axioms

MDPI

Article

A Strong Limit Theorem of the Largest Entries of a Sample Correlation Matrices under a Strong Mixing Assumption

Haozhu Zhao and Yong Zhang *

School of Mathematics, Jilin University, Changchun 130012, China; zhaohz19@mails.jlu.edu.cn
* Correspondence: zyong2661@jlu.edu.cn

Abstract: We are interested in an n by p matrix $\mathbf{X}_n$ where the n rows are strictly stationary α-mixing random vectors and each of the p columns is an independent and identically distributed random vector; $p = p_n$ goes to infinity as $n \to \infty$, satisfiying $0 < c_1 \leq p_n/n^\tau \leq c_2 < \infty$, where $\tau > 0$, $c_2 \geq c_1 > 0$. We obtain a logarithmic law of $L_n = \max_{1\leq i<j\leq p_n} |\rho_{ij}|$ using the Chen–Stein Poisson approximation method, where ρ_{ij} denotes the sample correlation coefficient between the ith column and the jth column of $\mathbf{X}_n$.

Keywords: Chen-SteinPoisson approximation method; mixing sequence; sample correlation matrices

MSC: 60F15

1. Introduction

Random matrix theory is being used in variety of fields, from physics to various areas of mathematics. Ref. [1] test the structure of the regression coefficient matrix under the multivariate linear regression model using the LRT statistic. The correlation coefficient matrix holds significance as a crucial statistic in the multivariate analysis, and its maximum likelihood estimator is the sample correlation matrix. Consider an n by p matrix $\mathbf{X}_n = (X_{k,i})$, $1 \leq k \leq n, 1 \leq i \leq p$, representing observations from a specific multivariate distribution. It has an unknown mean $\mu = (\mu_1, \cdots, \mu_p)$, unknown correlation coefficient matrix $\mathbf{R}$, and unknown covariance matrix Σ.

This paper shows the logarithmic law for the largest entries of a sample correlation matrices while an α-mixing assumption holds. This study is the promotion of the statistical hypothesis testing problem that [2] analyzed. Ref. [2] considered the statistical test as the sample size n and the sample dimension p are both large; the null hypothesis is $H_0 : \mathbf{R} = \mathbf{I}_p$, where $\mathbf{I}_p$ is the identity matrix. The null hypothesis of [2] postulates that the components of $\mathbf{X}_n = (\mathrm{X}^{(1)}, \cdots, \mathrm{X}^{(p)})$ are uncorrelated and, in the case of $\mathbf{X}_n$, have a p-variate normal distribution. Ref. [2]'s test statistic is

$$L_n = \max_{1\leq i<j\leq p_n} |\rho_{ij}|,$$

where

$$\rho_{ij} = \frac{\sum_{k=1}^n (X_{k,i} - \bar{X}^{(i)})(X_{k,j} - \bar{X}^{(j)})}{\sqrt{\sum_{k=1}^n (X_{k,i} - \bar{X}^{(i)})^2}\sqrt{\sum_{k=1}^n (X_{k,j} - \bar{X}^{(j)})^2}}$$

is the Pearson correlation coefficient between $\mathrm{X}^{(i)}$ and $\mathrm{X}^{(j)}$. Then, $\Gamma_n := (\rho_{ij})$ is the sample correlation matrix generated by $\mathbf{X}_n$.

Let $n \geq 2$; let $\bar{X}^{(i)} = \frac{\sum_{k=1}^n X_{k,i}}{n}$, $1 \leq i \leq p$. We have

check for updates

Citation: Zhao, H.; Zhang, Y. A Strong Limit Theorem of the Largest Entries of a Sample Correlation Matrices under a Strong Mixing Assumption. *Axioms* **2023**, *12*, 657. https://doi.org/10.3390/axioms12070657

Academic Editor: Hari Mohan Srivastava

Received: 17 May 2023
Revised: 19 June 2023
Accepted: 30 June 2023
Published: 2 July 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

$$\rho_{ij} = \frac{(\mathrm{X}^{(i)} - \bar{X}^{(i)}\mathrm{e})'(\mathrm{X}^{(j)} - \bar{X}^{(j)}\mathrm{e})}{\|\mathrm{X}^{(i)} - \bar{X}^{(i)}\mathrm{e}\| \cdot \|\mathrm{X}^{(j)} - \bar{X}^{(j)}\mathrm{e}\|} \tag{1}$$

where $\mathrm{e} = (1, \cdots, 1)' \in \mathbb{R}^n$, and $\|\cdot\|$ represent the Euclidean norm in $\mathbb{R}^n$.

Ref. [2] proved the following theorem, focusing on the test statistic L_n as $p = p_n$ and $\mathbf{X}_n = \{X_{k,i}; 1 \le k \le n, 1 \le i \le p\}$ is a set of i.i.d. random variables where the n rows are observations from a certain multivariate distribution and each of the p columns is an n observation from a variable of the population distribution.

Theorem 1. *Suppose that* $\{\xi, X_{k,i}; k \ge 1, i \ge 1\}$ *be i.i.d random variables. Let* $\mathbf{X}_n = (X_{k,i})$ *be an* $n \times p$ *matrix. For any* $\varepsilon > 0$, $E|\xi|^{30-\varepsilon} < \infty$. *If* $n/p \to \gamma \in (0, \infty)$, *then*

$$\lim_{n\to\infty} \sqrt{\frac{n}{\log n}} L_n = 2 \qquad a.s.$$

Theorem 2. *Suppose that* $\{\xi, X_{k,i}; k \ge 1, i \ge 1\}$ *be i.i.d. random variables. Let* $\mathbf{X}_n = (X_{k,i})$ *be an* $n \times p$ *matrix. For some* $\varepsilon > 0$, $E|\xi|^{30+\varepsilon} < \infty$. *If* $n/p \to \gamma$, *then, for any* $y \in \mathbb{R}$,

$$P(nL_n^2 - 4\log n + \log(\log n) \le y) \to e^{-Ke^{-y/2}}$$

as $n \to \infty$, *where* $K = (\gamma^2\sqrt{8\pi})^{-1}$.

Subsequently, many scholars consider the limiting property of the largest entries of the sample correlation matrices under the weaker moment conditions. Ref. [3] showed that the moment condition in Theorem 2 could be $x^6 P(|X_{1,1}X_{1,2}| \ge x) \to 0$ as $x \to \infty$ under $\limsup_{n\to\infty} p/n < \infty$. Ref. [4] showed that Theorem 2 holds under the moment condition $(x^6/\log^3 x)P(|X_{1,1}X_{1,2}| \ge x) \to 0$ as $x \to \infty$. Ref. [5] improved the moment condition and obtained the limit theorems of L_n. Refs. [6,7] showed the results as p/n bounded from zero to infinity under a more relaxed moment assumption. Meanwhile, p reaches infinity as $n \to \infty$; some scholars also pursue the limiting property of the largest entries of the sample correlation matrices, they are interested in the relationship of the sample dimension p and the sample size n. Ref. [8] obtained the limit theorem as $\log p = o(n^\alpha)$ for $0 < \alpha \le 1$ without the Gaussian assumption. Most of the work is based on the assumption of sample independence; it is reasonable to assume the independence of samples, but it is difficult to verify the independence of the samples. Therefore, it is necessary to study the largest entries of the sample correlation matrices under mixing assumptions. Ref. [9] showed the asymptotic distribution of L_n under a φ-mixing assumptions. Ref. [10] showed the asymptotic distribution of L_n under an α-mixing assumption. Under the dependent case, the logarithmic law for L_n remains largely unknown.

We will establish a logarithmic law of L_n under an α-mixing assumption. Let $\{X_n, n \ge 1\}$ be a random variables sequence on probability space (Ω, F, P). Let $\mathcal{F}_a^b$ represent the σ-field generated by the random variables $X_a, X_{a+1}, \cdots, X_b$. For any two σ-fields $\mathcal{A}, \mathcal{B} \subset F$, set $\alpha(\mathcal{A}, \mathcal{B}) := \sup\{|P(AB) - P(A)P(B)|; A \in \mathcal{A}, B \in \mathcal{B}\}$. The strong mixing coefficients of $\{X_n, n \ge 1\}$ are defined as: $\alpha(n) := \sup_{k\ge 1} \alpha(\mathcal{F}_1^k, \mathcal{F}_{k+n}^\infty)$, if $\alpha(n) \to 0$, $\{X_n, n \ge 1\}$ is called α-mixing (see [11]).

The first section introduces the background of, and the motivation for, this study. In Section 2, we show the main result of this paper. In Section 3, we introduce notations and present some classical or elementary facts, which we include for easier referencing. The proof of the main theorem is discussed in Section 4 and is the main novel ingredient of the paper. In Section 5, we present the significance of the main result and its applications.

2. Main Result

Assumption 1. *Let* $\mathbf{X}_k = \left(X_{k,1}, X_{k,2}, \cdots, X_{k,p}\right)$ *be a random vector sequence,* $Var(X_{1,1}) = 1$, *and assume* $\{\mathbf{X}_k; 1 \le k \le n\}$ *be a strictly stationary* α*-mixing random vector sequence, satisfied with* $\alpha(n) = O(e^{-n})$.

Assumption 2. *Let* $\mathbf{X}^{(i)} = (X_{1,i}, X_{2,i}, \cdots, X_{n,i})'$ *be a sequence of random vector, assume* $\{\mathbf{X}^{(i)}; 1 \le i \le p\}$ *be an i.i.d. random vector sequence.*

Remark 1. *For* $H_0 : \mathbf{R} = \mathbf{I}_p$, $X^{(1)}, X^{(2)}, \cdots, X^{(p)}$ *is independent. Let* $\boldsymbol{\mathcal{B}}_i := \{X_{k,i}; 1 \le k \le n\}$ *be a random sampling of* $X^{(i)}$*; it is reasonable to suppose* $\{\boldsymbol{\mathcal{B}}_i; 1 \le i \le p\}$ *is independent. Therefore, Assumption 2 can reasonably obtain the logarithmic law of* L_n.

