Submitted to the Annals of Statistics

SUPPLEMENT TO "LIMITING LAWS OF COHERENCE OF RANDOM MATRICES WITH APPLICATIONS TO TESTING COVARIANCE STRUCTURE AND CONSTRUCTION OF COMPRESSED SENSING MATRICES"

By T. Tony Cai * and Tiefeng Jiang†

University of Pennsylvania and University of Minnesota

In this supplement we first give more details on Remarks 2.3 and 2.4. Then we prove Propositions 4.1 and 6.2 and verify the conclusions in the three examples given in Section 4, and finally we prove Lemmas 6.5-6.12 which are used in the proofs of the main results.

Details on Remark 2.3. Consider $\Sigma = I_p$ with $p = 2n$ and $\tau = n$. So conditions (i) and (iii) in Theorem 4 hold, but (ii) does not. Observe

$$\Big\{(i,j);\ 1 \le i < j \le 2n,\ |i-j| \ge n\Big\} = n + (n-1) + \cdots + 1 = \frac{n(n+1)}{2} \sim \frac{p^2}{8}$$

as $n \to \infty$. So $L_{n,\tau}$ is the maximum of roughly $p^2/8$ random variables, and the dependence of any two of such random variables are less than that appeared in L_n in Theorem 3. The result in Theorem 3 can be rewritten as

$$nL_n^2 - 2\log\frac{p^2}{2} + \log\log\frac{p^2}{2} - \log 8 \ \text{ converges weakly to } F$$

as $n \to \infty$. Recalling L_n is the maximum of roughly $p^2/2$ weakly dependent random variables, replace L_n with $L_{n,\tau}$ and $p^2/2$ with $p^2/8$ to have $nL_{n,\tau}^2 - 2\log\frac{p^2}{8} + \log\log\frac{p^2}{8} - \log 8$ converges weakly to F, where F is as in Theorem 3. That is,

$$(nL_{n,\tau}^2 - 4\log p + \log\log p) + \log 16 \ \text{ converges weakly to } F \tag{76}$$

as $n \to \infty$ (This can be done rigorously by following the proof of Theorem 3). The difference between (76) and Theorem 4 is evident.

---

*The research of Tony Cai was supported in part by NSF FRG Grant DMS-0854973.

†The research of Tiefeng Jiang was supported in part by NSF FRG Grant DMS-0449365.

Details on Remark 2.4. Let $p = mn$ with integer $m \geq 2$. We consider the $p \times p$ matrix $\Sigma = \text{diag}\,(H_n, \cdots, H_n)$ where there are m H_n's in the diagonal of Σ and all of the entries of the $n \times n$ matrix H_n are equal to 1. Thus, if $(\zeta_1, \cdots, \zeta_p) \sim N_p(0, \Sigma)$, then $\zeta_{ln+1} = \zeta_{ln+2} = \cdots = \zeta_{(l+1)n}$ for all $0 \leq l \leq m-1$ and $\zeta_1, \zeta_{n+1}, \cdots, \zeta_{(m-1)n+1}$ are i.i.d. $N(0,1)$-distributed random variables. Let $\{\zeta_{ij};\, 1 \leq i \leq n, 1 \leq j \leq m\}$ be i.i.d. $N(0,1)$-distributed random variables. Then

$$(\underbrace{\zeta_{i1}, \cdots, \zeta_{i1}}_{n}, \underbrace{\zeta_{i2}, \cdots, \zeta_{i2}}_{n}, \cdots, \underbrace{\zeta_{i\,m}, \cdots, \zeta_{i\,m}}_{n})' \in \mathbb{R}^p,\ 1 \leq i \leq n,$$

are i.i.d. random vectors with distribution $N_p(0, \Sigma)$. Denote the corresponding data matrix by $(x_{ij})_{n\times p}$. Now, take $\tau = n$ and $m = [e^{n^{1/4}}]$. Notice $\Gamma_{p,\delta} = p$ for any $\delta > 0$. Since $p = mn$, both (i) and (ii) in Theorem 4 are satisfied, but (iii) does not. Obviously,

$$L_{n,\tau} = \max_{1\leq i<j\leq p,\, |i-j|\geq\tau} |\rho_{ij}| = \max_{1\leq i<j\leq m} |\hat{\rho}_{ij}|,$$

where $\hat{\rho}_{ij}$ is obtained from $(\zeta_{ij})_{n\times m}$ as in (1) (note that the mn entries of $(\zeta_{ij})_{n\times m}$ are i.i.d. with distribution $N(0,1)$). By Theorem 3 on $\max_{1\leq i<j\leq m} |\hat{\rho}_{ij}|$, we have that $nL_{n,\tau}^2 - 4\log m + \log\log m$ converges weakly to F, which is the same as the F in Theorem 4. Set $\log_2 x = \log\log x$ for $x > 1$. Notice

$$\begin{aligned} nL_{n,\tau}^2 - 4\log m + \log_2 m &= nL_{n,\tau}^2 - 4\log p + 4\log n + \log_2 m \\ &= (nL_{n,\tau}^2 - 4\log p + \log_2 p) + 4\log n + o(1) \end{aligned}$$

since $p = mn$ and $\log_2 p - \log_2 m \to 0$. Further, it is easy to check that $4\log n - 16\log_2 p \to 0$. Therefore, the previous conclusion is equivalent to that

$$(nL_{n,\tau}^2 - 4\log p + \log\log p) + 16\log\log p \text{ converges weakly to } F \tag{77}$$

as $n \to \infty$. This is different from the conclusion of Theorem 4.

