Coherence of high-dimensional random matrices in a Gaussian case : application of the Chen-Stein method

M Boucher, D Chauveau, M Zani

Coherence of high-dimensional random matrices in a Gaussian case : application of the Chen-Stein method

M. Boucher [*] , D. Chauveau [*], M. Zani [*†]

Abstract :
This paper studies the τ-coherence of a $(n \times p)$-observation matrix in a Gaussian framework. The τ-coherence is defined as the largest magnitude outside a diagonal bandwith of size τ of the empirical correlation coefficients associated to our observations. Using the Chen-Stein method we derive the limiting law of the normalized coherence and show the convergence towards a Gumbel distribution. We generalize here the results of Cai and Jiang [CJ11a]. We assume that the covariance matrix of the model is bandwise. Moreover, we provide numerical considerations highlighting issues from the high dimension hypotheses. We numerically illustrate the asymptotic behaviour of the coherence with Monte-Carlo experiment using a HPC splitting strategy for high dimensional correlation matrices.

Key words : coherence, high-dimensional matrices, correlation, Chen-Stein method, Gaussian, GPGPU, random matrices, sparsity.

1 Introduction

Random matrix theory has known a huge amount of breakthroughs for these last twenty years. Developments have been made in theoretical fields as well as in various applied domains. Among these applications, one can cite high-energy physics (e.g. [For10] on log–gases), electronic 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 matrices theory, see [BS10, Meh04, AGZ10].

In statistics more particularly, random matrices are useful for inference in a high dimensional framework. One can think about 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, the dimension p is much bigger than the sample size n.

We will be focusing here on the covariance structure of a certain type of random matrices. More precisely, we will be examining the coherence of random matrices with bandwise covariance of size $\tau > 1$. Let us define the model.

Let $\left(X^1, X^2, \ldots, X^p\right)$ be a p – dimensional Gaussian random vector with mean $\mu = {}^t(\mu^1, \cdots, \mu^p)$ in $\mathbb{R}^p$ and covariance matrix $\Sigma = (\sigma_{kj})_{kj} \in \mathbb{R}^p \times \mathbb{R}^p$. For $n \in \mathbb{N}^*$, we consider a random sample $\left(X_i^1, X_i^2, \ldots, X_i^p\right)_{i=1,\ldots,n}$ issued from $\left(X^1, X^2, \ldots, X^p\right)$, arranged in a (n, p)-matrix $\mathbb{X}$. From a statistical point of view, it means that each row can be seen as an individual and each column as a character. We will write $\mathbb{X}^k$ the k^{th} column of $\mathbb{X}$. We are interested in the

[*]Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, Route de Chartres, B.P. 6759, 45067, Orléans cedex 2, France
[†]Corresponding author: marguerite.zani@univ-orleans.fr

correlation terms of $\mathbb{X}$ in models where p is much larger than n. The classical empirical Pearson's correlation coefficient is defined by :

$$\rho_{kj} = \frac{\sum_{i=1}^{n} \left(X_i^k - \overline{\mathbb{X}^k}\right)\left(X_i^j - \overline{\mathbb{X}^j}\right)}{\sqrt{\sum_{i=1}^{n} \left(X_i^k - \overline{\mathbb{X}^k}\right)^2}\sqrt{\sum_{i=1}^{n} \left(X_i^j - \overline{\mathbb{X}^j}\right)^2}} = \frac{< \mathbb{X}^k - \overline{\mathbb{X}^k}\mathbf{1}_n, \mathbb{X}^j - \overline{\mathbb{X}^j}\mathbf{1}_n >}{\parallel \mathbb{X}^k - \overline{\mathbb{X}^k}\mathbf{1}_n \parallel . \parallel \mathbb{X}^j - \overline{\mathbb{X}^j}\mathbf{1}_n \parallel}, \tag{1}$$

where $\mathbf{1}_n$ is the identical 1–vector in $\mathbb{R}^n$:

$$\mathbf{1}_n =^t (1, 1, \ldots, 1) \in \mathbb{R}^n$$

and $\parallel x \parallel$ stands for the the Euclidian norm of the vector x and $\overline{\mathbb{X}^k}$ is the empirical mean of the k^{th} column $\mathbb{X}^k$:

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

Equivalently if the mean $\mu =^t \left(\mu^1, \ldots, \mu^p\right)$ is known,

$$\tilde{\rho}_{kj} = \frac{\sum_{i=1}^{n} \left(X_i^k - \mu^k\right)\left(X_i^j - \mu^j\right)}{\sqrt{\sum_{i=1}^{n} \left(X_i^k - \mu^k\right)^2}\sqrt{\sum_{i=1}^{n} \left(X_i^j - \mu^j\right)^2}} = \frac{< \mathbb{X}^k - \mu^k\mathbf{1}_n, \mathbb{X}^j - \mu^j\mathbf{1}_n >}{\parallel \mathbb{X}^k - \mu^k\mathbf{1}_n \parallel . \parallel \mathbb{X}^j - \mu^j\mathbf{1}_n \parallel} \tag{2}$$

The empirical correlation coefficients ρ_{kj} (resp. $\tilde{\rho}_{kj}$) are arranged in a (p,p)-matrix R_n (resp. $\tilde{R}_n$) which is the empirical correlation matrix of $\mathbb{X}$.

Definition 1.1. *With the notations above, we can define the largest magnitude of the off–diagonal terms of R_n and $\tilde{R}_n$:*

$$L_n = \max_{1 \le k < j \le p} |\rho_{kj}|, \quad \tilde{L}_n = \max_{1 \le k < j \le p} |\tilde{\rho}_{kj}| \tag{3}$$

The quantity $\tilde{L}_n$ is defined as the coherence of the matrix $\mathbb{X}_n$. With a slight abuse of terminology, we will call both L_n and $\tilde{L}_n$ coherence of $\mathbb{X}_n$.

The notion of coherence has first appeared in signal theory as an indicator of the sparsity of a matrix. More precisely, it is involved in the so-called Mutual Incoherence Property (MIP), which can be explained as follows: a mesurement (n,p) matrix X is used to recover a k–sparse signal β via linear mesurements $y = X\beta$ using a recovery algorithm. The condition

$$(2k-1)\tilde{L}_n < 1$$

ensures the exact recovery of β when β has at most k non zero entries. For details on this approach, see Donoho and Huo [DH01], Fuchs [Fuc04], Cai, Wang and Xu [CWX10b], and references therein.

Another domain where covariance and correlation matrices are highly used is statistic theory, for example testing $\Sigma = I$ against $\Sigma \neq I$. This issue has been considered in the case where n and p are of same order (i.e. $n/p \to \gamma \in (0, \infty)$) by Johnstone in the Gaussian case [Joh01], and Péché in the sub-Gaussian case [Pé09]. The test statistic relies – according to PCA methods – on

the largest eigenvalue of the empirical covariance matrix $\lambda_{max}(\hat{\Sigma}_n)$. The asymptotic distribution of this maximum eigenvalue is the Tracy–Widom law.

However, testing $\Sigma = I$ can seem too restrictive, and one think about independance versus non independence in terms of correlation matrix i.e. testing $R_n = I$ against $R_n \neq I$. According to previous results, one could think about using $\lambda_{max}(R_n)$. However, even if Tracy–Widom law is conjectured in this case (see [Jia04b], and see also [HM18] for a study on the i.i.d. case), one choose to study instead the coherence as a test statistic. Jiang in [Jia04a] 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 [Jia04a] 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]. 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 dimensional context.

Lately Cai and Jiang [CJ11a] (see also the supplement [CJ11b]) considered "ultra-high dimensions" i.e. p as large as e^{n^β}. In this paper, they also present a variant of the coherence, the so–called τ–coherence aimed to test whether the covariance Σ has a given bandwidth $\tau > 1$, where $\tau = 1$ would be a special case. We define it below:

Definition 1.2. *For any integer $\tau \geq 1$, we define the τ–coherence as:*

$$L_{n,\tau} = \max_{|k-j|\geqslant\tau} |\rho_{kj}| \tag{4}$$

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 normal hypothesis, improving assumptions on the moments of the entries and on the dimension p.