Theorem 3. *Under the Assumptions 1 and 2, let* $S_n = \Sigma_{k=1}^{n} X_{k,1} X_{k,2}$, *and define* $\sigma^2 := \lim_{n\to\infty} ES_n^2/n$. *Suppose that* $EX_{1,1} = 0$, *and* $0 < c_1 \le p_n/n^{\tau} \le c_2 < \infty$, *where* $\tau > 0$, $c_2 \ge c_1 > 0$. $E|X_{1,1}|^{12+16\tau+\varepsilon} < \infty$, *for some* $\varepsilon > 0$. *Then,*

$$\lim_{n\to\infty} \sqrt{\frac{n}{\log p_n}} L_n = 2\sigma \qquad a.s. \tag{2}$$

Corollary 1. *Suppose that* $\{\xi, X_{k,i}; k \ge 1, i \ge 1\}$ *are i.i.d. random variables,* $EX_{1,1} = 0$. *Let* $\mathbf{X}_n = (X_{k,i})$ *be an* $n \times p$ *matrix, let* $S_n = \Sigma_{k=1}^{n} X_{k,1} X_{k,2}$, *and define* $\sigma^2 := \lim_{n\to\infty} ES_n^2/n$. *Suppose that* $0 < c_1 \le p_n/n^{\tau} \le c_2 < \infty$, *where* $\tau > 0$, $c_2 \ge c_1 > 0$. $E|X_{1,1}|^{12+16\tau+\varepsilon} < \infty$, *for some* $\varepsilon > 0$. *Then,*

$$\lim_{n\to\infty} \sqrt{\frac{n}{\log p_n}} L_n = 2\sigma \qquad a.s. \tag{3}$$

Remark 2. *Theorem 3 considered* $p_n = O(n^{\tau})$, *where* $\tau > 0$. *Theorems 1 and 2 only considered the case* $\tau = 1$. *In Theorem 3, the n rows of* X_n *are strictly stationary* α*-mixing random vectors. This case is more complex than the i.i.d. case considered in Theorems 1 and 2 because of the dependence. Under the i.i.d. assumption, we obtained Corollary 1, and generalized the Theorem 1 from the same order n and* p_n *to* $p_n = O(n^{\tau})$*; the moment condition of Corollary 1 is weaker than Theorem 1.*

3. Preliminaries

The proofs of the main result are intricate and complex. We will gather and establish several technical tools that contribute to the proof of the main result. The subsequent lemmas play a crucial role in validating our findings. Lemma 1 could be seen from [12].

Lemma 1. *Suppose X and Y are random variables taking their values in Borel spaces* $\mathcal{S}_1$, $\mathcal{S}_2$, *respectively; suppose U is a uniform-*$[0,1]$ *random variable independent of* (X, Y). *Assume N is a positive integer,* $\mathcal{H} = \{H_1, H_2, \cdots, H_N\}$ *is a measurable partition of* $\mathcal{S}_2$. *There exists an* $\mathcal{S}_2$*-valued random variable* $Y^* = f(X, Y, U)$; *f is a measurable function from* $\mathcal{S}_1 \times \mathcal{S}_2 \times [0,1]$ *into* $\mathcal{S}_2$, *such that*
(i) Y^* *is independent of X,*
(ii) the probability distribution of Y^* *and Y on* $\mathcal{S}_2$ *are identical,*
(iii) $P(Y^* \text{ and } Y \text{ are not from the same } H_i \in \mathcal{H}) \le (8N)^{1/2} \alpha(\mathcal{B}(X), \mathcal{B}(Y))$.

Lemma 2. *Let* $\{X_i; i \ge 1\}$ *be a strictly stationary* α*-mixing sequence of random variables:* $EX_i = 0$. *If* $\sum_{i=1}^{\infty} [\alpha(i)]^{\frac{g-2}{g}} < \infty$, *for some* $g > 2$, $\sum_{i=1}^{n} i^{1+h} [\alpha(i)]^{1-h} < \infty$ *for some* $0 < h < 1$, *there exists* $K \in \mathbb{R}$ *such that*

$$E\left|\sum_{i=1}^{n} X_i\right|^g < K n^{g/2} \sup_{k\ge 1} E|X_k|^g, \quad n = 1, 2, \cdots.$$

Proof of Lemma 2. Let $Y_i = \frac{X_i}{\sup_{k\geq 1}\|X_k\|_g}$, where $\|X_k\|_g = (E|X_k|^g)^{1/g}$. Then, we have $E|Y_i|^g = \frac{E|X_i|^g}{\sup_{k\geq 1} E|X_k|^g} < \infty$. Using Theorem 1 of [13], we have

$$E\left|\sum_{i=1}^{n} Y_i\right|^g < Kn^{g/2}, \quad n = 1, 2, \cdots.$$

Then, we can obtain

$$E\left|\sum_{i=1}^{n} X_i\right|^g < Kn^{g/2} \sup_{k\geq 1} E|X_k|^g, \quad n = 1, 2, \cdots.$$

□

Lemma 3. *Assume* $\{X_n; n \geq 1\}$ *is a sequence of* α*-mixing random variables,* $X \in \mathcal{F}_{-\infty}^{k}$, $Y \in \mathcal{F}_{k+n}^{\infty}$ *and* $E|X|^p < \infty$, $E|Y|^q < \infty$, $\frac{1}{p} + \frac{1}{q} < 1$. *Then*

$$|EXY - EXEY| \leq 10\|X\|_p\|Y\|_q(\alpha(n))^{1-\frac{1}{p}-\frac{1}{q}}.$$

Proof of Lemma 3. See Corollary 6.1 in [11]. □

Lemma 4. *For any independent random variables sequence* $\{\xi_n; n \geq 1\}$ *with* $E\xi_n = 0$ *and finite variance, there exists an independent normal variables sequence* $\{\eta_n; n \geq 1\}$ *with* $E\eta_n = 0$, $E\eta_n^2 = E\xi_n^2$, *such that*

$$P\left(\max_{k\leq n}\left|\sum_{i=1}^{k}\xi_i - \sum_{i=1}^{k}\eta_i\right| \geq y\right) \leq (AQ)^Q y^{-Q}\sum_{i=1}^{n} E|\xi_i|^Q$$

for all $Q > 2$ *and* $y > 0$, *whenever* $E|\xi_i|^Q < \infty$, $i = 1, 2, \cdots, n$. *Here,* A *is a universal constant.*

Proof of Lemma 4. See [14]. □

Lemma 5. *Let* $\{\eta_k; 1 \leq k \leq n\}$ *be an independent symmetric random variables sequence and* $S_n = \sum_{k=1}^{n}\eta_k$. *Then, there exist positive numbers* C_j *and* D_j *depending only on* j *for each integer* $j \geq 1$, *for all* $t > 0$,

$$P(|S_n| \geq 2jt) \leq C_j P\left(\max_{1\leq j\leq n}|\eta_j| \geq t\right) + D_j(P(|S_n| \geq t))^j.$$

Proof of Lemma 5. See [15]. □

Lemma 6. *Let* $\{X_n; n \geq 1\}$ *be an independent random variables sequence.* $S_k = \sum_{i=1}^{k} X_i$. *For* $\forall\ x > 0$,

$$P\left(\max_{1\leq k\leq n}|S_k| > 2x\right) \leq \frac{P(|S_n| > x)}{\min_{1\leq k\leq n} P(|S_n - S_k| \leq x)}.$$

Proof of Lemma 6. See [16]. □

Lemma 7. *Let* $\{X_n; n \geq 1\}$ *be an independent random variables sequence and* $E|X_n|^g < \infty$ *for any* $g \geq 2$, $n \geq 1$; *then, there exists a positive constant* $C(g)$ *depending only on* g, *such that*

$$E\left(\max_{1\leq i\leq n}\left|\sum_{k=1}^{i}(X_k - EX_k)\right|\right)^g$$

$$\leq \quad C(g)\left(\sum_{k=1}^{n} E|X_k - EX_k|^g + \left(\sum_{k=1}^{n} E|X_k - EX_k|^2\right)^{g/2}\right)$$

Proof of Lemma 7. See [17]. □

Lemma 8. *If* $\{\eta_n; n \geq 1\}$ *is an i.i.d. random variables sequence with* $E\eta_1 = 0$, $E|\eta_1|^p < \infty$, $p \geq 1$, *and* $S_n = \sum_{i=1}^{n} \eta_i$. *Then,*

$$E|S_n|^p = \begin{cases} O(n^{p/2}), & \text{if } p \geq 2, \\ O(n), & \text{if } 1 \leq p < 2. \end{cases}$$

Proof of Lemma 8. See [16]. □

The following is the Chen–Stein method, as shown in [18].

Lemma 9. *Let* $\{\eta_\alpha; \alpha \in \mathbb{I}\}$ *be a random variables sequence on an index set* $\mathbb{I}$ *and* $\{B_\alpha; \alpha \in \mathbb{I}\}$ *is a set of subsets of* $\mathbb{I}$, *then for each* $\alpha \in \mathbb{I}$,$B_\alpha \subset \mathbb{I}$. *For any* $t \in \mathbb{R}$, *let* $\lambda = \sum_{\alpha\in\mathbb{I}} P(\eta_\alpha > t)$. *Then,*

$$\left|P\left(\max_{\alpha\in\mathbb{I}} \eta_\alpha \leq t\right) - e^{-\lambda}\right| \leq (1 \wedge \lambda^{-1})(b_1 + b_2 + b_3),$$

where

$$\begin{aligned} b_1 &= \sum_{\alpha\in\mathbb{I}} \sum_{\beta\in B_\alpha} P(\eta_\alpha > t)P(\eta_\beta > t), \\ b_2 &= \sum_{\alpha\in\mathbb{I}} \sum_{\alpha\neq\beta\in B_\alpha} P(\eta_\alpha > t, \eta_\beta > t), \\ b_3 &= \sum_{\alpha\in\mathbb{I}} E\left|P(\eta_\alpha > t|\sigma(\eta_\beta, \beta \notin B_\alpha)) - P(\eta_\alpha > t)\right|, \end{aligned}$$

and $\sigma(\eta_\beta; \beta \notin B_\alpha)$ *is the* $\sigma-$*algebra generated by* $\{\eta_\beta; \beta \notin B_\alpha\}$. *If* η_α *is independent of* $\{\eta_\beta; \beta \notin B_\alpha\}$ *for each* α, *then* b_3 *vanishes.*

Now we define

$$W_n = \max_{1\leq i<j\leq p_n} \left|\sum_{k=1}^{n} X_{k,i}X_{k,j}\right|, \quad n \geq 1. \tag{4}$$

Define $|||A||| = \max_{1\leq i\neq j\leq n} |a_{i,j}|$ for any square matrix $A = (a_{i,j})$.