Proof of Proposition 4.1. Recall the definition of $\tilde{L}_n$ in (3), to prove the conclusion, w.l.o.g., we assume $\mu = 0$ and $\sigma^2 = 1$. Evidently, by the i.i.d. assumption,

$$\begin{aligned} P(\tilde{L}_n \geq t) &\leq \frac{p^2}{2} P\Big(\frac{|x_1'x_2|}{\|x_1\|\cdot\|x_2\|} \geq t\Big) \\ &\leq \frac{p^2}{2} P\Big(\frac{|x_1'x_2|}{n} \geq \frac{t}{2}\Big) + \frac{p^2}{2}\cdot 2P\left(\frac{\|x_1\|^2}{n} \leq \frac{1}{2}\right) \qquad (78) \end{aligned}$$

where the event $\{\|x_{11}\|^2/n > 1/2,\ \|x_{12}\|^2/n > 1/2\}$ and its complement are used to get the last inequality. Since $\{x_{ij};\ i \geq 1,\ j \geq 1\}$ are i.i.d., the condition $Ee^{t_0|x_{11}|^2} < \infty$ implies $Ee^{t_0'|x_{11}x_{12}|} < \infty$ for some $t_0' > 0$. By the Chernoff bound (see, e.g., p. 27 from Dembo and Zeitouni (1998)) and noting that $E(x_{11}x_{12}) = 0$ and $Ex_{11}^2 = 1$, we have

$$P\Big(\frac{|x_1'x_2|}{n} \geq \frac{t}{2}\Big) \leq 2e^{-nI_1(t/2)} \text{ and } P\left(\frac{\|x_1\|^2}{n} \leq \frac{1}{2}\right) \leq 2e^{-nI_2(1/2)}$$

for any $n \geq 1$ and $t > 0$, where the following facts about rate functions $I_1(x)$ and $I_2(y)$ are used:

(i) $I_1(x) = 0$ if and only if $x = 0$; $I_2(y) = 0$ if and only if $y = 1$;

(ii) $I_1(x)$ is non-decreasing on $A := [0, \infty)$ and non-increasing on A^c. This is also true for $I_2(y)$ with $A = [1, \infty)$.

These and (78) conclude

$$P(\tilde{L}_n \geq t) \leq p^2e^{-nI_1(t/2)} + 2p^2e^{-nI_2(1/2)} \leq 3p^2e^{-ng(t)}$$

where $g(t) = \min\{I_1(t/2),\ I_2(1/2)\}$ for any $t > 0$. Obviously, $g(t) > 0$ for any $t > 0$ from (i) and (ii) above. ■

Proof of Proposition 6.2. We prove the proposition by following the outline of the proof of Proposition 6.1 step by step. It suffices to show

$$\lim_{n\to\infty} P\Big(\frac{W_n}{\sqrt{n\log p}} \geq 2 + 2\epsilon\Big) = 0 \quad \text{and} \tag{79}$$

$$\lim_{n\to\infty} P\Big(\frac{W_n}{\sqrt{n\log p}} \leq 2 - \epsilon\Big) = 0 \tag{80}$$

for any $\epsilon > 0$ small enough. Note that $|x_{11}x_{12}|^{\varrho} = |x_{11}|^{\varrho} \cdot |x_{12}|^{\varrho} \leq |x_{11}|^{2\varrho} + |x_{12}|^{2\varrho}$ for any $\varrho > 0$. From the given moment condition, we see that $E\exp\big(t_0|x_{11}|^{4\beta/(1-\beta)}\big) < \infty$. This implies that $E\exp\big(|x_{11}|^{\frac{4\beta}{1+\beta}}\big) < \infty$ and $E\exp\Big(|x_{11}x_{12}|^{\frac{2\beta}{1+\beta}}\Big) < \infty$. By (i) of Lemma 6.4, (28) holds for $\{p_n\}$ such that $p_n \to \infty$ and $\log p_n = o(n^{\beta})$. By using (27) and (29), we obtain (79).

By using condition $E\exp\{t_0|x_{11}|^{\frac{4\beta}{1+\beta}}\} < \infty$ again, we know (33) also holds for $\{p_n\}$ such that $p_n \to \infty$ and $\log p_n = o(n^{\beta})$. Then all statements after (30) and before (36) hold. Now, by Lemma 6.7, (37) holds for $\{p_n\}$ such that $p_n \to \infty$ and $\log p_n = o(n^{\beta})$, we then have (38). This implies (30), which is the same as (80). ■

To verify the assertions stated in the three examples in Section 4, we need the following lemma.

Lemma 0.1 *Let Z be a random variable with $EZ=0$, $EZ^2=1$ and $Ee^{t_0|Z|}<\infty$ for some $t_0>0$. Choose $\alpha>0$ such that $E(Z^2e^{\alpha|Z|})\leq 3/2$. Set $I(x)=\sup_{t\in\mathbb{R}}\{tx-\log Ee^{tZ}\}$. Then $I(x)\geq x^2/3$ for all $0\leq x\leq 3\alpha/2$.*

Proof. By the Taylor expansion, for any $x\in\mathbb{R}$, $e^x=1+x+\frac{x^2}{2}e^{\theta x}$ for some $\theta\in[0,1]$. It follows from $EZ=0$ that

$$Ee^{tZ}=1+\frac{t^2}{2}E(Z^2e^{\theta tZ})\leq 1+\frac{t^2}{2}E(Z^2e^{t|Z|})\leq 1+\frac{3}{4}t^2$$

for all $0\leq t\leq\alpha$. Use the inequality $\log(1+x)\leq x$ for all $x>-1$ to see that $\log Ee^{tZ}\leq 3t^2/4$ for every $0\leq t\leq\alpha$. Take $t_0=2x/3$ with $x>0$. Then $0\leq t_0\leq\alpha$ for all $0\leq x\leq 3\alpha/2$. It follows that

$$I(x)\geq t_0x-\frac{3}{4}t_0^2=\frac{x^2}{3}. \qquad \blacksquare$$

Verifications of Examples 1, 2, and 3 in Section 4. We consider the three examples one by one.
(i) If $x_{11}\sim N(0,n^{-1})$ as in (19), then ξ and η are i.i.d. with distribution $N(0,1)$. By Lemma 3.2 from Jiang (2005), $I_2(x)=(x-1-\log x)/2$ for $x>0$. So $I_2(1/2)>1/12$. Also, since $Ee^{\theta\xi\eta}=Ee^{\theta^2\xi^2/2}=(1-\theta^2)^{-1/2}$ for $|\theta|<1$. It is straightforward to get

$$I_1(x)=\frac{\sqrt{4x^2+1}-1}{2}-\frac{1}{2}\log\frac{\sqrt{4x^2+1}+1}{2},\quad x>0.$$