Our purpose is to generalize this model: we assume that $\Sigma = (\sigma_{kj})_{1\leqslant k,j\leqslant p}$ is defined as follows:

$$\sigma_{kj} = \begin{cases} \gamma_{kj}\sigma_k\sigma_j & \text{if} \quad |k-j| < \tau \\ \epsilon_n\sigma_k\sigma_j & \text{if} \quad \tau \leq |k-j| \leq \tau + K \\ 0 & \text{if} \quad \tau + K < |k-j| \end{cases} . \tag{5}$$

So, Σ is divided into three parts : a central band of size $\tau \in \mathbb{N}$, an outside part with null coefficients and a transitional bandwidth of size $K \in \mathbb{N}$. We have, for all $k \in [\![1; p]\!]$, $\sigma_k > 0$; for all $k, j \in [\![1; p]\!]$, $\gamma_{kj} \in [-1, 1]$ and $(\epsilon_n)_{n\geqslant 1}$ is a sequence of real numbers in $[-1, 1]$ such that $\lim_{n\to+\infty} |\epsilon_n| = 0$. To be more precise, the construction of Σ suggests that if we take two p-vectors which are close (in term of indexes, for example X^1 and X^2), they will be correlated. If they are far enough one to another, they will be independent. But, if they are not so close and not so far, their correlation will decrease to zero when n goes to infinity. This generalization of the model of [CJ11a] seemed to us more suited for real datasets. Later, we will see that both τ and K may depend of n and may go to infinity under sufficient hypotheses. Without loss of generality, we can assume that $\mu = 0_{\mathbb{R}^p}$ and for all $k \in [\![1, p]\!]$, $\sigma_k = 1$.

All previously cited studies are highly related to the Chen-Stein method which we also use here. It relies on a Poisson approximation of weakly dependent events. For references on this method, see [AGG89], [Pec12] and references therein.

This paper is organized as follows: section 2 presents the main results, further on section 3 gives some simulation results for our model, section 4 is devoted to the proof of the main result whereas section 5 gather 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 asymptoticaly when $n \to \infty$.

2 Main result

We focus on correlations not too big, i.e. not too close to 1 or -1. Hence we define the following set:

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

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

where $(r_{kj})_{p\times p}$ is the correlation matrix issued from $\Sigma = (\sigma_{kj})_{p\times p}$.

The main result of the paper is the following Theorem:

Theorem 2.1. *Let n be an integer, $p = p_n$ a sequence such that $p_n \underset{n\to+\infty}{\longrightarrow} +\infty$. 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* $n \to +\infty$

Hyp 2 : $\tau = \tau(n) = o(p_n^t)$ *as* $n \to +\infty$ *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* $n \to +\infty$ *and* $\gamma \in]-2+\sqrt{2}, 2-\sqrt{2}[$

Hyp 5 : $K = K(n) = O\left(p_n^{\nu}\right)$ *where* $\nu \in]0, c(\gamma,\delta)[$

and $c(\gamma,\delta) = min\left(\frac{1}{3}(\frac{1}{2}\gamma^2 - 2|\gamma| + 1), \frac{\delta^2(2-\delta)^2}{36}\right)$.
Under these conditions, we can show that :

$$nL_{n,\tau}^2 - 4\log(p_n) + \log(\log(p_n)) \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} Z \tag{6}$$

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

We can observe that it is the same distribution as in [CJ11a]. The band in the covariance matrix which contains terms ε_n is a smooth transition between the central band and the external part of the matrix with null terms. This model is an illustration of vanishing dependence when the components are too far one from each other. In [CJ11a], the central band of the matrix is as large as τ with the condition of theorem 2.1. If we suppose $K = cst < +\infty$, we boil down to the same model. Indeed, we have $\tilde{\tau} = \tau + K$ wich is the new width of the non–null bandwidth, and $\tilde{\tau}$ is still such that :

$$\forall t > 0, \, \tilde{\tau} = o(p^t) \text{ as } n \to +\infty. \tag{7}$$

3 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 C20]. 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 $p = \left[\exp\left(n^{\frac{1}{3.5}}\right)\right]$ 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).

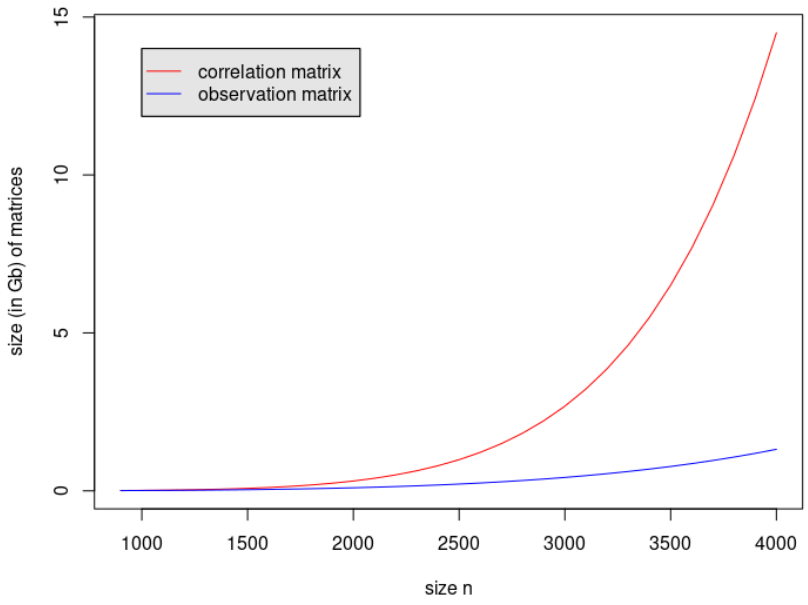

Figure 1: Size of $(p \times p)$-correlation matrix and $(n \times p)$-observation matrix according to n with $p = \left[\exp(n^{1/3.5})\right]$ for real numbers stored as double precision numbers.

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 choosen by the user. 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 wether the block corresponds to the central band (with bandwith τ) or not.

Using this strategy, we can generate correlation matrices even if p is very large, 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 :

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

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 the matrices of observations and computing the correlations per blocks. For each replications, we generate an observation matrix $\mathbf{X}$ of size $(n \times p)$ using the following numerical scheme :

$$\forall i \in [\![1,n]\!], \forall j \in [\![1,p]\!], X_i^j = \sum_{k=j}^{j+K-1} \varepsilon_n Y_i^k + \sum_{k=j+K}^{j+K+2\tau} r_k Y_i^k + \sum_{j+K+2\tau+1}^{j+2\tau+2K} \varepsilon_n Y_i^k \tag{8}$$

where all coefficients $(r_k)_{1\leqslant k\leqslant 1+2\tau}$ are real numbers in $[-1,1]$ (we take $r_1, \dots r_{1+2\tau} \overset{i.i.d}{\sim} \mathcal{U}_{[-1,1]}$ in the simulation), X_i^j is the coefficient of $\mathbf{X}$ on the i^{th} line and the j^{th} column and all random variable $Y_i^k \overset{i.i.d}{\sim} \mathcal{N}(0,1)$ 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.

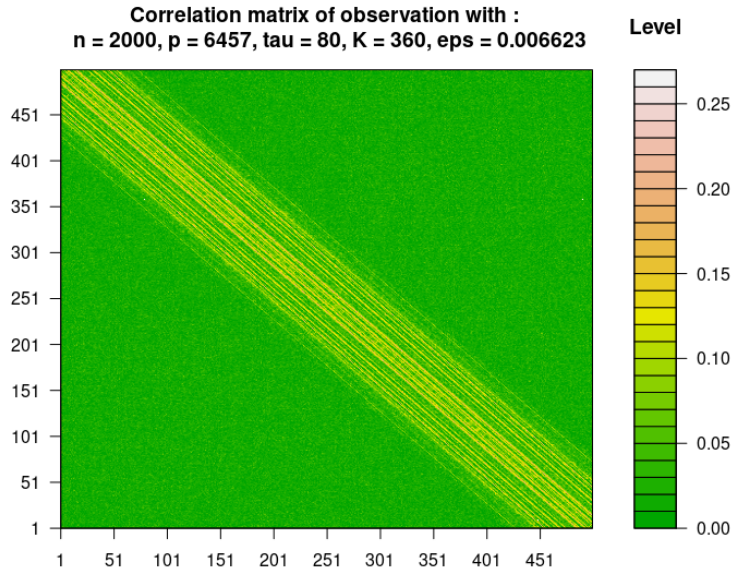

Figure 2: Level plot of the correlation's structure with observations : zoom on the square $1:500$

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. In fact, 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, we see that for n large enough, the sample distribution seems to approximate the limiting one.