Lemma 10. *We can see* $X^{(i)}$ *in (1). For each i, let* $h_i = \|X^{(i)} - \bar{X}^{(i)}\mathbf{e}\|/\sqrt{n}$. *Then,*

$$|||n\Gamma_n - X_n'X_n||| \leq (b_{n,1}^2 + 2b_{n,1})W_n b_{n,3}^{-2} + nb_{n,3}^{-2}b_{n,4}^2,$$

where

$$\begin{aligned} b_{n,1} &= \max_{1\leq i\leq p_n} |h_i - 1|, \quad W_n = \max_{1\leq i<j\leq p_n} |(X^{(i)})'X^{(j)}| \\ b_{n,3} &= \min_{1\leq i\leq p_n} h_i, \qquad b_{n,4} = \max_{1\leq i\leq p_n} |\bar{X}^{(i)}|. \end{aligned}$$

Proof of Lemma 10. See [19]. □

4. Proofs

The next lemma refers to Lemma 2 in [20]; we generalize its result to the α-mixing condition.

Lemma 11. *Under the Assumptions 1 and 2, let $a > 1/2$, $b \geq 0$ and $M > 0$ be constants.*

$$(1) \qquad E|X_{1,1}|^{(1+b)/a} < \infty;$$

$$(2) \quad c = \begin{cases} EX_{1,1}, & \text{if } a \leq 1, \\ \text{any number}, & \text{if } a > 1 \end{cases}$$

is the sufficient condition for

$$\max_{j \leq Mn^b} \left| n^{-a} \sum_{i=1}^{n} (X_{i,j} - c) \right| \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty$$

Proof of Lemma 11. Without loss of generality, suppose $c = 0$. Since, for $\eta > 0$ and $N \geq 1$,

$$\begin{aligned} & P\left\{ \max_{j \leq Mn^b} \left| \frac{1}{n^a} \sum_{i=1}^{n} X_{i,j} \right| \geq \eta, i.o. \right\} \\ \leq & \sum_{k \geq N} P\left\{ \max_{2^{k-1} < n \leq 2^k} \max_{j \leq M2^{kb}} \left| \sum_{i=1}^{n} X_{i,j} \right| \geq \eta' 2^{ka} \right\} \\ \leq & \sum_{k \geq N} M2^{kb} P\left\{ \max_{2^{k-1} < n \leq 2^k} \left| \sum_{i=1}^{n} X_{i,1} \right| \geq 2^{ka} \varepsilon' \right\}, \end{aligned}$$

where $\eta' = 2^{-a}\eta$. To conclude that the probability on the left-hand side of this inequality is equal to zero, it is sufficient to show that

$$\sum_{k=1}^{\infty} 2^{kb} P\left\{ \max_{n \leq 2^k} \left| \sum_{i=1}^{n} X_{i,1} \right| \geq 2^{ka} \eta \right\} < \infty. \tag{5}$$

Let $Y_{i,k} = X_{i,1} I\{|X_{i,1}| < 2^{ka}\}$ and $Z_{i,k} = Y_{i,k} - EY_{i,k}$. Then $|Z_{i,k}| \leq 2^{ka+1}$ and $EZ_{i,k} = 0$. Let g be an integer such that $g(a - 1/2) > b + 2a$. It is easy to see

$$\begin{aligned} X_{i,1} &= X_{i,1} I\{|X_{i,1}| < 2^{ka}\} + X_{i,1} I\{|X_{i,1}| \geq 2^{ka}\} \\ &= Z_{i,k} + EY_{i,k} + X_{i,1} I\{|X_{i,1}| \geq 2^{ka}\}. \end{aligned}$$

We have that

$$\begin{aligned} & \sum_{k=1}^{\infty} 2^{kb} P\left\{ \max_{n \leq 2^k} \left| \sum_{i=1}^{n} X_{i,1} \right| \geq 2^{ka} \eta \right\} \\ \leq & \sum_{k=1}^{\infty} 2^{kb} P\left\{ \max_{n \leq 2^k} \left| \sum_{i=1}^{n} Z_{i,k} \right| \geq \frac{2^{ka}\eta}{4} \right\} + \sum_{k=1}^{\infty} 2^{kb} P\left\{ \max_{n \leq 2^k} \left| \sum_{i=1}^{n} EY_{i,k} \right| \geq \frac{2^{ka}\eta}{2} \right\} \\ + & \sum_{k=1}^{\infty} 2^{kb} P\left\{ \max_{n \leq 2^k} \left| \sum_{i=1}^{n} X_{i,1} I\{|X_{i,1}| \geq 2^{ka}\} \right| \geq \frac{2^{ka}\eta}{4} \right\}. \end{aligned}$$

Then, note that $\{Z_{i,k}\}$ is a sequence of α-mixing random variables by Assumption 1. Using Lemma 2, we obtain

$$\begin{aligned} P\left\{ \max_{n \leq 2^k} \left| \sum_{i=1}^{n} Z_{i,k} \right| \geq \frac{\eta 2^{ka}}{4} \right\} &\leq C \frac{E\left| \sum_{i=1}^{2^k} Z_{i,k} \right|^g}{2^{kga}} \\ &\leq C \frac{2^{kg/2} E|Z_{1,k}^g|}{2^{kga}}, \end{aligned} \tag{6}$$

It is easy to see that $ga - b - 1 > g(a - 1/2) - (b + 2a) > 0$, $EZ_{1,k}^2 \le EX_{1,1}^2 < \infty$,

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{kb-kga+kg/2} E\left|Z_{1,k}^g\right| \le C \sum_{k=1}^{\infty} 2^{k(b-ga+g/2)} E\left|X_{1,1}^g\right| I\left\{|X_{1,1}| < 2^{ka}\right\} \\
&\le C \sum_{k=1}^{\infty} 2^{k(b-ga+g/2)} \left(\sum_{l=1}^{k} E|X_{1,1}|^g I\left\{2^{a(l-1)} \le |X_{1,1}| < 2^{al}\right\} + 1 \right) \\
&\le C \sum_{l=1}^{\infty} 2^{l(b-ga+g/2)} \left(E|X_{1,1}|^g I\left\{2^{a(l-1)} \le |X_{1,1}| < 2^{al}\right\} + 1 \right) \\
&\le C \sum_{l=1}^{\infty} 2^{l(b+2a-g(a-1/2))} E|X_{1,1}|^{(b+1)/a} I\left\{2^{a(l-1)} \le |X_{1,1}| < 2^{al}\right\} + C_1 < \infty.
\end{aligned}$$

where C_1 and C are constants; these can be different. We can obtain that

$$\sum_{k=N}^{\infty} 2^{kb} P\left\{ \max_{n \le 2^k} \left| \sum_{i=1}^{n} Z_{i,k} \right| \ge \eta 2^{ka} \right\} < \infty. \tag{7}$$

We could estimate $EY_{i,k}$ for a large k. We have

$$\begin{aligned}
&\max_{n \le 2^k} \left| \sum_{i=1}^{n} EY_{ik} \right| \le 2^k |EY_{1k}| \\
&\le \begin{cases} 2^k E|X_{11}| I\{|X_{11}| \ge 2^{ka}\} \le 2^{k(a-b)} E|X_{11}|^{\frac{(1+b)}{a}} I\{|X_{11}| \ge 2^{ka}\}, & \text{if } a \le 1+b \\ 2^k \log k + 2^{k(a-b)} E|X_{11}|^{\frac{(1+b)}{a}} I\{|X_{11}| > \log k\}, & \text{if } a > 1+b \end{cases} \\
&\le 2^{-1} \eta 2^{ka}
\end{aligned} \tag{8}$$

for all $\varepsilon > 0$. Hence,

$$\sum_{k=N}^{\infty} 2^{kb} P\left\{ \max_{n \le 2^k} \left| \sum_{i=1}^{n} EY_{ik} \right| \ge \frac{\eta 2^{ka}}{2} \right\} < \infty. \tag{9}$$

Finally, since $E|X_{11}|^{(b+1)/a} < \infty$, we have

$$\begin{aligned}
\sum_{k=1}^{\infty} 2^{kb} P\left(\bigcup_{i=1}^{2^k} \left\{ |X_{i1}| \ge 2^{ka} \right\} \right) &\le \sum_{k=1}^{\infty} 2^{k(b+1)} P\left\{ |X_{11}| \ge 2^{ka} \right\} \\
&< C E|X_{1,1}|^{\frac{b+1}{a}} < \infty.
\end{aligned} \tag{10}$$

Hence,

$$\sum_{k=N}^{\infty} 2^{kb} P\left\{ \max_{n \le 2^k} \left| \sum_{i=1}^{n} X_{i1} I\{|X_{i1}| \ge 2_{ka}\} \right| \ge \frac{\eta 2^{ka}}{4} \right\} < \infty. \tag{11}$$

Then, (5) follows from (7), (9) and (11). □

Lemma 12. *Under the Assumptions 1 and 2, $EX_{1,1} = 0$, $Var(X_{1,1}) = 1$. Suppose that $0 < c_1 \le p_n / n^\tau \le c_2 < \infty$, where $\tau > 0$, $c_2 \ge c_1 > 0$. If, for some $a \in (0, 1/2)$, $E|X_{1,1}|^{(2+2\tau)/(1-a)} < \infty$, then*

$$n^a b_{n,1} \to 0 \quad \text{a.s.}, \quad b_{n,3} \to 1 \quad \text{a.s.} \quad \text{and} \quad n^a b_{n,4} \to 0 \quad \text{a.s.}$$

as $n \to \infty$.