Let $y=\frac{\sqrt{4x^2+1}-1}{2}$. Then $y>2x^2/3$ for all $|x|\leq 4/5$. Thus, $I_1(x)=y-\frac{1}{2}\log(1+y)>\frac{y}{2}>\frac{x^2}{3}$ for $|x|\leq 4/5$. Therefore, $g(t)\geq\min\{I_1(\frac{t}{2}),\frac{1}{12}\}\geq\min\{\frac{t^2}{12},\frac{1}{12}\}=\frac{t^2}{12}$ for $|t|\leq 1$. Since $1/(2k-1)\leq 1$ if $k\geq 1$. By Proposition 4.1, we have

$$P\Big((2k-1)\tilde{L}_n<1\Big)\geq 1-3p^2\exp\Big\{-\frac{n}{12(2k-1)^2}\Big\} \tag{81}$$

for all $n\geq 2$ and $k\geq 1$, which is (23).
(ii) Let x_{11} be such that $P(x_{11}=\pm 1/\sqrt{n})=1/2$ as in (20). Then ξ and η in Proposition 4.1 are i.i.d. with $P(\xi=\pm 1)=1/2$. Hence, $P(\xi\eta=\pm 1)=1/2$ and $\xi^2=1$. Immediately, $I_2(1)=0$ and $I_2(x)=+\infty$ for all $x\neq 1$. If $\alpha=\log\frac{3}{2}\sim 0.405$, then $E(Z^2e^{\alpha|Z|})=e^{\alpha}\leq\frac{3}{2}$ with $Z=\xi\eta$. Thus, by Lemma 0.1, $I_1(x)\geq x^2/3$ for all $0\leq x\leq\frac{3}{5}\leq\frac{3\alpha}{2}$. Therefore, $g(t)\geq\frac{t^2}{12}$ for $0\leq t\leq\frac{6}{5}$. This gives that

$$P\Big((2k-1)\tilde{L}_n<1\Big)\geq 1-3p^2\exp\Big\{-\frac{n}{12(2k-1)^2}\Big\} \tag{82}$$

provided $\frac{1}{2k-1} \le \frac{6}{5}$, that is, $k \ge \frac{11}{12}$. We then obtain (23) for all $n \ge 2$ and $k \ge 1$.

(iii) Let x_{11} be such that $P(x_{11} = \pm\sqrt{3/n}) = 1/6$ and $P(x_{11} = 0) = 2/3$ as in (21). Then ξ and η in Proposition 4.1 are i.i.d. with $P(\xi = \pm\sqrt{3}) = 1/6$ and $P(\xi = 0) = 2/3$. It follows that $P(Z = \pm 3) = 1/18$ and $P(Z = 0) = 8/9$ with $Z = \xi\eta$. Take $\alpha = \frac{1}{3}\log\frac{3}{2} > 0.13$. Then $E(Z^2 e^{\alpha|Z|}) = \frac{2\times 9}{18}e^{3\alpha} = \frac{3}{2}$. Thus, by Lemma 0.1, $I_1(x) \ge x^2/3$ for all $0 \le x \le \frac{3\alpha}{2} = \frac{1}{2}\log\frac{3}{2} \sim 0.2027$. Now, $P(\xi^2 = 3) = \frac{1}{3} = 1 - P(\xi^2 = 0)$. Hence, $\xi^2/3 \sim Ber(p)$ with $p = \frac{1}{3}$. It follows that

$$\begin{aligned} I_2(x) &= \sup_{\theta\in\mathbb{R}}\Big\{(3\theta)\frac{x}{3} - \log Ee^{3\theta(\xi^2/3)}\Big\} \\ &= \Lambda^*\Big(\frac{x}{3}\Big) = \frac{x}{3}\log x + \Big(1 - \frac{x}{3}\Big)\log\frac{3-x}{2} \end{aligned}$$

for $0 \le x \le 3$ by (b) of Exercise 2.2.23 from Dembo and Zeitouni (1998). Thus, $I_2(\frac{1}{2}) = \frac{1}{6}\log\frac{1}{2} + \frac{5}{6}\log\frac{5}{4} \sim 0.0704 > \frac{1}{15}$. Now, for $0 \le t \le \frac{2}{5}$, we have

$$g(t) = \min\Big\{I_1(\frac{t}{2}),\, I_2(\frac{1}{2})\Big\} \ge \min\Big\{\frac{t^2}{12},\, \frac{1}{15}\Big\} = \frac{t^2}{12}.$$

Easily, $t := \frac{1}{2k-1} \le \frac{2}{5}$ if and only if $k \ge \frac{7}{4}$. Thus, by Proposition 4.1,

$$P\Big((2k-1)\tilde{L}_n < 1\Big) \ge 1 - 3p^2\exp\Big\{-\frac{n}{12(2k-1)^2}\Big\} \tag{83}$$

for all $n \ge 2$ and $k \ge \frac{7}{4}$. We finally conclude (23) for all $n \ge 2$ and $k \ge 2$. ■

Proof of Lemma 6.5. (i) First, since x_{ij}'s are i.i.d. bounded random variables with mean zero and variance one, by (i) of Lemma 6.4,

$$\begin{aligned} P(\sqrt{n/\log p}\, b_{n,4} \ge K) &= P\Big(\max_{1\le i\le p}\Big|\frac{1}{\sqrt{n\log p}}\sum_{k=1}^n x_{ki}\Big| \ge K\Big) \qquad (84) \\ &\le p\cdot P\Big(\Big|\frac{1}{\sqrt{n\log p}}\sum_{k=1}^n x_{k1}\Big| \ge K\Big) \\ &\le p\cdot e^{-(K^2/3)\log p} = \frac{1}{p^{K^2/3-1}} \to 0 \qquad (85) \end{aligned}$$

as $n \to \infty$ for any $K > \sqrt{3}$. This says that $\{\sqrt{n/\log p}\, b_{n,4}\}$ are tight.