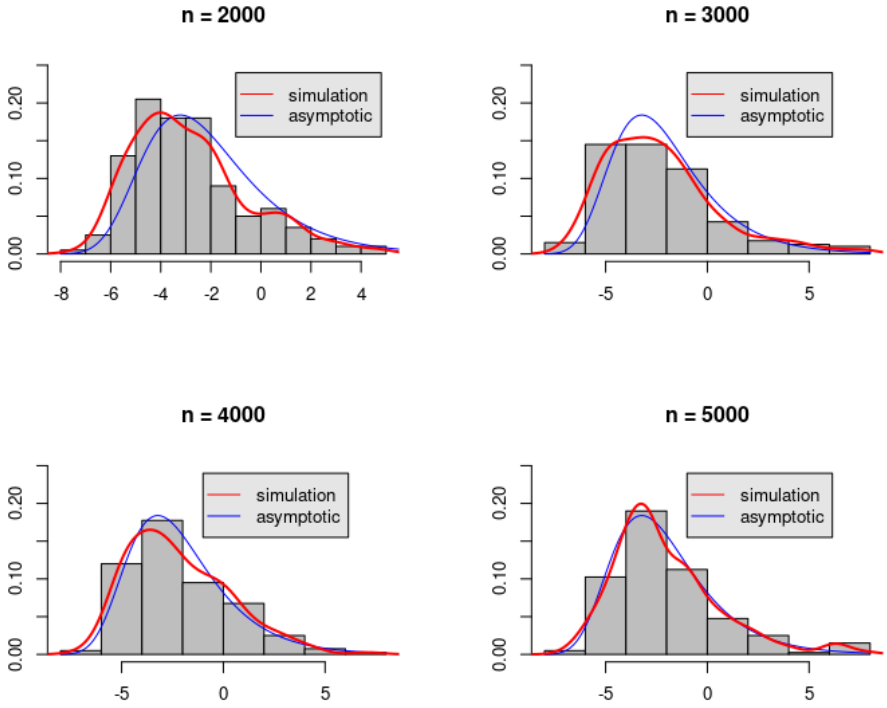

Figure 3: Histograms and Kernel density estimates for $n = 2000, 3000, 4000, 5000$ and for $R = 200$ replications

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. We remind, respectively, the definition of these norms :

$$d_{KS}(\hat{f}, f) = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|, \tag{9}$$

$$d_2(\hat{f}, f) = \int \left|\hat{f}(x) - f(x)\right|^2 dx, \tag{10}$$

$$d_{TV}(\hat{f}, f) = \frac{1}{2} \int \left| \hat{f}(x) - f(x) \right| dx. \tag{11}$$

We observe, in the results displayed in Figure 4, that the difference between both distributions decreases to 0 when n is increasing.

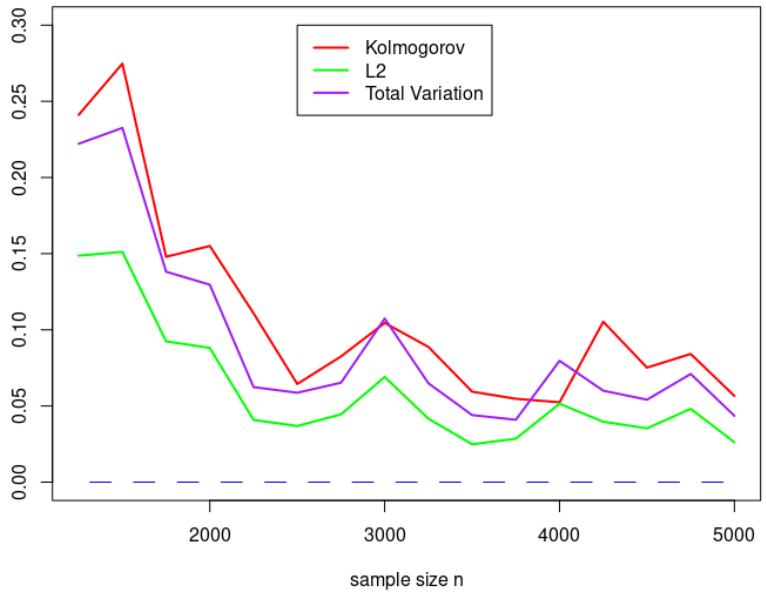

Figure 4: Evolution of Kolmogorov, $\mathbb{L}^2$ and Total Variation norm between simulated and asymptotic behavior

These results provide numerical evidence that our limiting distribution is adequate. We also higlight the fact that the procedure we proposed here allows to compute τ-coherence corresponding to any large matrix X 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 $n = 4000$ (and so $p = 44112$), requires about 90 min to obtain the value of one τ-coherence. In order to obtain more usable (i.e fast) codes in perspective of real-size applications, we are currently exploring HPC strategies to compute correlation blocks using GPGPU computation. We are very confident into the use of GPU to reduce simulation's time.

4 Proof of the main result

In this section, we describe the proof of our main result. First, we would like to higlight 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 this paper.

4.1 Notations

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

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. $I_0 \cup I_K = \Lambda_p^K \cup \Lambda_p^0 \cup \left[E_\delta \cap \overline{I_\tau}\right]$

The following Lemma gives the sizes of these three sets:

Lemma 4.1. *With the previous notations*

$$|I_\tau| = (\tau - 1)\left(\frac{2p-\tau}{2}\right), \tag{12}$$

$$|I_K| = (K+1)\left(\frac{2p-K-2\tau}{2}\right), \tag{13}$$

$$|I_0| = \frac{(p-\tau-K-1)(p-\tau-K)}{2}. \tag{14}$$

4.2 Auxiliary variables

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

$$V_{n,\tau} = \max_{1\leqslant k<j\leqslant p, |k-j|\geqslant\tau} |{}^t X^k X^j| = \max_{\alpha=(k,j)\in I_0\cup I_K} \left|{}^t X^k X^j\right| \tag{15}$$

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

Proposition 4.1. *Under the assumptions of Theorem theorem 2.1, we have*

$$\frac{n^2 L_{n,\tau}^2 - V_{n,\tau}^2}{n} \xrightarrow[n\to+\infty]{\mathbb{P}} 0. \tag{16}$$

The proof of this Proposition is postponed to section 4.

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:

$$V'_{n,\tau} = \max_{\alpha \in \Lambda_p^0 \cup \Lambda_p^K} (Z_\alpha) \tag{17}$$

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

Proposition 4.2. *Let*

$$a_n(y) = \sqrt{4n\log(p_n) - n\log\log(p_n) + ny} \text{ with } y \in \mathbb{R}. \tag{18}$$

We have :

$$\mathbb{P}\left(V'_{n,\tau} > a_n(y)\right) \leqslant \mathbb{P}\left(V_{n,\tau} > a_n(y)\right) \leqslant \mathbb{P}\left(V'_{n,\tau} > a_n(y)\right) + o(1) \tag{19}$$

Proof. (For seek of simplicity, we will denote $a_n(y)$ by a_n in the sequel).
To proove this result, we need the two following technical results whose proofs are postponed to section 4.

Lemma 4.2. *Let a_n be as in formula* (18)*. Then,*

$$\mathbb{P}_0 := \mathbb{P}\left(\left|{}^t X^1 X^{\tau+K+2}\right| > a_n\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{p_n^2}\left(1 + o(1)\right) = O_{n\to+\infty}\left(\frac{1}{p_n^2}\right) \tag{20}$$

Lemma 4.3. *Let a_n be as in formula* (18) *and let us define $c_\gamma := \frac{1}{2}\gamma^2 - 2|\gamma| + 2$ with γ defined in theorem 2.1. Then, for any $d \in [0; c_\gamma[$ and $n \to +\infty$:*

$$\mathbb{P}_K := \mathbb{P}\left(\left|{}^t X^1 X^{\tau+1}\right| > a_n\right) = o\left(p_n^{-d}\right) \tag{21}$$

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

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

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

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

We can use the following straightforward result :

Lemma 4.4.