Proof of Lemma 12. It is easy to see that $\|\mathbf{X}^{(i)} - \bar{X}^{(i)}\|^2 = (\mathbf{X}^{(i)})'\mathbf{X}^{(i)} - n|\bar{X}^{(i)}|^2$. For any $x > 0$, we know that $|x-1| \le |x^2-1|$. We obtain

$$\begin{aligned} n^a b_{n,1} &= n^a \max_{1\le i\le p_n} \left| \sqrt{\frac{(\mathbf{X}^{(i)})'\mathbf{X}^{(i)}}{n} - |\bar{X}^{(i)}|^2} - 1 \right| \\ &\le \max_{1\le i\le p_n} \left| \frac{(\mathbf{X}^{(i)})'\mathbf{X}^{(i)} - n}{n^{1-a}} \right| + \left(n^{\frac{a}{2}} \max_{1\le i\le p_n} |\bar{X}^{(i)}| \right)^2. \end{aligned} \tag{12}$$

Note that $(\mathbf{X}^{(i)})'\mathbf{X}^{(i)} = \sum_{k=1}^n X_{k,i}^2$. By Lemma 11, when $E|X_{1,1}|^{(2+2\tau)/(1-a)} < \infty$, the first and second maximum presented above reach zero. This is true under the assumptions of Theorem 3. Therefore, the first limit is proved. We can obtain the second limit through the first one. Since $E|X_{1,1}|^{(1+\tau)/(1-a)} < \infty$, we have

$$n^a b_{n,4} = n^a \max_{1\le i\le p_n} |\bar{X}^{(i)}| = \max_{1\le i\le p_n} \left| \frac{\sum_{k=1}^n X_{k,i}}{n^{1-a}} \right|$$

the limit that $n^a b_{n,4} \to 0$ a.s. is proved by the relationship between $n^a b_{n,4}$ and the rightmost term in (12). □

Let $\frac{1}{4} - \delta < \mu < \frac{1}{4}$, where $0 < \delta < \frac{1+8\varepsilon}{4(9+16\tau+\varepsilon)}$ is sufficiently small for $\varepsilon > 0$.

$$\begin{aligned} S_{n,i,j} &= \sum_{k=1}^n X_{k,i}X_{k,j}, \quad Y_{k,i,j} = X_{k,i}X_{k,j} I\{|X_{k,i}X_{k,j}| \le n^\mu\}, \\ S'_{n,i,j} &= \sum_{k=1}^n (Y_{k,i,j} - EY_{k,i,j}). \end{aligned}$$

Let $0 < \rho < 1$, ρ move close to $1/2$, $0 < \alpha < 1$. α moves close to 0, such that $\rho - \alpha\rho - 2\mu > 0$. Let $z = z_n = [n^\rho]$, $m_n = [n/(z_n + q_n)] \sim n^{1-\rho}$, $q = q_n = [n^{\alpha\rho}]$, $N_n = m_n(z_n + q_n)$. Let

$$\begin{aligned} I_{i,n} &= \{j : (i+1)z + iq + 1 \le j \le (i+1)(z+q)\}, \\ H_{i,n} &= \{j : i(z+q) + 1 \le j \le (i+1)z + iq\}, \end{aligned}$$

let $v_{m,i,j} = \sum_{k\in I_{m,n}} (Y_{k,i,j} - EY_{k,i,j})$, $u_{m,i,j} = \sum_{k\in H_{m,n}} (Y_{k,i,j} - EY_{k,i,j})$, $1 \le m \le m_n$,.

Lemma 13. *Under the condition of Theorem 3, let* $\{Y^*_{m,i,j}; m = 1, 2, \cdots, m_n\}$ *be i.i.d. normal random variables,* $E|Y^*_{m,i,j}| = 0$ *and variance* $Eu^2_{m,i,j}$, $\sigma^2 = \lim_{n\to\infty} ES_n^2/n$. *Then*

$$\lim_{n\to\infty} \frac{ES_n^2}{\sum_{m=1}^{m_n} EY^{*2}_{m,i,j}} = 1.$$

Proof of Lemma 13. We have $ES_n^2 = E\left(S'_{n,1,2} + \sum_{k=1}^n X_{k1}X_{k2} I\{|X_{k1}X_{k2}| \ge n^\mu\}\right)^2$, where we could write $S'_{n,1,2} = \sum_{m=1}^{m_n} u_{m,1,2} + \sum_{m=1}^{m_n} v_{m,1,2} + \sum_{k=N_{m_n}+1}^n (Y_{k,1,2} - EY_{k,1,2})$. Since $E|X_{1,1}|^{12+16\tau+\varepsilon} < \infty$, for some $\varepsilon > 0$, it follows that

$$\begin{aligned} & \sum_{k=1}^n E|X_{k,1}X_{k,2}| I\{|X_{k,1}X_{k,2}| \ge n^\mu\} \\ \le & \sum_{k=1}^n \left[E|X_{k,1}| I\left\{|X_{k,1}| \ge n^{\frac{\mu}{2}}\right\} \right] E|X_{k,2}| + \sum_{k=1}^n E\left[|X_{k,2}| I\left\{|X_{k,2}| \ge n^{\frac{\mu}{2}}\right\} \right] E|X_{k,1}| \\ \le & \sum_{k=1}^n E|X_{k,1}|^{12+16\tau+\varepsilon} E|X_{k,1}|^{1-(12+16\tau+\varepsilon)} I\left\{|X_{k,1}| \ge n^{\frac{\mu}{2}}\right\} E|X_{k,2}| \end{aligned}$$

$$
\begin{aligned}
&+\sum_{k=1}^{n} E\left|X_{k,2}\right|^{12+16\tau+\varepsilon} E\left|X_{k,2}\right|^{1-(12+16\tau+\varepsilon)} I\left\{\left|X_{k,2}\right| \geq n^{\frac{\mu}{2}}\right\} E|X_{k,1}| \\
\leq\ & Cn^{1-\frac{\mu}{2}(11+16\tau+\varepsilon)} = o\left(\sqrt{\frac{n}{\log n}}\right),
\end{aligned}
$$

as $n \to \infty$. Therefore,

$$
\sum_{k=1}^{n} E|X_{k,1}X_{k,2}| I\{|X_{k,1}X_{k,2}| \geq n^{\mu}\} = o\left(\sqrt{\frac{n}{\log n}}\right),
$$

as $n \to \infty$. By Lemma 3, it follows that

$$
\begin{aligned}
& E\left(\sum_{m=1}^{m_n} v_{m,1,2}\right)^2 = \sum_{m=1}^{m_n} E(v_{m,1,2})^2 + 2\sum_{j=2}^{m_n}\sum_{m=1}^{j-1}\left(Ev_{m,1,2}v_{j,1,2} - Ev_{m,1,2}Ev_{j,1,2}\right) \\
\leq\ & \sum_{m=1}^{m_n}\sum_{k\in I_{m,n}} E(Y_{k,1,2} - EY_{k,1,2})^2 + 2\sum_{m=1}^{m_n}\sum_{\substack{k<j \\ k,j\in I_{m,n}}}\left(EY_{k,1,2}Y_{j,1,2} - EY_{k,1,2}EY_{j,1,2}\right) \\
& +2\sum_{j=2}^{m_n}\sum_{m=1}^{j-1}\left(Ev_{m,1,2}v_{j,1,2} - Ev_{m,1,2}Ev_{j,1,2}\right) \\
\leq\ & Cm_nq_n + 2\sum_{m=1}^{m_n}\sum_{\substack{k<j \\ k,j\in I_{m,n}}} 10(E|Y_{k,1,2}|^p)^{1/p}(E|Y_{j,1,2}|^{q'})^{1/q'}(\alpha(I_{m,n}))^{1-\frac{1}{p}-\frac{1}{q'}} \\
+\ & 2\sum_{j=2}^{m_n}\sum_{m=1}^{j-1} 10\left(E|v_{m,1,2}|^p\right)^{1/p}\left(E\left|v_{j,1,2}\right|^{q'}\right)^{1/q'}\left(\alpha(|I_{m,n}|)\right)^{1-\frac{1}{p}-\frac{1}{q'}} \\
\leq\ & C\left(m_nq_n + m_n^2q_n^2(e^{-n^{\alpha\rho}})^{1-\frac{1}{p}-\frac{1}{q'}} + m_n^2\left(n^{\frac{(\alpha\rho+2\mu)p}{2}}\right)^{1/p}\left(n^{\frac{(\alpha\rho+2\mu)q'}{2}}\right)^{1/q'}\left(e^{-n^{\alpha\rho}}\right)^{1-\frac{1}{p}-\frac{1}{q'}}\right) \\
\leq\ & Cn^{1-\rho+\alpha\rho} = o\left(\frac{n}{\log n}\right),
\end{aligned}
$$

as $n \to \infty$, where $2 < p < 12 + 16\tau + \varepsilon$, $2 < q' < 12 + 16\tau + \varepsilon$, $1/p + 1/q' < 1$. We have