Second, noticing that $|t-1| \le |t^2-1|$ for any $t>0$ and $nh_i^2 = \|x_i - \bar{x}_i\|^2 = x_i^T x_i - n|\bar{x}_i|^2$, we get that

$$b_{n,1} \le \max_{1\le i\le p} |h_i^2 - 1| \le \max_{1\le i\le p}\left|\frac{1}{n}\sum_{k=1}^n (x_{ki}^2-1)\right| + \max_{1\le i\le p}\left|\frac{1}{n}\sum_{k=1}^n x_{ki}\right|^2$$
$$= Z_n + b_{n,4}^2 \tag{86}$$

where $Z_n = \max_{1\le i\le p}\left|\frac{1}{n}\sum_{k=1}^n (x_{ki}^2-1)\right|$. Therefore,

$$\sqrt{\frac{n}{\log p}}\,b_{n,1} \le \sqrt{\frac{n}{\log p}}\,Z_n + \sqrt{\frac{\log p}{n}}\cdot\Big(\sqrt{\frac{n}{\log p}}\,b_{n,4}\Big)^2. \tag{87}$$

Replacing "x_{ki}" in (84) with "x_{ki}^2-1" and using the same argument, we obtain that $\{\sqrt{n/\log p}\,Z_n\}$ are tight. Since $\log p = o(n)$ and $\{\sqrt{n/\log p}\,b_{n,4}\}$ are tight, using (25) we know the second term on the right hand side of (87) goes to zero in probability as $n\to\infty$. Hence, we conclude from (87) that $\{\sqrt{n/\log p}\,b_{n,1}\}$ are tight.

Finally, since $\log p = o(n)$ and $\{\sqrt{n/\log p}\,b_{n,1}\}$ are tight, use (25) to have $b_{n,1}\to 0$ in probability as $n\to\infty$. This implies that $b_{n,3}\to 1$ in probability as $n\to\infty$.

(ii) We first claim that

$$\begin{aligned}&b_{n,3}\xrightarrow{P} 1 \text{ as } n\to\infty, \text{ and } \{\sqrt{n/\log p}\,b_{n,1}\} \text{ and } \{\sqrt{n/\log p}\,b_{n,4}\}\\ &\text{are tight if } Ee^{t_0|x_{11}|^\alpha}<\infty \text{ for some } 0<\alpha\le 2 \text{ and } t_0>0, \text{ and}\\ &p_n\to\infty \text{ and } \log p_n = o(n^{\beta_1}) \text{ as } n\to\infty, \text{ where } \beta_1 = \alpha/(4-\alpha).\end{aligned} \tag{88}$$

If the claim holds and $0<\alpha\le 2$, recalling $\beta = \alpha/(4+\alpha) < \alpha/(4-\alpha) = \beta_1$, then $\log p_n = o(n^\beta) = o(n^{\beta_1})$ as $n\to\infty$, the desired conclusions follow.

If claim (88) holds and $\alpha>2$, then $Ee^{t_0|x_{11}|^2}<\infty$. It follows that $\{\sqrt{n/\log p}\,b_{n,1}\}$ and $\{\sqrt{n/\log p}\,b_{n,4}\}$ are all tight with $\log p_n = o(n)$. Noticing $\beta = \frac{\alpha}{4+\alpha} < 1$, we see that $\{\sqrt{n/\log p}\,b_{n,1}\}$ and $\{\sqrt{n/\log p}\,b_{n,4}\}$ are all tight with $\log p_n = o(n^\beta)$. We also have that $b_{n,3}\to 1$ in probability as $n\to\infty$ by the same argument as in the last paragraph of the proof of (i) above. Now we turn to prove claim (88).

By (85) and (87), to prove claim (88), it is enough to show, for some constant $K>0$,

$$p\cdot P\Big(\Big|\frac{1}{\sqrt{n\log p}}\sum_{k=1}^n x_{k1}\Big|\ge K\Big)\to 0 \quad\text{and} \tag{89}$$

$$p\cdot P\Big(\Big|\frac{1}{\sqrt{n\log p}}\sum_{k=1}^n (x_{k1}^2-1)\Big|\ge K\Big)\to 0 \tag{90}$$

as $n \to \infty$. Using $a_n := \sqrt{\log p_n} = o(n^{\beta/2})$ and (i) of Lemma 6.4, we have

$$P\Big(\Big|\frac{1}{\sqrt{n\log p}}\sum_{k=1}^{n} x_{k1}\Big| \geq K\Big) \leq \frac{1}{p^{K^2/3}} \quad \text{and}$$
$$P\Big(\Big|\frac{1}{\sqrt{n\log p}}\sum_{k=1}^{n} (x_{k1}^2-1)\Big| \geq K\Big) \leq \frac{1}{p^{K^2/3}}$$

as n is sufficiently large, where the first inequality holds provided

$$E\exp\left(t_0|x_{11}|^{2\beta/(1+\beta)}\right) = E\exp(t_0|x_{11}|^{\alpha/2}) < \infty;$$

the second holds since $E\exp\left(t_0|x_{11}^2-1|^{2\beta/(1+\beta)}\right) = E\exp(t_0|x_{11}^2-1|^{\alpha/2}) < \infty$ for some $t_0 > 0$, which is equivalent to $Ee^{t_0'|x_{11}|^{\alpha}} < \infty$ for some $t_0' > 0$. We then get (89) and (90) by taking $K = 2$. ■

Proof of Lemma 6.6. Let $G_n = \{|\sum_{k=1}^n x_{k1}^2/n - 1| < \delta\}$. Then, by the Chernoff bound (see, e.g., p. 27 from Dembo and Zeitouni (1998)), for any $\delta \in (0,1)$, there exists a constant $C_\delta > 0$ such that $P(G_n^c) \leq 2e^{-nC_\delta}$ for all $n \geq 1$. Set $a_n = t_n\sqrt{n\log p}$. Then

$$\Psi_n \leq E\Big\{P^1\Big(|\sum_{k=1}^n x_{k1}x_{k2}| > a_n\Big)^2 I_{G_n}\Big\} + 2e^{-nC_\delta} \tag{91}$$

for all $n \geq 1$. Evidently, $|x_{k1}x_{k2}| \leq C^2$, $E^1(x_{k1}x_{k2}) = 0$ and $E^1(x_{k1}x_{k2})^2 = x_{k1}^2$, where E^1 stands for the conditional expectation given $\{x_{k1},\ 1 \leq k \leq n\}$. By the Bernstein inequality (see, e.g., p.111 from Chow and Teicher (1997)),

$$\begin{aligned} P^1\Big(|\sum_{k=1}^n x_{k1}x_{k2}| > a_n\Big)^2 I_{G_n} &\leq 4\cdot\exp\Big\{-\frac{a_n^2}{(\sum_{k=1}^n x_{k1}^2 + C^2a_n)}\Big\} I_{G_n} \\ &\leq 4\cdot\exp\Big\{-\frac{a_n^2}{((1+\delta)n + C^2a_n)}\Big\} \\ &\leq \frac{1}{p^{t^2/(1+2\delta)}} \end{aligned} \tag{92}$$

as n is sufficiently large, since $a_n^2/(n(1+\delta)+C^2a_n) \sim t^2(\log p)/(1+\delta)$ as $n \to \infty$. Recalling (91), the conclusion then follows by taking δ small enough. ■

Proof of Lemma 6.7. Let P^2 stand for the conditional probability given $\{x_{k2},\ 1 \leq k \leq n\}$. Since $\{x_{ij};\ i \geq 1,\ j \geq 1\}$ are i.i.d., to prove the lemma, it is enough to prove

$$\Psi_n := E\Big\{P^2\Big(|\sum_{k=1}^{n} x_{k1}x_{k2}| > t_n\sqrt{n\log p}\Big)^2\Big\} = O\left(\frac{1}{p^{t^2-\epsilon}}\right) \tag{93}$$

as $n \to \infty$. Here we use the notation "P^2" instead of "P^1" simply because of the convenience of notation.