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

Hence from assumption 3 of theorem 2.1, we have :

$$|E_\delta| = o(p_n^2) \tag{22}$$

Now, we need to prove that $|I_K|\,\mathbb{P}_K + |E_\delta|\,\mathbb{P}_0 \underset{n\to+\infty}{\longrightarrow} 0$. First, from lemma 4.2 and (22),

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

Secondly, using lemma 4.1 (more precisely eq. (13)) and lemma 4.3 we have :

$$|I_K|\,\mathbb{P}_K \underset{n\to+\infty}{\longrightarrow} 0 \quad\Leftrightarrow\quad \nu < c_\gamma - 1 \tag{23}$$

$$\tag{24}$$

and this is fullfilled from assumptions on theorem 2.1. Finally, we obtain :

$$|I_K|\,\mathbb{P}_K + |E_\delta|\,\mathbb{P}_0 \underset{n\to+\infty}{\longrightarrow} 0. \tag{25}$$

Then,

$$\mathbb{P}\left(V_{n,\tau} > a_n\right) \leqslant \mathbb{P}\left(V'_{n,\tau} > a_n\right) + o(1) \tag{26}$$

Also, it is easy to see that :

$$\mathbb{P}\left(V'_{n,\tau} > a_n\right) \leqslant \mathbb{P}\left(V_{n,\tau} > a_n\right). \tag{27}$$

□

Remark 1. *The main constraint so far is $p_nK\mathbb{P}_K \to 0$ when $n \to \infty$ which leads to*

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

Moreover, it also implies the following condition:

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

4.3 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 remind this result, which can be found in [AGG89], in the following lemma:

Lemma 4.5. *The Chen-Stein Method*
Let $\mathcal{I}$ be a set of indices. Let $\alpha \in \mathcal{I}$ and B_α a set of subset of $\mathcal{I}$ (i.e. for all α, $B_\alpha \subset \mathcal{I}$). Let η_α be random variables. For a given $t \in \mathbb{R}$, we define $\lambda := \sum_{\alpha \in \mathcal{I}} \mathbb{P}(\eta_\alpha > t)$. Then,

$$\left| \mathbb{P}\left(\max_{\alpha \in \mathcal{I}} (\eta_\alpha) \leqslant t \right) - e^{-\lambda} \right| \leqslant \min\left(1, \frac{1}{\lambda}\right) . (b_1 + b_2 + b_3), \tag{28}$$

where

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

As we said, this method is an approximation of weekly 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 weekly dependent events to have b_1, b_2 and b_3 small (even null or asymptotically null).

In our case, notations are :

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

First of all, we compute λ_n to assure it converges (as $n \to +\infty$) to a finite value. Then, we compute $b_{1,n}$, $b_{2,n}$ and $b_{3,n}$. Let us start with a preliminary lemma.

Lemma 4.6. *Considering the previous notations, with straightforward computations we obtain the following results :*

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

4.3.1 Computation of λ_n

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

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

According to theorem 2.1, lemma 4.3 and lemma 4.6, we have

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

while, according to lemma 4.2 and lemma 4.6,

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

Finally, we obtain :

$$\lim_{n\to+\infty} (\lambda_n) = \frac{1}{\sqrt{8\pi}}e^{y/2} \tag{29}$$

This quantity appears in the distribution function of the asymptotic Gumbel random variable.

4.3.2 Computation of $b_{1,n}$

We add some notations :

- $B_\alpha^0 := B_\alpha \cap \Lambda_p^0$ and $|B_\alpha^0| \leqslant |B_\alpha| \leqslant 8(\tau+K)p_n$
- $B_\alpha^K := B_\alpha \cap \Lambda_p^K$ and $|B_\alpha^K| \leqslant K^2$
- $\mathbb{P}_\alpha := \mathbb{P}\left(Z_\alpha > a_n\right)$

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 :

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

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

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

$$|\Lambda_p^0|.|B_\alpha^0|.\left(\mathbb{P}_0\right)^2 \quad \sim \quad \frac{1}{2}p_n^2.|B_\alpha^0|.\left(\mathbb{P}_0\right)^2 \leqslant 4(\tau+K)p_n^3.O\left(\frac{1}{p_n^4}\right) = O\left(p_n^{\nu-1}\right) \tag{30}$$

From assumptions on ν we have $\lim\limits_{n\to+\infty}\left[|\Lambda_p^0|.|B_\alpha^0|.\left(\mathbb{P}_0\right)^2\right] = 0$

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

$$|\Lambda_p^0|.|B_\alpha^K|\mathbb{P}_0\mathbb{P}_K \quad \sim \quad \frac{1}{2}p_n^2|B_\alpha^K|\mathbb{P}_0\mathbb{P}_K \leqslant \frac{1}{2}K^2p_n^2\mathbb{P}_0\mathbb{P}_K \leqslant O\left(p_n^{2+2\nu}\right)O\left(p_n^{-2}\right)\mathbb{P}_K = O\left(p_n^{2\nu}\mathbb{P}_K\right)$$

According to lemma 4.3, we will have $\lim\limits_{n\to+\infty}\left[p_n^{2\nu}\mathbb{P}_K\right] = 0$ iff $2\nu < c_\gamma$ which is true from hypothesis 5 in theorem 2.1. Then, we obtain :

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

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

$$|\Lambda_p^K|.|B_\alpha^0|\mathbb{P}_K\mathbb{P}_0 \quad \leqslant \quad p_nK|B_\alpha^0|\mathbb{P}_K\mathbb{P}_0 \leqslant 8p_nK(\tau+K)p_n\mathbb{P}_K\mathbb{P}_0 = O\left(p_n^{2\nu}\mathbb{P}_K\right) \tag{31}$$

So, from previous assumptions on ν, we have :

$$\lim_{n\to+\infty}\left[|\Lambda_p^K|.|B_\alpha^0|\mathbb{P}_K\mathbb{P}_0\right] = 0\,.$$

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

$$|\Lambda_p^K|.|B_\alpha^K|.\left(\mathbb{P}_K\right)^2 \quad \leqslant \quad pK^3\left(\mathbb{P}_K\right)^2 = O\left(p_n^{1+3\nu}\right)\left(\mathbb{P}_K\right)^2 \tag{32}$$

According to lemma 4.3, if $1+3\nu < 2c_\gamma$, we have

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

To conclude, we finally obtain :

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

Remark 2. *The main constraint here is* $p_nK^3\left(\mathbb{P}_K\right)^2 \to 0$ *which is true from condition* $p_nK\mathbb{P}_K \to 0$ *of remark 1.*

4.3.3 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 $b_{2,n}$:

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

Here we introduce some new notations :

- $\mathbb{P}_{\alpha\beta} := \mathbb{P}\left(Z_\alpha > a_n, Z_\beta > a_n\right)$
- $\mathbb{P}_{0i} := \mathbb{P}_{\alpha\beta}\mathbf{1}_{\alpha\in\Lambda_p^0}\mathbf{1}_{\beta\in\Omega_i}$ 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 $\lim_{n\to+\infty}[b_{2,n}] = 0$, 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 :

$$b_{2,n} = \underbrace{\sum_{\alpha\in\Lambda_p^0}\sum_{\beta\in B_\alpha^0}\mathbb{P}_{\alpha\beta}}_{:=Q_1} + \underbrace{\sum_{\alpha\in\Lambda_p^0}\sum_{\beta\in B_\alpha^K}\mathbb{P}_{\alpha\beta}}_{:=Q_2} + \underbrace{\sum_{\alpha\in\Lambda_p^K}\sum_{\beta\in B_\alpha^0}\mathbb{P}_{\alpha\beta}}_{:=Q_3} + \underbrace{\sum_{\alpha\in\Lambda_p^K}\sum_{\beta\in B_\alpha^K}\mathbb{P}_{\alpha\beta}}_{:=Q_4} \tag{33}$$