$$
\begin{aligned}
& E\left(\sum_{k=N_{m_n}+1}^{n}(Y_{k,1,2} - EY_{k,1,2})\right)^2 \\
\leq\ & \sum_{k=N_{m_n}+1}^{n} E(Y_{k,1,2} - EY_{k,1,2})^2 + 2\sum_{j=N_{m_n}+2}^{n}\sum_{k=N_{m_n}+1}^{j-1}\left(EY_{k,1,2}Y_{j,1,2} - EY_{k,1,2}EY_{j,1,2}\right) \\
\leq\ & C(n^{\rho} + n^{\alpha\rho}) + 2\sum_{j=N_{m_n}+2}^{n}\sum_{k=N_{m_n}+1}^{j-1} 10(E|Y_{k,1,2}|^p)^{1/p}(E|Y_{j,1,2}|^{q'})^{1/q'}(\alpha(I_{m,n}))^{1-\frac{1}{p}-\frac{1}{q'}} \\
\leq\ & C(n^{\rho} + n^{\alpha\rho}) + C(n^{\rho} + n^{\alpha\rho})^2(e^{-n^{\alpha\rho}})^{1-\frac{1}{p}-\frac{1}{q'}} = o\left(\frac{n}{\log n}\right)
\end{aligned}
$$

as $n \to \infty$, where $2 < p < 12 + 16\tau + \varepsilon$, $2 < q' < 12 + 16\tau + \varepsilon$, $1/p + 1/q' < 1$. We have that $ES_n^2 = E\left(\sum_{m=1}^{m_n}(u_{m,1,2} - Eu_{m,1,2})\right)^2 + o(n/\log n)$, by Lemma 3, we can obtain

$$
\begin{aligned}
& \left|E\left(\sum_{m=1}^{m_n}(u_{m,1,2} - Eu_{m,1,2})\right)^2 - \sum_{m=1}^{m_n} EY_{m,1,2}^{*2}\right| \\
=\ & \left|E\sum_{m=1}^{m_n}(u_{m,1,2} - Eu_{m,1,2})^2 + 2E\sum_{j=2}^{m_n}\sum_{m=1}^{j-1}(u_{m,1,2} - Eu_{m,1,2})(u_{j,1,2} - Eu_{j,1,2}) - \sum_{m=1}^{m_n} EY_{m,1,2}^{*2}\right|
\end{aligned}
$$

$$\begin{aligned}
&\leq 2\left|\sum_{j=2}^{m_n}\sum_{m=1}^{j-1}\left(Eu_{m,1,2}u_{j,1,2}-Eu_{m,1,2}Eu_{j,1,2}\right)\right|\\
&\leq 2\sum_{j=2}^{m_n}\sum_{m=1}^{j-1}10\left(E|u_{m,1,2}|^{p}\right)^{1/p}\left(E\left|u_{j,1,2}\right|^{q'}\right)^{1/q'}\left(\alpha(|I_{m,n}|)\right)^{1-\frac{1}{p}-\frac{1}{q'}}\\
&\leq Cm_n^2\left(n^{\frac{(\rho+2\mu)p}{2}}\right)^{1/p}\left(n^{\frac{(\rho+2\mu)q'}{2}}\right)^{1/q'}\left(e^{-n^{\alpha\rho}}\right)^{1-\frac{1}{p}-\frac{1}{q'}}\\
&\leq Cn^{2-\rho+2\mu}\left(e^{-n^{\alpha\rho}}\right)^{1-\frac{1}{p}-\frac{1}{q'}}=o\left(\frac{n}{\log n}\right)
\end{aligned}$$

as $n\to\infty$, where $1/p+1/q'<1$. Hence,

$$\left|ES_n^2-\sum_{m=1}^{m_n}EY_{m,i,j}^{*2}\right|=o\left(\frac{n}{\log n}\right),$$

as $n\to\infty$. Recall that $\lim_{n\to\infty}\frac{ES_n^2}{n}=\sigma^2$, Conclusively,

$$\lim_{n\to\infty}\frac{ES_n^2}{\sum_{m=1}^{m_n}EY_{m,i,j}^{*2}}=1,$$

where $EY_{m,i,j}^{*2}=Eu_{m,i,j}^2$. □

Lemma 14. *Let* $\{Y_{m,i,j}^{*};m=1,2,\cdots,m_n\}$ *be i.i.d. normal random variables,* $EY_{1,1,1}^{*}=0$ *and variance* $Eu_{m,i,j}^2$. *Then,*

$$\limsup_{n\to\infty}\frac{\max_{1\leq i<j\leq p_n}\left|\sum_{m=1}^{m_n}Y_{m,i,j}^{*}\right|}{\sqrt{n\log p_n}}\leq 2\sigma\qquad\text{a.s.}\tag{13}$$

Proof of Lemma 14. Choose $t\in(0,1)$, set $\omega_n=(2+t)\sigma\sqrt{n\log p_n}$; we can suppose that $\sigma=1$. Using Lemma 13, we have $\lim_{n\to\infty}ES_n^2/\sum_{m=1}^{m_n}EY_{m,i,j}^{*2}=1$, as $n\to\infty$, where $EY_{m,i,j}^{*2}=Eu_{m,i,j}^2$. Then, we can obtain

$$\begin{aligned}
&\max_{1\leq i\neq j<\infty}P\left(\left|\sum_{m=1}^{m_n}Y_{m,i,j}^{*}\right|>\omega_n\right)=2\left[1-\Phi\left(\frac{(2+t)\sqrt{n\log p_n}}{\sqrt{\sum_{m=1}^{m_n}Eu_{m,i,j}^{*2}}}\right)\right]\\
\leq& C\frac{\sqrt{\sum_{m=1}^{m_n}Eu_{m,i,j}^{*2}}}{\sqrt{2\pi}(2+t)\sqrt{n\log p_n}}\exp\left(-\frac{(2+t)^2n\log p_n}{2\sum_{m=1}^{m_n}Eu_{m,i,j}^{*2}}\right)\\
\leq& C\frac{1}{(2+t)\sqrt{2\pi\log p_n}p_n^{(2+t)^2/2}}=O\left(\frac{1}{n^{\tau(2+t)^2/2}}\right),
\end{aligned}\tag{14}$$

as n is large, where we use

$$1-\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_x^{\infty}e^{-t^2/2}dt\sim\frac{1}{\sqrt{2\pi}x}e^{-x^2/2}\tag{15}$$

as $x\to+\infty$ (shown in [16]). We define $W_n'=\max_{1\leq i<j\leq p_n}\left|\sum_{m=1}^{m_n}Y_{m,i,j}^{*}\right|$, $n_k=k^g$ for any integer $g>\left(2+\tau(2+t)^2\right)/\left(t^2+4t\right)\tau$,

$$\max_{n_k\leq n\leq n_{k+1}}W_n'\leq\max_{1\leq i\neq j\leq p_{n_{k+1}}}\left|\sum_{m=1}^{m_{n_k}}Y_{m,i,j}^{*}\right|+r_n,\tag{16}$$

where

$$r_n = \max_{1\le i\ne j\le p_{n_{k+1}}} \max_{n_k\le n\le n_{k+1}} \left| \sum_{m=1}^{m_n} Y^*_{m,i,j} - \sum_{m=1}^{m_{n_k}} Y^*_{m,i,j} \right|. \tag{17}$$

By (14),

$$\begin{aligned} & P\left(\max_{1\le i\ne j\le p_{n_{k+1}}} \left| \sum_{m=1}^{m_{n_k}} Y^*_{m,i,j} \right| > \omega_{n_k} \right) \\ \le\ & p^2_{n_{k+1}} P\left(\frac{\left|\sum_{m=1}^{m_{n_k}} Y^*_{m,1,2}\right|}{\sqrt{\sum_{m=1}^{m_{n_k}} Eu^{*2}_{m,i,j}}} > \frac{\omega_{n_k}}{\sqrt{\sum_{m=1}^{m_{n_k}} Eu^{*2}_{m,i,j}}} \right) \\ \le\ & 2p^2_{n_{k+1}} \left(1 - \Phi\left(\frac{(2+t)\sqrt{n_k \log p_{n_k}}}{\sqrt{\sum_{m=1}^{m_{n_k}} Eu^{*2}_{m,i,j}}} \right) \right) \\ \le\ & C(k+1)^{2g\tau} \frac{\sqrt{\sum_{m=1}^{m_{n_k}} Eu^{*2}_{m,i,j}}}{\sqrt{2\pi}(2+t)\sqrt{n_k \log p_{n_k}}} \exp\left(-\frac{(2+t)^2 n_k \log p_{n_k}}{2\sum_{m=1}^{m_{n_k}} Eu^{*2}_{m,i,j}} \right) \\ =\ & O\left(k^{-\frac{(t^2+4t)g\tau}{2}} \right). \end{aligned}$$

Since $\sum_k k^{-(t^2+4t)g\tau/2} < \infty$, using the Borel–Cantelli lemma,

$$\limsup_{k\to\infty} \frac{\max_{1\le i\ne j\le p_{n_k}} \left|\sum_{m=1}^{m_{n_k}} Y^*_{m,i,j}\right|}{\sqrt{n_k \log p_{n_k}}} \le 2+t \quad \text{a.s.} \tag{18}$$

Now, let us estimate r_n as in (17).