Step 1. For any $x > 0$, by the Markov inequality

$$P(\max_{1\le k\le n} |x_{k2}| \ge x) \le nP(|x_{12}| \ge x) \le Cne^{-t_0x^{\alpha}} \tag{94}$$

where $C = Ee^{t_0|x_{11}|^{\alpha}} < \infty$. Second, we know that $Ee^{t|x_{11}|^{4\beta/(1+\beta)}} < \infty$ for any $t > 0$ from the given condition. For any $\epsilon > 0$, by (ii) of Lemma 6.4, there exists a constant $C = C_{\epsilon} > 0$ such that

$$P\Big(\frac{|\sum_{k=1}^{n} x_{k2}^2 - n|}{n^{(\beta+1)/2}} \ge \epsilon\Big) \le e^{-C_{\epsilon}n^{\beta}} \tag{95}$$

for each $n \ge 1$.

Set $h_n = n^{(1-\beta)/4}$, $\mu_n = Ex_{ij}I(|x_{ij}| \le h_n)$,

$$\begin{aligned} y_{ij} &= x_{ij}I(|x_{ij}| \le h_n) - Ex_{ij}I(|x_{ij}| \le h_n) \\ z_{ij} &= x_{ij}I(|x_{ij}| > h_n) - Ex_{ij}I(|x_{ij}| > h_n) \end{aligned} \tag{96}$$

for all $i \ge 1$ and $j \ge 1$. Then, $x_{ij} = y_{ij} + z_{ij}$ for all $i, j \ge 1$. Use the inequality $P(U + V \ge u + v) \le P(U \ge u) + P(V \ge v)$ to obtain

$$\begin{aligned} & P^2\Big(|\sum_{k=1}^{n} x_{k1}x_{k2}| > t_n\sqrt{n\log p}\Big)^2 \\ \le\ & 2P^2\Big(|\sum_{k=1}^{n} y_{k1}x_{k2}| > (t_n-\delta)\sqrt{n\log p}\Big)^2 \\ & + 2P^2\Big(|\sum_{k=1}^{n} z_{k1}x_{k2}| > \delta\sqrt{n\log p}\Big)^2 := 2A_n + 2B_n \end{aligned} \tag{97}$$

for any $\delta > 0$ small enough. Hence,

$$\Psi_n \le 2EA_n + 2EB_n \tag{98}$$

for all $n \ge 2$.

Step 2: the bound of A_n. Now, if $\max_{1\le k\le n}|x_{k2}|\le h_n$, then $|y_{k1}x_{k2}|\le 2h_n^2$ for all $k\ge 1$. It then follows from the Bernstein inequality (see, e.g., p. 111 from Chow and Teicher (1997)) that

$$\begin{aligned} A_n &= P^2\Big(|\sum_{k=1}^n y_{k1}x_{k2}| > (t_n-\delta)\sqrt{n\log p}\Big)^2 \\ &\le 4\cdot\exp\Big\{-\frac{(t_n-\delta)^2 n\log p}{E(y_{11}^2)\sum_{k=1}^n x_{k2}^2 + 2h_n^2(t_n-\delta)\sqrt{n\log p}}\Big\} \\ &\le 4\cdot\exp\Big\{-\frac{(t_n-\delta)^2 n\log p}{E(y_{11}^2)(n+\epsilon n^{(\beta+1)/2}) + 2h_n^2(t_n-\delta)\sqrt{n\log p}}\Big\} \end{aligned}$$

for $0<\delta<t_n$ and $\frac{|\sum_{k=1}^n x_{k2}^2 - n|}{n^{(\beta+1)/2}}<\epsilon$. Notice $E(y_{11}^2)\to 1$ and $2h_n^2(t_n-\delta)\sqrt{n\log p}/3 = o(n)$ as $n\to\infty$. Thus,

$$\frac{(t_n-\delta)^2 n\log p}{E(y_{11}^2)(n+\epsilon n^{(\beta+1)/2}) + 2h_n^2(t_n-\delta)\sqrt{n\log p}} \sim (t-\delta)^2\log p$$

as $n\to\infty$. In summary, if $\max_{1\le k\le n}|x_{k2}|\le h_n$ and $\frac{|\sum_{k=1}^n x_{k2}^2 - n|}{n^{(\beta+1)/2}}\le\epsilon$, then for any $\delta\in(0,t/2)$,

$$A_n \le \frac{1}{p^{t^2-2t\delta}} \tag{99}$$

as n is sufficiently large. Therefore, for any $\epsilon>0$ small enough, take δ sufficiently small to obtain

$$\begin{aligned} EA_n &= E\Big\{P^2\Big(|\sum_{k=1}^n y_{k1}x_{k2}| > (t_n-\delta)\sqrt{n\log p}\Big)^2\Big\} \\ &\le \frac{1}{p^{t^2-\epsilon}} + P(\max_{1\le k\le n}|x_{k2}|\ge h_n) + P\Big(\frac{|\sum_{k=1}^n x_{k2}^2 - n|}{n^{(\beta+1)/2}}\ge\epsilon\Big) \\ &\le \frac{1}{p^{t^2-\epsilon}} + Cne^{-h_n^\alpha} + e^{-C_\epsilon n^\beta} = O\left(\frac{1}{p^{t^2-\epsilon}}\right) \qquad (100) \end{aligned}$$

as $n\to\infty$, where the second inequality follows from (94) and (95), and the last identity follows from the fact that $h_n^\alpha = n^\beta$ and the assumption $\log p = o(n^\beta)$.