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

We have :

$$Q_1 \leqslant 4\sum_{i=1}^{8}\sum_{\alpha\in\Lambda_p^0}\sum_{\beta\in\Omega_i}\mathbb{P}_{\alpha\beta} \tag{34}$$

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

$$\begin{aligned} Q_1 \quad \leqslant \quad & \left|\Lambda_p^0\right|.|\Omega_1|\,\mathbb{P}_{01} + \left|\Lambda_p^0\right|.|\Omega_2|\,\mathbb{P}_{02} + \left|\Lambda_p^0\right|.|\Omega_3|\,\mathbb{P}_{03} + \left|\Lambda_p^0\right|.|\Omega_4|\,\mathbb{P}_{04} && (35)\\ & + \left|\Lambda_p^0\right|.|\Omega_5|\,\mathbb{P}_{05} + \left|\Lambda_p^0\right|.|\Omega_6|\,\mathbb{P}_{06} + \left|\Lambda_p^0\right|.|\Omega_7|\,\mathbb{P}_{07} + \left|\Lambda_p^0\right|.|\Omega_8|\,\mathbb{P}_{08} && (36)\end{aligned}$$

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

Lemma 4.7. *Using the previous notations, we have, as $n \to +\infty$:*

$$\left|\Lambda_p^0\right|.|\Omega_1|\,\mathbb{P}_{01} \to 0 \tag{37}$$

Proof. We have :

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

where we use lemma 4.6 for the equivalent. We can write :

$$\mathbb{P}_{01} \quad = \quad \mathbb{P}\left(\left|\sum_{k=1}^{n} u_k^1 u_k^2\right| > a_n, \left|\sum_{k=1}^{n} u_k^3 u_k^4\right| > a_n\right) \tag{38}$$

$$\tag{39}$$

where $\left(u_k^1, u_k^2, u_k^3, u_k^4\right)_{1 \leqslant k \leqslant n} \overset{i.i.d}{\sim} \mathcal{N}_4\left(0, \Sigma_4\right)$ 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

$$\mathbb{P}_{01} \quad \leqslant \quad O\left(p_n^{-2b^2+\varepsilon_1}\right) + O\left(p_n^{-2-2c^2+\varepsilon_2}\right) \text{ as } n \to +\infty \tag{40}$$

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 :

$$\varepsilon_1 < 2b^2 - 2 - 2t\,. \tag{41}$$

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

$$\varepsilon_2 < 2\left(c^2 - t\right)\,. \tag{42}$$

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

$$\left|\Lambda_p^0\right| . \left|\Omega_1\right| \mathbb{P}_{01} \to 0 \text{ as } n \to +\infty\,. \tag{43}$$

□

Lemma 4.8. *Using previous notations, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_2\right| \mathbb{P}_{02} \to 0 \tag{44}$$

Proof. We have :

$$\left|\Lambda_p^0\right| . \left|\Omega_2\right| \mathbb{P}_{02} \quad \leqslant \quad \tau K \left|\Lambda_p^0\right| \mathbb{P}_{02} \sim \frac{1}{2} p_n^2 \tau K \mathbb{P}_{02} = O\left(\tau p_n^{2+\nu} \mathbb{P}_{02}\right)$$

where we use the lemma 4.6 for the equivalence above. In this proof, we almost have the same case than in the proof of lemma 4.7. 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

$$p_n^{2+\nu}\mathbb{P}_{02} \to 0 \tag{45}$$

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 $\tau = o\left(p_n^t\right)$ for any $t > 0$, we have :

$$\left|\Lambda_p^0\right| . \left|\Omega_2\right| \mathbb{P}_{02} \to 0 \,. \tag{46}$$

□

Lemma 4.9. *Using notations previously introduced, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_3\right| \mathbb{P}_{03} \to 0 \tag{47}$$

Proof. This proof is exactly the same than for lemma 4.8 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 4.10. *Using notations previously introduced, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_4\right| \mathbb{P}_{04} \to 0 \tag{48}$$

Proof. We have :

$$\left|\Lambda_p^0\right| . \left|\Omega_4\right| \mathbb{P}_{04} \leqslant \tau p_n \left|\Lambda_p^0\right| \mathbb{P}_{04} \sim \tau p_n^3 \mathbb{P}_{04} \tag{49}$$

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 $\mathbb{P}_{04} = O\left(p_n^{-4+\varepsilon}\right)$ 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 $n \to \infty$ since $\tau = o\left(p_n^t\right)$ for any $t > 0$. □

Lemma 4.11. *Using notations previously introduced, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_5\right| \mathbb{P}_{05} \to 0 \,. \tag{50}$$

Proof. This proof is exactly the same than for lemma 4.10 considering the correlation matrix

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

□

Lemma 4.12. *Using notations previously introduced, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_6\right| \mathbb{P}_{06} \to 0 \tag{51}$$

Proof. This proof is exactly the same than for lemma 4.8 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 :

$$\left|\Lambda_p^0\right| . \left|\Omega_6\right| \mathbb{P}_{06} \leqslant K^2 \left|\Lambda_p^0\right| \mathbb{P}_{06} \sim O\left(p_n^{2+2\nu} \mathbb{P}_{06}\right) \tag{52}$$

with Σ_4 as correlation matrix for the 4-uplet in $\mathbb{P}_{06}$. As for lemma 4.8, we have the following conditions

$$\varepsilon_1 < 2b^2 - 2 - 2\nu \text{ and } \varepsilon_2 < 2\left(c^2 - \nu\right) \tag{53}$$

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

$$\left|\Lambda_p^0\right| . \left|\Omega_6\right| \mathbb{P}_{06} \to 0 \tag{54}$$

□

Lemma 4.13. *Using notations previously introduced, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_7\right| \mathbb{P}_{07} \to 0 \tag{55}$$

Proof. This proof is exactly the same than for lemma 4.10 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 4.14. *Using notations previously introduced, we have :*

$$\left|\Lambda_p^0\right| . \left|\Omega_8\right| \mathbb{P}_{08} \to 0 \tag{56}$$

Proof. This proof is exactly the same than for lemma 4.10 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 3. *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 will 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 :

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

and

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

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 $|r_i| \leqslant 1 - \delta$ 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 $\lim\limits_{n \to +\infty} (\varepsilon_n) = 0$ then, for n large enough, we still have $|r_i| \leqslant 1 - \delta$. Using Cauchy-Schwarz inequality, we can bound :

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

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 :

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

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

$$\sum_{\alpha\in\Lambda^0_{p,I}} \sum_{\beta\in B^K_\alpha} \mathbb{P}_{\alpha\beta} \leqslant 3K^3 p\sqrt{\mathbb{P}_0\mathbb{P}_K} \leqslant 3K^3\mathbb{P}_K^{1/2} O(1) = O\left(p^{3\nu}\mathbb{P}^{1/2}\right) \text{ as } n \to +\infty.$$

Now, using lemma 4.3:

$$p^{3\nu}\mathbb{P}^{1/2} \underset{n\to+\infty}{\longrightarrow} 0 \Leftrightarrow 3\nu < \frac{1}{2}\left(\frac{1}{2}\gamma^2 - 2|\gamma| + 2\right) \Leftrightarrow \nu < \frac{1}{6}c_\gamma \tag{57}$$

which is true according to assumptions of theorem 2.1.

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

- $\Omega^2_1 := \{(u,v) \in B^K_\alpha : u - i < \tau \text{ and } j - v > \tau + K\}$ and $|\Omega^2_1| \leqslant K\tau$
- $\Omega^2_2 := \{(u,v) \in B^K_\alpha : \tau + K < u - i \text{ and } j - v < \tau\}$ and $|\Omega^2_2| \leqslant K\tau$
- $\Omega^2_3 := \{(u,v) \in B^K_\alpha : \tau \leqslant u - i \leqslant \tau + K \text{ and } j - v > \tau + K\}$ and $|\Omega^2_2| \leqslant K^2$
- $\Omega^2_4 := \{(u,v) \in B^K_\alpha : \tau + K < u - i \text{ and } \tau \leqslant j - v \leqslant \tau + K\}$ and $|\Omega^2_4| \leqslant K^2$

We have :

$$\sum_{\alpha\in\Lambda^0_{p,II}} \sum_{\beta\in B^K_\alpha} \mathbb{P}_{\alpha\beta} \leqslant 4\sum_{i=1}^{4} \sum_{\alpha\in\Lambda^0_{p,II}} \sum_{\beta\in\Omega^2_i} \mathbb{P}_{\alpha\beta}$$

.