Let partial sums $S_0 = 0$ and $S_l = \sum_{m=1}^{l} Y^*_{m,i,j}$. Observe that the distribution of $\sum_{m=1}^{m_n} Y^*_{m,i,j} - \sum_{m=1}^{m_{n_k}} Y^*_{m,i,j}$ is equal to that of $S_{m_n - m_{n_k}}$ for all $m_n \ge m_{n_k}$. Thus, by Lemma 6, we have

$$\begin{aligned} & P\left(r_n \ge t\sqrt{n_k \log p_{n_k}} \right) \\ \le & p^2_{n_{k+1}} P\left(\max_{1\le l\le n_{k+1}-n_k} |S_l| \ge t\sqrt{n_k \log p_{n_k}} \right) \\ \le & 2p^2_{n_{k+1}} P\left(\left|S_{n_{k+1}-n_k}\right| \ge (t/2)\sqrt{n_k \log p_{n_k}} \right) \end{aligned} \tag{19}$$

as n is sufficiently large, since $\min_{1\le l\le n_{k+1}-n_k} P\left(\left|S_{n_{k+1}-n_k} - S_l\right| \le (t/2)\sqrt{n_k \log p_{n_k}}\right) \ge 1/2$, where Ottaviani's inequality in Lemma 6 is used in the last inequality. Note that, for fixed g and t, $(t/2)\sqrt{n_k \log p_{n_k}} \ge (2+t)\sqrt{(n_{k+1}-n_k)\log(p_{n_{k+1}} - p_{n_k})}$ as n is sufficiently large. From (15), we have

$$2p^2_{n_{k+1}} P\left(\left|S_{n_{k+1}-n_k}\right| \ge (t/2)\sqrt{n_k \log p_{n_k}} \right) = O\left(k^{-(\tau g-1)(2+t)^2/2 + 2g\tau} \right)$$

Therefore,

$$P\left(r_n \ge t\sqrt{n_k \log p_{n_k}} \right) = O\left(k^{-u}\right),$$

where $u = (\tau g - 1)(2+t)^2/2 - 2g\tau$, since g is chosen such that $u > 1$. Using the Borel–Cantelli lemma again, we obtain

$$\limsup_{k\to\infty} \frac{r_n}{\sqrt{n_k \log p_{n_k}}} \le t \quad \text{a.s.} \tag{20}$$

Using (16), (18) and (20), we obtain that

$$\limsup_{k\to\infty} \frac{\max_{n_k\le n\le n_{k+1}} W'_n}{\sqrt{n_k \log p_{n_k}}} \le 2+2t \quad \text{a.s.}$$

for any sufficiently small $t>0$. Then, we have the inequality (13) in Lemma 14. □

Lemma 15. *Let* $\{Y^*_{m,i,j}; m=1,2,\cdots,m_n\}$ *be i.i.d. normal random variables,* $EY^*_{m,i,j}=0$ *and variance* $Eu^2_{m,i,j}$. *Then,*

$$\liminf_{n\to\infty} \frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} Y^*_{m,i,j}\right|}{\sqrt{n\log p_n}} \ge 2\sigma \qquad \text{a.s.} \tag{21}$$

Proof of Lemma 15. First prove that

$$P(W'_n \le v_n) = O\left(\frac{1}{n^{t'}}\right) \tag{22}$$

as $n\to\infty$, for $t'>0$, depending only on t and the distribution of $X_{1,1}X_{1,2}$ only. For any $t\in(0,1)$, set $v_n=(2-t)\sigma\sqrt{n\log p_n}$, and $\sigma^2=\lim_{n\to\infty} ES^2_n/n$, we can suppose $\sigma=1$. Take an integer g satisfying $g>1/t'$. Then, we have $P(W'_{n_k}\le v_{n_k})=O\left(1/k^{t'g}\right)$. Since $\sum_k k^{-t'g}<\infty$, using the Borel–Cantelli lemma, we obtain

$$\liminf_{k\to\infty} \frac{W'_{n_k}}{\sqrt{n_k\log p_{n_k}}} \ge 2-t \quad \text{a.s.} \tag{23}$$

for any $t\in(0,1)$. We can see the definition of r_n in (17), and obtain

$$\inf_{n_k\le n\le n_{k+1}} W'_n \ge W'_{n_k} - r_n.$$

From (20) and (23), we have that

$$\liminf_{k\to\infty} \frac{\inf_{n_k\le n\le n_{k+1}} W'_n}{\sqrt{n_k\log p_{n_k}}} \ge 2-2t \quad \text{a.s.}$$

for any $t>0$ that is small enough. This suggests (21) of Lemma 15.

Now, we prove (22) using Lemma 9.

Let $I=\{(i,j);1\le i<j\le p\}$. Set $\alpha=(i,j)\in I$, $B_\alpha=\{(k,l)\in I;$ *one of* k *and* $l=i$ *or* j *but* $(k,l)\ne\alpha\}$, $\eta_\alpha=|\sum_{m=1}^{m_n} Y^*_{m,i,j}|$, $t=v_n$ and $A_\alpha=A_{ij}=\{|\sum_{m=1}^{m_n} Y^*_{m,i,j}|>v_n\}$. Usnig Lemma 9,

$$P(W'_n\le v_n)\le e^{-\lambda_n}+b_{1,n}+b_{2,n}. \tag{24}$$

Evidently,

$$\begin{aligned} \lambda_n &= \frac{p(p-1)}{2}P(A_{12}),\\ b_{1,n} &\le 2p^3P(A_{12})^2,\\ b_{2,n} &\le 2p^3P(A_{12}A_{13}). \end{aligned} \tag{25}$$

Recall that $\sum_{m=1}^{m_n} Y^*_{m,i,j}$ is a sum of i.i.d. normal random variables. Recall (15). We have

$$\begin{aligned} P(A_{12}) &= P\left(\left|\sum_{m=1}^{m_n} Y^*_{m,1,2}\right| > v_n\right) \\ &= P\left(\frac{\left|\sum_{m=1}^{m_n} Y^*_{m,1,2}\right|}{\sqrt{\sum_{m=1}^{m_n} Eu^{*2}_{m,1,2}}} > \frac{(2-t)\sqrt{n\log p_n}}{\sqrt{\sum_{m=1}^{m_n} Eu^{*2}_{m,1,2}}}\right) \\ &= 2\left[1-\Phi\left(\frac{(2-t)\sqrt{n\log p_n}}{\sqrt{\sum_{m=1}^{m_n} Eu^{*2}_{m,1,2}}}\right)\right] \\ &\le C\frac{\sqrt{\sum_{m=1}^{m_n} Eu^{*2}_{m,1,2}}}{(2-t)\sqrt{2\pi n\log p_n}}\exp\left(-\frac{(2-t)^2 n\log p_n}{2\sum_{m=1}^{m_n} Eu^{*2}_{m,1,2}}\right) = O\left(\frac{1}{n^{\tau(2-t)^2/2}}\right) \end{aligned} \tag{26}$$

as $n \to \infty$. Provided that $E|X_{1,1}|^{12+16\tau+\varepsilon} < \infty$ and $v_n/\sqrt{n\log p_n} \to 2-t$,

$$P(A_{12}A_{13}) = P\left(\left|\sum_{m=1}^{m_n} Y^*_{m,1,2}\right| \ge v_n, \left|\sum_{m=1}^{m_n} Y^*_{m,1,3}\right| \ge v_n\right). \tag{27}$$

The two events in (27) are conditionally independent given $Y^*_{m,i,j}$'s. P^1 and E^1 represent the conditional probability and expectation of $\{Y^*_{m,i,j}; 1 \le m \le m_n, 1 \le i,j \le p_n\}$, respectively. Then, the probability in (27) is

$$E\left[P^1\left(\left|\sum_{m=1}^{m_n} Y^*_{m,1,2}\right| \ge (2-t)\sqrt{n\log p_n}\right)^2\right]. \tag{28}$$

Set

$$A_n(s) := \left\{\frac{1}{n}\left|\sum_{m=1}^{m_n}\left(\left|Y^*_{m,1,2}\right|^s - E\left|Y^*_{m,1,2}\right|^s\right)\right| \le \tilde{\delta}\right\}$$

for $s \ge 2$ and $\tilde{\delta} \in (1, \frac{1}{2})$. Choose $\beta \in (a^2+2, q/(a^2+1))$ and $r = a^2+1$. Let $\zeta_m = \left|Y^*_{m,1,2}\right|^\beta - E\left|Y^*_{m,1,2}\right|^\beta$ for $m = 1, 2, \cdots, m_n$. Then, $E|\zeta_1|^r < \infty$. Using the Chebyshev inequality and Lemma 8,

$$P(A_n(\beta)^c) = O(n^{-f(r)}) \tag{29}$$

as $n \to \infty$, where

$$f(r) = \begin{cases} r-1, & 1 < r \le 2 \\ r/2, & r \ge 2 \end{cases} \tag{30}$$

Let $\{\zeta'_m; 1 \le m \le m_n\}$ be an independent copy of $\{\zeta_m; 1 \le m \le m_n\}$. Using (29), $P(|\sum_{m=1}^{m_n} \zeta_m| \le n\tilde{\delta}/2) \ge 1/2$ for any n that is large enough, we have

$$P\left(\left|\sum_{m=1}^{m_n} \zeta_m\right| > n\tilde{\delta}\right) = O\left(n^{-f(r)}\right) \tag{31}$$

by repeating (29). Choose an integer $j \ge 1$, set $\nu = n\tilde{\delta}/4j$. Lemma 5 implies that there are positive constants C_j and D_j, satisfying

$$P\left(\left|\sum_{m=1}^{m_n}(\zeta_m - \zeta'_m)\right| > n\tilde{\delta}/2\right) = P\left(\left|\sum_{m=1}^{m_n}(\zeta_m - \zeta'_m)\right| > 2j\nu\right)$$

$$\leq \quad C_j P\left(\max_{1\leq m\leq m_n}|\zeta_m-\zeta'_m|>\nu\right)+D_j P\left(\left|\sum_{m=1}^{m_n}(\zeta_m-\zeta'_m)\right|>\nu\right)^j.$$

Since $E|\zeta_1|^r<\infty$, $P(\max_{1\leq m\leq m_n}|\zeta_m-\zeta'_m|>\nu)=O(n^{1-r})$. From the equality in (31), we have that

$$\left(P\left(\left|\sum_{m=1}^{m_n}(\zeta_m-\zeta'_m)\right|>\nu\right)\right)^j=O(n^{-jf(r)})$$

Take $j=[\frac{(r-1)}{f(r)}]+1$. We obtain

$$P\left(\left|\sum_{m=1}^{m_n}(\zeta_m-\zeta'_m)\right|>n\tilde{\delta}/2\right)=O(n^{1-r}) \tag{32}$$

as $n\to\infty$. Then, we have

$$P(A_n(\beta)^c)=O(n^{1-r})$$

as $n\to\infty$, since (29), (31) and (32). Hence,

$$\begin{aligned}&P(A_{12}A_{13})\\ \leq& E\left[P^1\left(\left|\sum_{m=1}^{m_n}Y^*_{m,1,2}\right|\geq(2-t)\sqrt{n\log p_n}\right)^2 I_{A_n(s)\cap A_n(2)}\right]+P(A_n(s)^c).\end{aligned} \tag{33}$$