Step 3: the bound of B_n. Recalling the definition of z_{ij} and μ_n in (96), we have

$$
\begin{aligned}
\sqrt{B_n} &= P^2\Big(|\sum_{k=1}^n z_{k1}x_{k2}| > \delta\sqrt{n\log p}\Big)\\
&\leq P^2\Big(|\sum_{k=1}^n x_{k1}x_{k2}I\{|x_{k1}|>h_n\}| > \delta\sqrt{n\log p}/2\Big)\\
&\quad +I\Big(|\sum_{k=1}^n x_{k2}| > \frac{\delta\sqrt{n\log p}}{2(e^{-n}+|\mu_n|)}\Big) := C_n + D_n. \qquad (101)
\end{aligned}
$$

Now, by (94),

$$
C_n \leq P(\max_{1\leq k\leq n}|x_{k1}|>h_n) \leq Cne^{-t_0h_n^{\alpha}} = Cne^{-t_0n^{\beta}}. \qquad (102)
$$

Easily, $|\mu_n| \leq E|x_{11}|I(|x_{11}|>h_n) \leq e^{-t_0h_n^{\alpha}/2}E(|x_{11}|e^{t_0|x_{11}|^{\alpha}/2}) = Ce^{-t_0n^{\beta}/2}$. Also, $P(|\sum_{k=1}^n\eta_k| \geq x) \leq \sum_{k=1}^n P(|\eta_k|\geq x/n)$ for any random variables $\{\eta_i\}$ and $x>0$. We then have

$$
\begin{aligned}
ED_n &= P\Big(|\sum_{k=1}^n x_{k2}| > \frac{\delta\sqrt{n\log p}}{2(e^{-n}+|\mu_n|)}\Big)\\
&\leq nP\Big(|x_{11}| > \frac{\delta\sqrt{n\log p}}{2n(e^{-n}+|\mu_n|)}\Big)\\
&\leq nP\Big(|x_{11}| > e^{t_0n^{\beta}/3}\Big) \leq e^{-n} \qquad (103)
\end{aligned}
$$

as n is sufficiently large, where the last inequality is from condition $Ee^{t_0|x_{11}|^{\alpha}} < \infty$. Consequently,

$$
EB_n \leq 2E(C_n^2) + 2E(D_n^2) = 2E(C_n^2) + 2E(D_n) \leq e^{-Cn^{\beta}} \qquad (104)
$$

as n is sufficiently large. This joint with (98) and (100) yields (93). ■

Proof of Lemma 6.8. Take $\gamma = (1-\beta)/2 \in [1/3, 1/2)$. Set

$$
\eta_i = \xi_i I(|\xi_i|\leq n^{\gamma}),\ \mu_n = E\eta_1 \ \text{ and }\ \sigma_n^2 = Var(\eta_1),\quad 1\leq i\leq n. \qquad (105)
$$

Since the desired result is a conclusion about $n\to\infty$, without loss of generality, assume $\sigma_n>0$ for all $n\geq 1$. We first claim that there exists a constant $C>0$ such that

$$
\max\Big\{|\mu_n|,\ |\sigma_n-1|,\ P(|\xi_1|>n^{\gamma})\Big\} \leq Ce^{-n^{\beta}/C} \qquad (106)
$$

for all $n \geq 1$. In fact, since $E\xi_1 = 0$ and $\alpha\gamma = \beta$,

$$\begin{aligned} |\mu_n| = |E\xi_1 I(|\xi_1| > n^\gamma)| &\leq E|\xi_1| I(|\xi_1| > n^\gamma) \\ &\leq E\Big(|\xi_1| e^{t_0|\xi_1|^\alpha/2}\Big) \cdot e^{-t_0 n^\beta/2} \end{aligned} \tag{107}$$

for all $n \geq 1$. Note that $|\sigma_n - 1| \leq |\sigma_n^2 - 1| = \mu_n^2 + E\xi_1^2 I(|\xi_1| > n^\gamma)$, by the same argument as in (107), we know both $|\sigma_n - 1|$ and $P(|\xi_1| > n^\gamma)$ are bounded by $Ce^{-n^\beta/C}$ for some $C > 0$. Then (106) follows.
Step 1. We prove that, for some constant $C > 0$,

$$\Big|P\Big(\frac{S_n}{\sqrt{n\log p_n}} \geq y_n\Big) - P\Big(\frac{\sum_{i=1}^n \eta_i}{\sqrt{n\log p_n}} \geq y_n\Big)\Big| \leq 2e^{-n^\beta/C} \tag{108}$$

for all $n \geq 1$. Observe

$$\xi_i \equiv \eta_i \quad \text{for} \quad 1 \leq i \leq n \quad \text{if} \quad \max_{1\leq i\leq n} |\xi_i| \leq n^\gamma. \tag{109}$$

Then, by (106),

$$\begin{aligned} P\Big(\frac{S_n}{\sqrt{n\log p_n}} \geq y_n\Big) &\leq P\Big(\frac{S_n}{\sqrt{n\log p_n}} \geq y_n, \max_{1\leq i\leq n}|\xi_i| \leq n^\gamma\Big) + P\Big(\bigcup_{i=1}^n \{|\xi_i| > n^\gamma\}\Big) \\ &\leq P\Big(\frac{\sum_{i=1}^n \eta_i}{\sqrt{n\log p_n}} \geq y_n\Big) + Cne^{-n^\beta/C} \end{aligned} \tag{110}$$

for all $n \geq 1$. Use inequality that $P(AB) \geq P(A) - P(B^c)$ for any events A and B to have

$$\begin{aligned} P\Big(\frac{S_n}{\sqrt{n\log p_n}} \geq y_n\Big) &\geq P\Big(\frac{S_n}{\sqrt{n\log p_n}} \geq y_n, \max_{1\leq i\leq n}|\xi_i| \leq n^\gamma\Big) \\ &= P\Big(\frac{\sum_{i=1}^n \eta_i}{\sqrt{n\log p_n}} \geq y_n, \max_{1\leq i\leq n}|\xi_i| \leq n^\gamma\Big) \\ &\geq P\Big(\frac{\sum_{i=1}^n \eta_i}{\sqrt{n\log p_n}} \geq y_n\Big) - Cne^{-n^\beta/C} \end{aligned}$$

where in the last step the inequality $P(\max_{1\leq i\leq n}|\xi_i| > n^\gamma) \leq Cne^{-n^\beta/C}$ is used as in (110). This and (110) concludes (108).
Step 2. Now we prove

$$P\Big(\frac{\sum_{i=1}^n \eta_i}{\sqrt{n\log p_n}} \geq y_n\Big) \sim \frac{e^{-x_n^2/2}}{\sqrt{2\pi}x_n} \tag{111}$$

as $n \to \infty$, where

$$x_n = y_n' \sqrt{\log p_n} \ \text{ and } \ y_n' = \frac{1}{\sigma_n}\left(y_n - \sqrt{\frac{n}{\log p_n}}\,\mu_n\right). \tag{112}$$