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

Lemma 4.15. *Considering the same notations as previously :* $\left(u^1_k, u^2_k, u^3_k, u^4_k\right)_{1\leqslant k\leqslant n} \overset{i.i.d}{\sim} \mathcal{N}_4\left(0, \Sigma_4\right)$.

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 :*

$$\sum_{\alpha\in\Lambda^0_{p,II}} \sum_{\beta\in\Omega^2_1} \mathbb{P}_{\alpha\beta} \leqslant O\left(p^{t+\nu+\epsilon}\mathbb{P}_K^{1/2}\right) \tag{58}$$

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

Proof. In order to prove this result, we observe that in this case, for all $k \geqslant 1$, u^2_k is independent of $\{u^1_k, u^3_k, u^4_k\}$. It means that conditionally on u^1_k, we have independence between Z_{12} and Z_{34}.

By consequence, using Cauchy-Schwarz, we obtain :

$$\begin{aligned}
\mathbb{P}_{\alpha\beta} &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{Z_{12}>a_n}\mathbf{1}_{Z_{34}>a_n}|u_k^1, k=1,\dots,n\right]\right] &(59)\\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{Z_{12}>a_n}|u_k^1, k=1,\dots,n\right]\mathbb{E}\left[\mathbf{1}_{Z_{34}>a_n}|u_k^1, k=1,\dots,n\right]\right] &(60)\\
&\leqslant \sqrt{\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{Z_{12}>a_n}|u_k^1, k=1,\dots,n\right]^2\right]\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{Z_{34}>a_n}|u_k^1, k=1,\dots,n\right]^2\right]} &(61)
\end{aligned}$$

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 :

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

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

$$\sum_{\alpha\in\Lambda_{p,II}^0}\sum_{\beta\in\Omega_1^2}\mathbb{P}_{\alpha\beta} \leqslant O\left(p^{t+\nu+\epsilon}\mathbb{P}_K^{1/2}\right) \tag{62}$$

□

Now, from lemma 4.15 we have the condition :

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

which can be fullfilled from condition in eq. (57) and for well-chosen $t > 0$ and $\epsilon > 0$. Finally we obtain, with our condition on ν that :

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

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 :

$$\Sigma_4^2 = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & x & r\\ 0 & x & 1 & \varepsilon_n\\ 0 & r & \varepsilon_n & 1 \end{pmatrix}, \Sigma_4^3 = \begin{pmatrix} 1 & 0 & \varepsilon_n & x\\ 0 & 1 & 0 & 0\\ \varepsilon_n & 0 & 1 & \varepsilon_n\\ x & 0 & \varepsilon_n & 1 \end{pmatrix}, \Sigma_4^4 = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & x & \varepsilon_n\\ 0 & x & 1 & \varepsilon_n\\ 0 & \varepsilon_n & \varepsilon_n & 1 \end{pmatrix} \tag{63}$$

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 4.15, 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 4.15 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:

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

It means that we need to have, according to lemma 4.3 :

$$2\nu<\frac{1}{2}\left(\frac{1}{2}\gamma^2-2|\gamma|+2\right)\Leftrightarrow\nu<\frac{1}{4}\left(\frac{1}{2}\gamma^2-2|\gamma|+2\right). \tag{64}$$

With this condition

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

and

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

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 :

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

and then

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

Remark 4. *The main constaint here is* $\nu<\frac{1}{6}c_\gamma$.

Computation of Q_3 :

We focuse 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 :

- $\Omega_1^3=\{(u,v)\in\Lambda_p^0 : i-u<\tau \text{ and } v-j<\tau\}$ and $|\Omega_1^3|\leqslant\tau^2$.
- $\Omega_2^3=\{(u,v)\in\Lambda_p^0 : i-u<\tau \text{ and } \tau\leqslant v-j\leqslant\tau+K\}$ and $|\Omega_2^3|\leqslant\tau K$.
- $\Omega_3^3=\{(u,v)\in\Lambda_p^0 : i-u<\tau \text{ and } \tau+K<v-j\}$ and $|\Omega_3^3|\leqslant\tau p$.
- $\Omega_4^3=\{(u,v)\in\Lambda_p^0 : \tau\leqslant i-u\leqslant\tau+K \text{ and } v-j<\tau\}$ and $|\Omega_4^3|\leqslant\tau K$.
- $\Omega_5^3=\{(u,v)\in\Lambda_p^0 : \tau+K<i-u \text{ and } v-j<\tau\}$ and $|\Omega_5^3|\leqslant\tau p$.

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

Then, we have :

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

For Q_3, we use the computation of Q_2. Indeed, covariance matrices which are involved in the computation fo 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 :

$$|\Omega_1^3|, |\Omega_2^3|, |\Omega_4^3| \leqslant |\Omega_6^3| \leqslant K^2 \tag{65}$$

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

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

However, from condition in eq. (57), we have $\nu < \frac{1}{3}\left(\frac{1}{4}\gamma^2 - |\gamma| + 1\right)$. But, $\gamma \in]-2+\sqrt{2}, 2-\sqrt{2}[$. Then, $\frac{1}{3}\left(\frac{1}{4}\gamma^2 - |\gamma| + 1\right) \in]\frac{1}{6}, \frac{1}{3}[$. It leads, in particular, to $\nu < \frac{1}{3}$ and then :

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

And so, from equation eq. (65) :

$$\lim_{n \to +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_1^3} \mathbb{P}_{\alpha\beta} \right] = \lim_{n \to +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_2^3} \mathbb{P}_{\alpha\beta} \right] = \lim_{n \to +\infty} \left[\sum_{\alpha \in \Lambda_p^K} \sum_{\beta \in \Omega_4^3} \mathbb{P}_{\alpha\beta} \right] = 0 \tag{66}$$

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 $|r| \leqslant 1-\delta$. 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^3).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 $|\Omega_i^3| \leqslant Kp$ for $i = 3, 5, 7, 8$, we can write :

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

Then, we have exactly the same condition on ν that for equation eq. (64). It means that

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

and it induces that

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

Computation of Q_4 :

This last quantity is simpler because we can write :

$$Q_4 = \sum_{\alpha\in\Lambda_p^K}\sum_{\beta\in B_\alpha^K}\mathbb{P}_{\alpha\beta} \leqslant \sum_{\alpha\in\Lambda_p^K}\sum_{\beta\in B_\alpha^K}\mathbb{P}_K = \left|\Lambda_p^K\right|.\left|B_\alpha^K\right|\mathbb{P}_K \leqslant pK^3\mathbb{P}_k = O\left(p^{1+3\nu}\mathbb{P}_K\right) \text{ as } n\to+\infty$$

So, to have $\lim_{n\to+\infty}(Q_4)=0$ we must have, according to lemma 4.3:

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

Under this assumption on ν we have

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

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

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

Remark 5. *The constraint on ν is $\nu < \frac{1}{3}(c_\gamma - 1)$.*

4.3.4 Computation of $b_{3,n}$

We have :

$$\begin{aligned} b_{3,n} &= \sum_{\alpha\in\Lambda_p^0\cup\Lambda_p^K}\mathbb{E}\left[\left|\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}|\sigma\left(Z_\beta,\beta\in(\Lambda_p^0\cup\Lambda_p^K)\backslash B_\alpha\right)\right]-\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}\right]\right|\right] \\ &= \sum_{\alpha\in\Lambda_p^0}\mathbb{E}\left[\left|\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}|\sigma\left(Z_\beta,\beta\in(\Lambda_p^0\cup\Lambda_p^K)\backslash B_\alpha\right)\right]-\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}\right]\right|\right] \\ &+ \sum_{\alpha\in\Lambda_p^K}\mathbb{E}\left[\left|\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}|\sigma\left(Z_\beta,\beta\in(\Lambda_p^0\cup\Lambda_p^K)\backslash B_\alpha\right)\right]-\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}\right]\right|\right] \end{aligned} \tag{67}$$