Since

$$\begin{aligned}&P^1\left(\left|\sum_{m=1}^{m_n}Y^*_{m,1,2}\right|\geq(2-t)\sqrt{n\log p_n}\right)\\ \leq& C\exp\left(-\frac{(2-t)^2 n\log p_n}{2\sum_{m=1}^{m_n}Eu^{*2}_{m,1,2}}\right),\end{aligned} \tag{34}$$

we can obtain

$$\begin{aligned}&P^1\left(\left|\sum_{m=1}^{m_n}Y^*_{m,1,2}\right|\geq(2-t)\sqrt{n\log p_n}\right)^2 I_{A_n(s)\cap A_n(2)}\\ \leq& C\exp\left(-\frac{(2-t)^2 n\log p_n}{\sum_{m=1}^{m_n}Eu^{*2}_{m,1,2}}\right)=O\left(n^{b-(2-t)^2}\right)\end{aligned} \tag{35}$$

for any $b>0$. If both b and t are small enough, we have that

$$e^{-\lambda_n}\leq e^{-n^t},\qquad b_{1,n}\leq\frac{1}{\sqrt{n}}\quad and\quad b_{2,n}\leq\frac{1}{\sqrt{n}} \tag{36}$$

for any n that is large enough. Since (24) and (36), (22) holds. □

Lemma 16. *Under the condition of Theorem 3, take* $T_n=\max_{1\leq i<j\leq p_n}\left|S'_{n,i,j}\right|$*; then,*

$$\limsup_{n\to\infty}\frac{T_n}{\sqrt{n\log p_n}}\leq 2\sigma\qquad\text{a.s.} \tag{37}$$

$$\liminf_{n\to\infty}\frac{T_n}{\sqrt{n\log p_n}}\geq 2\sigma\qquad\text{a.s.} \tag{38}$$

Proof of Lemma 16. Using Markov inequality and Lemma 2, for $\forall\ \delta' > 0$, we can obtain

$$\begin{aligned} & P\left(\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} v_{m,i,j}\right| \ge \delta'\sqrt{n\log p_n}\right) \le C\frac{p_n^2 E\left|\sum_{m=1}^{m_n} v_{m,1,2}\right|^q}{(n\log p_n)^{q/2}} \\ \le\ & C\frac{p_n^2 m_n^{q/2} q_n^{q/2} E|X_{1,1}X_{1,2}|^q I\{|X_{1,1}X_{1,2}| \le n^{\mu}\}}{(n\log n)^{q/2}} \\ \le\ & C\frac{1}{(\log n)^{\frac{q}{2}} n^{(\rho-\alpha\rho-2\mu))_2^q-2\tau}} = O\left(\frac{1}{n^{1+\varepsilon'}}\right) \end{aligned}$$

for $\varepsilon' > 0$ and sufficiently large q, using the Borel–Cantelli lemma, we have

$$\lim_{n\to\infty}\frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} v_{m,i,j}\right|}{\sqrt{n\log p_n}} = 0, \quad \text{a.s.}$$

and using the Markov inequality and Lemma 2, for $\forall\ \delta' > 0$,

$$\begin{aligned} & P\left(\max_{1\le i<j\le p_n}\left|\sum_{k=N_{m_n}+1}^{n}\left(Y_{k,i,j} - EY_{k,i,j}\right)\right| \ge \delta'\sqrt{n\log p_n}\right) \\ \le\ & C\frac{p_n^2 E\left|\sum_{k=N_{m_n}+1}^{n}(Y_{k,1,2} - EY_{k,1,2})\right|^q}{(n\log p_n)^{q/2}} \\ \le\ & C\frac{p_n^2 (z_n+q_n)^{q/2} E|X_{1,1}X_{1,2}|^q I\{|X_{1,1}X_{1,2}| \le n^{\mu}\}}{(n\log n)^{q/2}} \\ \le\ & C\frac{n^{2\tau}(n^{\rho}+n^{\alpha\rho})^{q/2}}{n^{(1-2\mu)q}(\log n)^{q/2}} = O\left(\frac{1}{n^{1+\varepsilon'}}\right), \end{aligned}$$

for sufficiently large q, using the Borel–Cantelli lemma, we can obtain

$$\lim_{n\to\infty}\frac{\max_{1\le i<j\le p_n}\left|\sum_{k=N_{m_n}+1}^{n}\left(Y_{k,i,j} - EY_{k,i,j}\right)\right|}{\sqrt{n\log p_n}} = 0, \quad \text{a.s.}$$

Hence, we only need to prove

$$\lim_{n\to\infty}\frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} u_{m,i,j}\right|}{\sqrt{n\log p_n}} = 2\sigma, \quad \text{a.s.}$$

Since Lemma 1, we have an independent random variables sequence $\{u^*_{m,i,j}; 1 \le m \le m_n\}$, $\{u^*_{m,i,j}; 1 \le m \le m_n\}$ and $\{u_{m,i,j}; 1 \le m \le m_n\}$ have same distribution, and we could obtain that $P\left(u^*_{i,1,2} \text{ and } u_{i,1,2} \text{ are not from the same } H_i \in \mathcal{H}\right) \le \left(8N\right)^{1/2}\alpha(\mathcal{B}(X), \mathcal{B}(Y)\right)$. We can prove that

$$\begin{aligned} & P\left(\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n}\left(u_{m,i,j} - u^*_{m,i,j}\right)\right| \ge \delta'\sqrt{n\log p_n}\right) \\ \le\ & p_n^2 m_n (8m_n)^{1/2}\alpha(|I_{1,n}|) \\ \le\ & C\frac{1}{n^{\frac{3}{2}\rho-\frac{3}{2}-2\tau} e^{n^{\alpha\rho}}} = o\left(\frac{1}{n^{1+\varepsilon'}}\right), \end{aligned}$$

Using the Borel–Cantelli lemma, we only need to prove

$$\lim_{n\to\infty}\frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} u^*_{m,i,j}\right|}{\sqrt{n\log p_n}} = 2\sigma, \quad \text{a.s.}$$

Let $u^*_{m,i,j} = \sum_{k\in H_{m,n}}(Y^i_{k,i,j} - EY^i_{k,i,j})$, $1 \le m \le m_n$, where $Y^i_{k,i,j} = X^i_{k,i}X^i_{k,j}I\{|X^i_{k,i}X^i_{k,j}| \le n^\mu\}$, $\{X^i_{k,j}; k \in H_{i,n}\}$ is an independent replication of $\{X_{k,j}; k \in H_{i,n}\}$. Thus, we only need to prove

$$(1)\quad \limsup_{n\to\infty} \frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} u^*_{m,i,j}\right|}{\sqrt{n\log p_n}} \le 2\sigma \qquad \text{a.s.}$$

$$(2)\quad \liminf_{n\to\infty} \frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n} u^*_{m,i,j}\right|}{\sqrt{n\log p_n}} \ge 2\sigma \qquad \text{a.s.}$$

Let $Y^*_{m,i,j}, 1 \le m \le m_n$ be an independent normal random variables sequence with variance $Var\left(u^*_{m,i,j}\right)$. Since Lemma 4, we have that

$$\begin{aligned} & P\left(\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n}\left(u^*_{m,i,j} - Y^*_{m,i,j}\right)\right| \ge \delta'\sqrt{n\log p_n}\right) \le C\frac{p_n^2\sum_{i=1}^{m_n} E\left|u^*_{i,1,2}\right|^q}{(n\log n)^{\frac{q}{2}}} \\ \le\ & C\frac{p_n^2 m_n z_n^{\frac{q}{2}}(E|X^i_{k,1}X^i_{k,2}|^2 I\{|X^i_{k,1}X^i_{k,2}| \le n^\mu\})^{\frac{q}{2}}}{(n\log p_n)^{\frac{q}{2}}} \\ +\ & C\frac{p_n^2 m_n z_n E|X^i_{k,1}X^i_{k,2}|^q I\{|X^i_{k,1}X^i_{k,2}| \le n^\mu\}}{(n\log p_n)^{\frac{q}{2}}} \\ \le\ & C\frac{1}{(\log p_n)^{\frac{q}{2}} n^{(1-\rho)\frac{q}{2}+\rho-1-2\tau}} + C\frac{1}{(\log p_n)^{\frac{q}{2}} n^{(1-2\mu)\frac{q}{2}-1-2\tau}} = O\left(\frac{1}{n^{1+\varepsilon'}}\right). \end{aligned}$$

for sufficiently large q. Using the Borel–Cantelli lemma,

$$\lim_{n\to\infty} \frac{\max_{1\le i<j\le p_n}\left|\sum_{m=1}^{m_n}\left(u^*_{m,i,j} - Y^*_{m,i,j}\right)\right|}{\sqrt{n\log p_n}} = 0, \quad \text{a.s.}$$

Thus, with Lemmas 14 and 15 and Borel–Cantelli lemma, we could obtain the result. □

Proof of Lemma 3. We could find W_n in (4). Let $a = 1/3$, since $E|X_{1,1}|^{12+16\tau+\varepsilon} < \infty$. Using the triangle inequality and Lemmas 10 and 12, we have that

$$|nL_n - W_n| \le |||n\Gamma_n - \mathbf{X}'_n\mathbf{X}_n||| \le 4n^{-1/3}W_n + 2n^{1/3} \qquad \text{a.s.} \tag{39}$$

as n is large enough. Hence,

$$\begin{aligned} \left|\frac{\sqrt{n}L_n}{\sqrt{\log p_n}} - \frac{W_n}{\sqrt{n\log p_n}}\right| &= \frac{1}{\sqrt{n\log p_n}}|nL_n - W_n| \le \frac{4n^{-1/3}W_n}{\sqrt{n\log p_n}} + \frac{2n^{1/3}}{\sqrt{n\log p_n}} \\ &\le \frac{4W_n}{n^{5/6}\sqrt{\log p_n}} + \frac{1}{n^{1/6}\sqrt{\log p_n}} \to 0, \qquad \text{a.s.} \end{aligned}$$