First, by (106),

$$|y_n' - y_n| \le \frac{|1-\sigma_n|}{\sigma_n} y_n + \frac{1}{\sigma_n}\cdot\sqrt{\frac{n}{\log p_n}}\,|\mu_n| \le Ce^{-n^{\beta}/C} \tag{113}$$

for all $n \ge 1$ since both σ_n and y_n have limits and $p_n \to \infty$. In particular, since $\log p_n = o(n^{\beta})$,

$$x_n = o(n^{\beta/2}) \tag{114}$$

as $n \to \infty$. Now, set

$$\eta_i' = \frac{\eta_i - \mu_n}{\sigma_n}$$

for $1 \le i \le n$. Easily

$$P\Big(\frac{\sum_{i=1}^n \eta_i}{\sqrt{n\log p_n}} \ge y_n\Big) = P\Big(\frac{\sum_{i=1}^n \eta_i'}{\sqrt{n\log p_n}} \ge y_n'\Big) \tag{115}$$

for all $n \ge 1$. Reviewing (105), for some constant $K > 0$, we have $|\eta_i'| \le Kn^{\gamma}$ for $1 \le i \le n$. Take $c_n = Kn^{\gamma - 1/2}$. Recalling x_n in (112). It is easy to check that

$$s_n := \Big(\sum_{i=1}^n E\eta_i'^2\Big)^{1/2} = \sqrt{n},\ \ \varrho_n := \sum_{i=1}^n E|\eta_i'|^3 \sim nC, \quad |\eta_i'| \le c_n s_n \text{ and}$$
$$0 < c_n \le 1$$

as n is sufficiently large. Recall $\gamma = (1-\beta)/2$, it is easy to see from (114) that

$$0 < x_n < \frac{1}{18c_n}$$

for n large enough. Now, let $\gamma(x)$ be as in Lemma 6.3, since $\beta \le 1/3$, by the lemma and (114),

$$\Big|\gamma\big(\frac{x_n}{s_n}\big)\Big| \le \frac{2x_n^3\varrho_n}{s_n^3} = o\left(n^{\frac{3\beta}{2}-\frac{1}{2}}\right) \to 0 \ \text{ and } \ \frac{(1+x_n)\varrho_n}{s_n^3} = O(n^{(\beta-1)/2}) \to 0$$

as $n \to \infty$. By (112) and (113), $x_n s_n = y_n' \sqrt{n \log p_n}$ and $x_n \to \infty$ as $n \to \infty$. Use Lemma 6.3 and the fact $1 - \Phi(t) = \frac{1}{\sqrt{2\pi}t} e^{-t^2/2}$ as $t \to +\infty$ to obtain

$$(116)\quad P\Big(\frac{\sum_{i=1}^n \eta_i'}{\sqrt{n \log p_n}} \geq y_n'\Big) = P\Big(\sum_{i=1}^n \eta_i' \geq x_n s_n\Big) \sim 1 - \Phi(x_n) \sim \frac{e^{-x_n^2/2}}{\sqrt{2\pi} x_n}$$

as $n \to \infty$. This and (115) conclude (111).
Step 3. Now we show

$$(117)\qquad \frac{e^{-x_n^2/2}}{\sqrt{2\pi x_n}} \sim \frac{p_n^{-y_n^2/2} (\log p_n)^{-1/2}}{\sqrt{2\pi} y} := \omega_n$$

as $n \to \infty$. Since $y_n \to y$ and $\sigma_n \to 1$, we know from (113) that

$$(118)\qquad \sqrt{2\pi} x_n = \sqrt{2\pi} y_n' (\log p_n)^{1/2} \sim \sqrt{2\pi} y\, (\log p_n)^{1/2}$$

as $n \to \infty$. Further, by (112),

$$(119)\quad \frac{e^{-x_n^2/2}}{p_n^{-y_n^2/2}} = \exp\Big\{-\frac{x_n^2}{2} + \frac{y_n^2}{2} \log p_n\Big\} = \exp\Big\{\frac{1}{2}\Big(y_n^2 - y_n'^2\Big) \log p_n\Big\}.$$

Since $y_n \to y$, by (113), both $\{y_n\}$ and $\{y_n'\}$ are bounded. It follows from (113) again that $|y_n^2 - y_n'^2| \leq C|y_n - y_n'| = O(e^{-n^\beta/C})$ as $n \to \infty$. With assumption $\log p_n = o(n^\beta)$ we get $e^{-x_n^2/2} \sim p_n^{-y_n^2/2}$ as $n \to \infty$, which combining with (118) yields (117).

Finally, we compare the right hand sides of (108) and (117). Choose $C' > \max\{y_n^2;\, n \geq 1\}$, since $\log p_n = o(n^\beta)$, recall ω_n in (117),

$$\begin{aligned}
\frac{2e^{-n^\beta/C}}{\omega_n} &= 2\sqrt{2\pi}\, y\, (\log p_n)^{1/2} p_n^{y_n^2/2} e^{-n^\beta/C} \\
&= O\left(n^{\beta/2} \cdot \exp\Big\{C' \log p_n - \frac{n^\beta}{C}\Big\}\right) \\
&= O\left(n^{\beta/2} \cdot \exp\Big\{-\frac{n^\beta}{2C}\Big\}\right) \to 0
\end{aligned}$$

as $n \to \infty$ for any constant $C > 0$. This fact joint with (108), (111) and (117) proves the lemma. ■