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

$$\begin{aligned} b_{3,n} &= \sum_{\alpha\in\Lambda_p^K}\mathbb{E}\left[\left|\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}|\sigma\left(Z_\beta,\beta\in(\Lambda_p^0\cup\Lambda_p^K)\backslash B_\alpha\right)\right]-\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}\right]\right|\right] && (68)\\ &\leqslant \sum_{\alpha\in\Lambda_p^K}\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}|\sigma\left(Z_\beta,\beta\in(\Lambda_p^0\cup\Lambda_p^K)\backslash B_\alpha\right)\right]\right]+\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{Z_\alpha>a_n}\right]\right] && (69)\\ &\leqslant 2\left|\Lambda_p^K\right|\mathbb{P}_K && (70)\\ & && (71) \end{aligned}$$

According to the hypotheses in theorem 2.1, we have

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

Finally, we obtain :

$$\lim_{n\to+\infty}\left[b_{3,n}\right]=0$$

Remark 6. *Gathering remark 1 to 5 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 2.1 for* ν*.*

□

4.4 Proof of Theorem 2.1

We showed in Section 4.3 that

$$\mathbb{P}\left(V'_{n,\tau}\leqslant a_n\right)=\exp\left(-\lambda_n\right)+o\left(1\right)\ \text{ as } n\to+\infty \tag{72}$$

Thanks to lemma 4.2, we have, for n large enough :

$$\mathbb{P}\left(V_{n,\tau}\leqslant a_n\right)=\exp\left(-\lambda_n\right)+o\left(1\right)\ \text{ as } n\to+\infty \tag{73}$$

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

$$\frac{1}{n}V_{n,\tau}^2-4\log\left(p_n\right)+\log\log\left(p_n\right)\xrightarrow[n\to+\infty]{\mathcal{L}}Z \tag{74}$$

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

$$\frac{1}{n\log\left(p_n\right)}V_{n,\tau}^2-4\xrightarrow[n\to+\infty]{\mathbb{P}}0 \tag{75}$$

Then, from Proposition 4.1 we have (6).

□

5 Proofs of technical results

5.1 Proof of Proposition 4.1

We first recall here the basic definition of tightness :

Definition 5.1. *Let* (u_n) *be a real sequence. We say that* (u_n) *is a tight sequence if :*

$$\forall\epsilon>0,\exists K>0,\ \sup_{n\geq1}\left(\mathbb{P}\left(|u_n|\geq K\right)\right)<\epsilon \tag{76}$$

For the proof of proposition 4.1 we use the following lemma which is proved in [Jia04a] (see Lemma 2.2):

Lemma 5.1. *Let τ be an integer. We define $|||A||| = \max_{1\leq i<j\leq p, |i-j|>\tau} |A_{ij}|$ for a $(p\times p)$-matrix. Let $\mathbf{X}$ be a random (n,p)-matrix where $(X^1, X^2, \ldots, X^p)$ are the p columns in $\mathbb{R}^n$ and R_n the empirical correlation matrix of $\mathbf{X}$. Let's define, for any $k \in [\![1,p]\!]$:*

- $h_k = \frac{1}{\sqrt{n}} \parallel X^k - \overline{X^k}\mathbf{1}_n \parallel$ *with* $\overline{X^k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k$
- $c_{n,1} = \max_{1\leqslant k\leqslant p} |h_k - 1|$
- $c_{n,2} = \min_{1\leqslant k\leqslant p} h_k$
- $c_{n,3} = \max_{1\leqslant k\leqslant p} |\overline{X^k}|$
- $V_{n,\tau} = \max_{1\leqslant k<j\leqslant p, |k-j|\geqslant\tau} |{}^tX^kX^j| = \max_{1\leqslant k<j\leqslant p, |k-j|\geqslant\tau} |\sum_{i=1}^{n} X_i^kX_i^j|$

Then,

$$|||nR_n - {}^t\mathbf{X}\mathbf{X}||| \leqslant \frac{c_{n,1}^2 + 2c_{n,1}}{c_{n,2}^2} V_{n,\tau} + n\left(\frac{c_{n,3}}{c_{n,2}}\right)^2 \tag{77}$$

We denote here by $\Delta_n := |nL_{n,\tau} - V_{n,\tau}|$, for any $n \geqslant 1$. We have :

$$\left|n^2L_{n,\tau}^2 - V_{n,\tau}^2\right| = |nL_{n,\tau} - V_{n,\tau}|\,.\,|nL_{n,\tau} + V_{n,\tau}| \leqslant \Delta_n.\,(\Delta_n + 2V_{n,\tau}) \tag{78}$$

Now, we can notice :

$$\Delta_n \leqslant |||nR_n - {}^t\mathbf{X}\mathbf{X}||| \leqslant \frac{c_{n,1}^2 + 2c_{n,1}}{c_{n,2}^2} V_{n,\tau} + n\left(\frac{c_{n,3}}{c_{n,2}}\right)^2 \tag{79}$$

where the upper bound above refers to lemma 5.1. Using definition 5.1, we see that $\left(c_{n,1}\sqrt{\frac{n}{\log(p_n)}}\right)_{n\geqslant 1} := \left(c'_{n,1}\right)_{n\geqslant 1}$ and $\left(c_{n,3}\sqrt{\frac{n}{\log(p_n)}}\right)_{n\geqslant 1} := \left(c'_{n,3}\right)_{n\geqslant 1}$ are both tight sequences. In that way, both sequences $c_{1,n} = c'_{n,1}\sqrt{\frac{\log(p_n)}{n}}$ and $c_{4,n} = c'_{4,n}\sqrt{\frac{\log(p)}{n}}$ are tight too (from Hyp 1 in theorem 2.1, $\frac{\log(p_n)}{n} \to 0$ when $n \to +\infty$).
So,

$$\begin{aligned}
\Delta_n &\leqslant \frac{\frac{\log(p_n)}{n}c_{n,1}'^2 + 2\sqrt{\frac{\log(p_n)}{n}}c'_{n,1}}{c_{n,2}^2} V_{n,\tau} + n\frac{\log(p_n)}{n}\left(\frac{c'_{n,3}}{c_{n,2}}\right)^2 \\
\frac{\Delta_n}{\log(p_n)} &\leqslant \frac{\frac{1}{n}c_{n,1}'^2 + 2\sqrt{\frac{1}{n}}c'_{n,1}}{c_{n,2}^2} V_{n,\tau} + \left(\frac{c'_{n,3}}{c_{n,2}}\right)^2 \\
\frac{\Delta_n}{\log(p_n)} &\leqslant \frac{V_{n,\tau}}{\sqrt{n\log(p_n)}} \frac{c_{n,1}'^2\sqrt{\frac{\log(p_n)}{n}} + 2c'_{n,1}}{c_{n,2}^2} + \left(\frac{c'_{n,3}}{c_{n,2}}\right)^2
\end{aligned}$$

With this inequality, we notice that the sequence $(\Delta'_n)_{n\geqslant 1} := \left(\frac{\Delta_n}{\log(p_n)}\right)_{n\geqslant 1}$ is tight. Finally, with tightness of $(\Delta'_n)_{n\geqslant 1}$, we have from assumptions in theorem 2.1 :

$$\begin{aligned}\frac{\left|n^2L_{n,\tau}^2 - V_{n,\tau}^2\right|}{n} &\leqslant \frac{\Delta_n}{n}\left(\Delta_n + 2V_{n,\tau}\right) = \frac{\Delta'_n\log(p_n)}{n}\left(\Delta'_n\log(p_n) + 2V_{n,\tau}\right)\\ &\leqslant 2\sqrt{\frac{\left(\log(p_n)\right)^3}{n}}\left(\Delta'_n\sqrt{\frac{\log(p_n)}{n}} + V_{n,\tau}\right) \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0.\end{aligned}$$

