Finally, we obtain :

$$\lim_{n \rightarrow +\infty} [b_{3,n}] = 0$$

Remark 7. *Gathering remark 2 to 6 and noticing that on $\gamma \in]-2 + \sqrt{2}, 2 - \sqrt{2}[$ we have $\frac{1}{3}(c_\gamma - 1) < \frac{1}{6}c_\gamma$ we get the final assumptions of theorem 1 for ν .*

□