If $\lim_{n\to\infty}\frac{W_n}{\sqrt{n\log p_n}} = 2\sigma$, a.s. then $\lim_{n\to\infty}\frac{W_n}{n^{5/6}\sqrt{\log p_n}} = 0$, a.s. Hence, in order to prove Theorem 3, we need to show that

$$\lim_{n\to\infty}\frac{W_n}{\sqrt{n\log p_n}} = 2\sigma \qquad \text{a.s.} \tag{40}$$

Take $T_n = \max_{1\le i<j\le p_n}\left|S'_{n,i,j}\right|$. We have that

$$|W_n - T_n| \le \max_{1\le i<j\le p_n}\left|\sum_{k=1}^{n} X_{k,i}X_{k,j}I\left\{|X_{k,i}X_{k,j}| \ge n^{\mu}\right\}\right| =: U_n,$$

Using Lemma 2, let $q = 3$ for any $\delta' > 0$, since $1/4 - \delta < \mu < 1/4$, where $0 < \delta < \frac{1+8\varepsilon}{4(9+16\tau+\varepsilon)}$; for some $\varepsilon > 0$, we have

$$\begin{aligned}
&P\left(U_n \ge \delta'\sqrt{n\log p_n}\right)\\
\le& P\left(\max_{1\le i<j\le p_n}\left|\sum_{k=1}^{n} X_{k,i}X_{k,j}I\left\{\left|X_{k,i}X_{k,j}\right| \ge n^{\mu}\right\}\right| \ge \delta'\sqrt{n\log p_n}\right)\\
\le& p_n^2\frac{E\left|\sum_{k=1}^{n} X_{k,1}X_{k,2}I\left\{\left|X_{k,1}X_{k,2}\right| \ge n^{\mu}\right\}\right|^q}{\left(\delta'\sqrt{n\log p_n}\right)^q}\\
\le& \frac{Cn^{2\tau}n^{\frac{3}{2}}E|X_{1,1}X_{1,2}|^3 I\{|X_{1,1}X_{1,2}| \ge n^{\mu}\}}{n^{3/2}(\log p_n)^{3/2}}\\
\le& \frac{Cn^{2\tau}E|X_{1,1}|^3 I\left\{|X_{1,1}| \ge n^{\frac{\mu}{2}}\right\}E|X_{1,2}|^3}{(\log p_n)^{3/2}} + \frac{Cn^{2\tau}E|X_{1,2}|^3 I\left\{|X_{1,2}| \ge n^{\frac{\mu}{2}}\right\}E|X_{1,1}|^3}{(\log p_n)^{3/2}} \qquad (41)\\
\le& \frac{Cn^{2\tau}E|X_{1,1}|^{12+16\tau+\varepsilon}E|X_{1,1}|^{3-(12+16\tau+\varepsilon)} I\left\{|X_{1,1}| \ge n^{\frac{\mu}{2}}\right\}E|X_{1,2}|^3}{(\log p_n)^{3/2}}\\
&+ \frac{Cn^{2\tau}E|X_{1,2}|^{12+16\tau+\varepsilon}E|X_{1,2}|^{3-(12+16\tau+\varepsilon)} I\left\{|X_{1,2}| \ge n^{\frac{\mu}{2}}\right\}E|X_{1,1}|^3}{(\log p_n)^{3/2}}\\
\le& C\frac{1}{(\log p_n)^{3/2}n^{\frac{\mu}{2}(9+16\tau+\varepsilon)-2\tau}} = o\left(\frac{1}{n^{1+\varepsilon'}}\right),
\end{aligned}$$

Using the Borel–Cantelli lemma,

$$\lim_{n\to\infty}\frac{U_n}{\sqrt{n\log p_n}} = 0 \qquad \text{a.s.}$$

To prove (40), we need to show that

$$\lim_{n\to\infty}\frac{T_n}{\sqrt{n\log p_n}} = 2\sigma \qquad \text{a.s.}$$

We could obtain $\lim_{n\to\infty} T_n/\sqrt{n\log p_n} = 2\sigma$ by Lemma 16. (40) holds from Lemma 16. Since (40), we can obtain $4n^{-1/3}W_n = O(n^{1/6}\log p_n)$ a.s.. Therefore, $nLn - W_n = O(n^{1/3})$. We could prove Theorem 3 by (40). □

5. Examples

In certain applications, such as the construction of compressed sensing matrices, the means $\mu_i = EX^{(i)}$ and $\mu_j = EX^{(j)}$ are given and one is interested in

$$\tilde{\rho}_{ij} = \frac{(\mathrm{X}^{(i)} - \mu_i)^T(\mathrm{X}^{(j)} - \mu_j)}{\|\mathrm{X}^{(i)} - \mu_i\| \cdot \|\mathrm{X}^{(j)} - \mu_j\|},\ 1 \le i, j \le p$$

The corresponding coherence is defined by

$$\tilde{L}_n = \max_{1\le i<j\le p}|\tilde{\rho}_{ij}|.$$

Compressed sensing is a rapidly evolving field, aiming to construct measurement matrices $X_{n\times p}$ that enable the exact recovery of any k-sparse signal $\beta \in \mathbb{R}^p$ from linear measurements $y = X\beta$ using computationally efficient recovery algorithms.

Two commonly employed conditions in compressed sensing are the Restricted Isometry Property (RIP) and the Mutual Incoherence Property (MIP). In this paper, the derived limiting laws can be utilized to assess the likelihood of a random matrix satisfying the MIP condition, as demonstrated by Cai and Jiang [19].

Example 1. *The MIP condition, which is frequently utilized, requires the pairwise correlations among the column vectors of X_n, denoted as $X_n = (X^{(1)}, X^{(2)}, \cdots, X^{(p)}) = (X_{k,i})_{n\times p}$, to be small. It has been established that the condition*

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

guarantees the exact recovery of a k-sparse signal β in the absence of noise, where $y = X\beta$, and enables the stable recovery of a sparse signal in the presence of noise, where $y = X\beta + z$. Here, z represents an error vector that may not necessarily be random.

Author Contributions: Writing–original draft, H.Z.; Writing–review & editing, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Natural Science Foundation of China under Grant [No. 11771178, 12171198]; the Science and Technology Development Program of Jilin Province under Grant [No. 20210101467JC]; and Fundamental Research Funds for the Central Universities.

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bai, Y.; Zhang, Y.; Liu, C. Moderate deviation principle for likelihood ratio test in multivariate linear regression model. *J. Multivar. Anal.* **2023**, *194*, 105139. [CrossRef]
2. Jiang, T. The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **2004**, *14*, 865–880. [CrossRef]
3. Zhou, W. Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Am. Math. Soc.* **2007**, *359*, 5345–5363. [CrossRef]
4. Liu, W.; Lin, Z.; Shao, Q. 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.* **2008**, *18*, 2337–2366. [CrossRef]
5. Li, D.; Rosalsky, A. Some strong limit theorems for the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **2006**, *16*, 423–447. [CrossRef]
6. Li, D.; Liu, W.; Rosalsky, A. Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. *Probab. Theory Relat. Fields* **2010**, *148*, 5–35. [CrossRef]
7. Li, D.; Qi, Y.; Rosalsky, A. On Jiang's asymptotic distribution of the largest entry of a sample correlation matrix. *J. Multivar. Anal.* **2012**, *111*, 256–270. [CrossRef]
8. Shao, Q.; Zhou, W. Necessary and sufficient conditions for the asymptotic distributions of coherence of ultra-high dimensional random matrices. *Ann. Probab.* **2014**, *42*, 623–648. [CrossRef]
9. Liu, W.; Lin, Z. Asymptotic distributions of the largest entries of sample correlation matrices under dependence assumptions. *Chin. Ann. Math. Ser.* **2008**, *29*, 543–556.
10. Zhao, H.; Zhang Y. The asymptotic distributions of the largest entries of sample correlation matrices under an α-mixing assumption. *Acta. Math. Sin.-Engl. Ser.* **2022**, *38*, 2039–2056. [CrossRef]
11. Lin, Z.; Lu, C. *Limit Theory on Mixing Dependent Random Variables*; Kluwer Academic Publishers: Dordrecht, The Netherland, 1997.
12. Bradley, R. Approximation theorems for strongly mixing random variables. *Mich. Math. J.* **1983**, *30*, 69–81. [CrossRef]
13. Kim, T. A note on moment bounds for strong mixing sequences. *Stat. Probab. Lett.* **1993**, *16*, 163–168. [CrossRef]
14. Sakhanenko, A. On the accuracy of normal approximation in the invariance principle. *Sib. Adv. Math.* **1991**, *1*, 58–91.
15. Li, D.; Rao, M.; Jiang, T.; Wang, X. Complete convergence and almost sure convergence of weighted sums of random variables. *J. Theoret. Probab.* **1995**, *8*, 49–76. [CrossRef]
16. Chow, Y.; Teicher, H. *Probability Theory, Independence, Interchangeability, Martingales*, 2nd ed.; Springer: New York, NY, USA, 1988.
17. Allan, G. *Probability: A Graduate Course*; Springer: New York, NY, USA, 2005.

18. Arratia, R.; Goldstein, L.; Gordon, L. Two moments suffice for Poisson approximations: The Chen-Stein method. *Ann. Probab.* **1989**, *17*, 9–25. [CrossRef]
19. Cai, T.; Jiang, T. Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *Ann. Stat.* **2011**, *39*, 1496–1525. [CrossRef]
20. Bai, Z.; Yin, Y. Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.* **1993**, *21*, 1275–1294. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.