5.2 Proof of Lemma 4.2

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 5.2. *We consider the following hypotheses :*

1. $\xi_1,\ldots,\xi_n$ *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* $\mathbb{E}\left[e^{t_0|\xi_1|^\alpha}\right] < +\infty$.
3. $(p_n)_{n\in\mathbb{N}^*}$ *such that* $p_n \underset{n\to+\infty}{\longrightarrow} +\infty$ *and* $\log(p_n) = o\left(n^{\frac{\alpha}{2+\alpha}}\right)$ *as* $n\to+\infty$
4. $(y_n)_{n\geqslant 1}$ *such that* $y_n \underset{n\to+\infty}{\longrightarrow} y > 0$

Then,

$$\mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\sum_{k=1}^{n}\xi_k \geqslant y_n\right) \underset{n\to+\infty}{\sim} \frac{1}{y\sqrt{2\pi}}p_n^{-\frac{1}{2}y_n^2}\sqrt{\log(p_n)}^{-1} \tag{80}$$

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

$$\mathbb{P}\left(\left|{}^tX^1X^{\tau+K+2}\right| > a_n\right) = \mathbb{P}(\left(\left|\sum_{i=1}^{n}X_i^1X_i^{\tau+K+2}\right| > a_n\right) \tag{81}$$

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

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

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

4. We have $w_n := \frac{a_n}{\sqrt{n\log(p)}} \underset{n\to+\infty}{\longrightarrow} \sqrt{4} = 2 > 0$
5. According to the hypothesis 1 from theorem 2.1 : $\log(p_n) = o(n^{\frac{1}{3}})$ as $n \underset{n\to+\infty}{\longrightarrow} +\infty$

So we have all hypothesis needed to apply the lemma 5.2, and then :

$$\begin{aligned}
\mathbb{P}\left(\left|{}^tX^1X^{\tau+K+2}\right|>a_n\right) &= \mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\left|{}^tX^1X^{\tau+K+2}\right|>\frac{a_n}{\sqrt{n\log(p_n)}}\right)\\
&= \mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\left|\sum_{i=1}^n X_i^1X_i^{\tau+K+2}\right|>\frac{a_n}{\sqrt{n\log(p_n)}}\right)\\
= \mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\sum_{i=1}^n X_i^1X_i^{\tau+K+2}>\frac{a_n}{\sqrt{n\log(p_n)}}\right)&+\mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\sum_{i=1}^n X_i^1X_i^{\tau+K+2}<-\frac{a_n}{\sqrt{n\log(p_n)}}\right)\\
&= \mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\sum_{i=1}^n X_i^1X_i^{\tau+K+2}>w_n\right)+\mathbb{P}\left(-\frac{1}{\sqrt{n\log(p_n)}}\sum_{i=1}^n X_i^1X_i^{\tau+K+2}>w_n\right) \quad (82)
\end{aligned}$$

From lemma 5.2

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

5.3 Proof of Lemma lemma 4.3

We remind that $\mathbb{P}_K := \mathbb{P}\left(\left|{}^tX^1X^{\tau+1}\right|>a_n\right)$. We will apply once again lemma 5.2, with new quantities ξ and ω:

- $\mathbb{P}_K^+ := \mathbb{P}\left({}^tX^1X^{\tau+1}>a_n\right)$
- $\mathbb{P}_K^- := \mathbb{P}\left({}^tX^1X^{\tau+1}<-a_n\right)$
- $\xi_k := X_k^1X_k^{\tau+1}$
- $w_k := \dfrac{1}{\sqrt{1+\varepsilon_n^2\left(\xi_k-\varepsilon_n\right)}}$.

Notice that $(\xi_k)_{k\geqslant 1}$ are independent due to the independence between each line of $\mathbf{X}_n$. First we compute $\mathbb{E}\left[\xi_k\right]=\varepsilon_n$ and $\mathrm{var}\left(\xi_k\right)=1+\varepsilon_n^2$. So, $\mathbb{E}\left[w_k\right]=0$ and $\mathrm{var}\left(w_k\right)=1$.We will apply the lemma 5.2 with w_k. Then,

$$\mathbb{P}_K^+ = \mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\sum_{k=1}^n w_k > \underbrace{\frac{a_n-n\varepsilon_n}{\sqrt{(1+\varepsilon_n^2)n\log(p_n)}}}_{:=z_n}\right) \quad (83)$$

From hypotheses of theorem 2.1, we have $\lim\limits_{n\to+\infty}\left[z_n\right] := z = 2-\gamma > 0$. 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 :

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

$$p_n^a \mathbb{P}_K^+ \sim \frac{1}{z\sqrt{2\pi}} \exp\left[\left(a - 2 - \frac{1}{2}\gamma^2 + 2\gamma + o(1)\right)\log(p_n)\right] \tag{84}$$

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

$$\lim_{n\to+\infty}\left[p_n^a \mathbb{P}_K^+\right] = 0 \Leftrightarrow a < 2 + \frac{1}{2}\gamma^2 - 2\gamma \tag{85}$$

Analogously, we have :

$$\mathbb{P}_K^- = \mathbb{P}\left(\frac{1}{\sqrt{n\log(p_n)}}\sum_{k=1}^n w_k < -\frac{a_n + n\varepsilon_n}{\sqrt{(1+\varepsilon_n^2)n\log(p_n)}}\right) \tag{86}$$

$$= \mathbb{P}\left(\frac{-1}{\sqrt{n\log(p_n)}}\sum_{k=1}^n w_k > \underbrace{\frac{a_n + n\varepsilon_n}{\sqrt{(1+\varepsilon_n^2)n\log(p_n)}}}_{:=\tilde{z_n}}\right) \tag{87}$$

Thanks to lemma 5.2 and because $\lim_{n\to+\infty}[\tilde{z_n}] := \tilde{z} = 2 + \gamma > 0$, we have :

$$\mathbb{P}_K^- \sim \frac{1}{\tilde{z}\sqrt{2\pi}} \exp\left[\frac{-2}{1+\varepsilon_n^2}\log(p_n)\left(1 + \frac{\varepsilon_n^2}{4}\frac{\log\log(p_n)}{\log(p_n)} + \frac{y}{4}\frac{1}{\log(p_n)} + \frac{n\varepsilon_n^2}{4\log(p_n)} + \frac{\varepsilon_n}{2}\frac{a_n}{\log(p_n)}\right)\right] \tag{88}$$

$$p_n^b \mathbb{P}_K^- \sim \frac{1}{\tilde{z}\sqrt{2\pi}} \exp\left[\left(b - 2 - \frac{1}{2}\gamma^2 - 2\gamma + o(1)\right)\log(p_n)\right] \tag{89}$$

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

$$\lim_{n\to+\infty}\left[p_n^b \mathbb{P}_K^-\right] = 0 \Leftrightarrow b < 2 + \frac{1}{2}\gamma^2 + 2\gamma \tag{90}$$

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. (90) and eq. (85), we obtain, for all $d \in [0; c_\gamma[$ and as $n \to +\infty$:

$$\mathbb{P}_K := \mathbb{P}\left(\left|{}^t X^1 X^{\tau+1}\right| > a_n\right) = o\left(p_n^{-d}\right). \tag{91}$$

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

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.

[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.

[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.

[Don06] D.L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52:1289–1306, 2006.

[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.

[HM18] Johannes Heiny and Thomas Mikosch. Almost sure convergence of the largest and smallest eigenvalues of high-dimensional sample correlation matrices. *Stochastic Processes Appl.*, 128(8):2779–2815, 2018.

[Jia04a] Tiefeng Jiang. The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.*, 14(2):865–880, 2004.

[Jia04b] Tiefeng Jiang. The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā*, 66(1):35–48, 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.

[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.

[Pé09] Sandrine Péché. Universality results for the largest eigenvalues of some sample covariance matrix ensembles. *Probab. Theory Relat. Fields*, 143(3-4):481–516, 2009.

[Pec12] Giovanni Peccati. The Chen-Stein method for Poisson functionals, https://hal.archives-ouvertes.fr/hal-00654235. April 2012.

[R C20] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.

[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.

[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.