Proof of Lemma 6.9. For any Borel set $A \subset \mathbb{R}$, set

$$P_2(A) = P(A | u_{k1}, u_{k3},\, 1 \leq k \leq n),$$

the conditional probability of A with respect to u_{k1}, u_{k3}, $1 \leq k \leq n$. Observe from the expression of Σ_4 that three sets of random variables $\{u_{k1}, u_{k3};\ 1 \leq k \leq n\}$, $\{u_{k2};\ 1 \leq k \leq n\}$ and $\{u_{k4};\ 1 \leq k \leq n\}$ are independent. Then

$$\begin{aligned}
& P\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n,\ |\sum_{k=1}^n u_{k3}u_{k4}| > a_n\Big) \\
= \ & E\Big\{P_2\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n\Big)P_2\Big(|\sum_{k=1}^n u_{k3}u_{k4}| > a_n\Big)\Big\} \\
\leq \ & \Big\{E\, P_2\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n\Big)^2\Big\}^{1/2} \cdot \Big\{E\, P_2\Big(|\sum_{k=1}^n u_{k3}u_{k4}| > a_n\Big)^2\Big\}^{1/2}
\end{aligned}$$

by the Cauchy-Schwartz inequality. Use the same independence again

$$P_2\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n\Big) = P\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n\Big|u_{k1},\ 1 \leq k \leq n\Big); \tag{120}$$

$$P_2\Big(|\sum_{k=1}^n u_{k3}u_{k4}| > a_n\Big) = P\Big(|\sum_{k=1}^n u_{k3}u_{k4}| > a_n\Big|u_{k3},\ 1 \leq k \leq n\Big). \tag{121}$$

These can be also seen from Proposition 27 in Fristedt and Gray (1997). It follows that

$$\begin{aligned}
& \sup_{|r|\leq 1} P\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n,\ |\sum_{k=1}^n u_{k3}u_{k4}| > a_n\Big) \\
\leq \ & E\Big\{P\Big(|\sum_{k=1}^n u_{k1}u_{k2}| > a_n\Big|u_{11}, \cdots, u_{n1}\Big)^2\Big\}.
\end{aligned}$$

Since $\{u_{k1};\ 1 \leq k \leq n\}$ and $\{u_{k2};\ 1 \leq k \leq n\}$ are independent, and $t_n := a_n/\sqrt{n \log p} \to t = 2$, taking $\alpha = 2$ in Lemma 6.7, we obtain the desired conclusion from the lemma. ■

Proof of Lemma 6.10. Since Σ_4 is always non-negative definite, the determinant of the first 3×3 minor of Σ_4 is non-negative: $1 - r_1^2 - r_2^2 \geq 0$. Let $r_3 = \sqrt{1 - r_1^2 - r_2^2}$ and $\{u_{k5};\ 1 \leq k \leq n\}$ be i.i.d. standard normals which are independent of $\{u_{ki};\ 1 \leq i \leq 4;\ 1 \leq k \leq n\}$. Then,

$$(u_{11}, u_{12}, u_{13}, u_{14}) \stackrel{d}{=} (u_{11}, u_{12}, r_1u_{11} + r_2u_{12} + r_3u_{15}, u_{14}).$$

Define $Z_{ij} = |\sum_{k=1}^n u_{ki}u_{kj}|$ for $1 \le i, j \le 5$ and $r_5 = r_3$. By the Cauchy-Schwartz inequality,

$$
\begin{aligned}
|\sum_{k=1}^{n}(r_1u_{k1} + r_2u_{k2} + r_3u_{k5})u_{k4}| &\le \sum_{i\in\{1,2,5\}} |r_i|\cdot|\sum_{k=1}^{n} u_{ki}u_{k4}| \\
&\le \Big(r_1^2 + r_2^2 + r_3^2\Big)^{1/2}\Big(Z_{14}^2 + Z_{24}^2 + Z_{54}^2\Big)^{1/2} \\
&\le \sqrt{3}\cdot\max\{Z_{14}, Z_{24}, Z_{54}\}.
\end{aligned}
$$

It follows from the above two facts that

$$
\begin{aligned}
& P\Big(|\sum_{k=1}^{n} u_{k1}u_{k2}| > a_n,\, |\sum_{k=1}^{n} u_{k3}u_{k4}| > a_n\Big) \\
\le\ & P\Big(Z_{12} > a_n,\, \max\{Z_{14}, Z_{24}, Z_{54}\} > \frac{a_n}{\sqrt{3}}\Big) \\
\le\ & \sum_{i\in\{1,2,5\}} P\Big(Z_{12} > a_n,\, Z_{i4} > \frac{a_n}{\sqrt{3}}\Big) \\
=\ & 2P\Big(Z_{12} > a_n,\, Z_{14} > \frac{a_n}{\sqrt{3}}\Big) + P\Big(Z_{12} > a_n\Big)\cdot P\Big(Z_{54} > \frac{a_n}{\sqrt{3}}\Big) \qquad (122)
\end{aligned}
$$

by symmetry and independence. For any Borel set $A \subset \mathbb{R}$, set $P^1(A) = P(A|u_{k1},\, 1 \le k \le n)$, the conditional probability of A with respect to u_{k1}, $1 \le k \le n$. For any $s > 0$, from the fact that $\{u_{k1}\}, \{u_{k2}\}$ and $\{u_{k4}\}$ are independent, we see that

$$
\begin{aligned}
P\Big(Z_{12} > a_n,\, Z_{14} > sa_n\Big) &= E\Big(P^1(Z_{12} > a_n)\cdot P^1(Z_{14} > sa_n)\Big) \\
&\le \Big\{E\,P^1(Z_{12} > a_n)^2\Big\}^{1/2}\cdot\Big\{E\,P^1(Z_{14} > sa_n)^2\Big\}^{1/2}
\end{aligned}
$$

by the Cauchy-Schwartz inequality. Taking $t_n := a_n/\sqrt{n\log p} \to t = 2$ and $t_n := sa_n/\sqrt{n\log p} \to t = 2s$ in Lemma 6.7, respectively, we get

$$
E\,P^1(Z_{12} > a_n)^2 = O\Big(p^{-4+\epsilon}\Big) \quad\text{and}\quad EP^1(Z_{14} > sa_n)^2 = O\Big(p^{-4s^2+\epsilon}\Big)
$$

as $n \to \infty$ for any $\epsilon > 0$. This implies that, for any $s > 0$ and $\epsilon > 0$,

$$
P\Big(Z_{12} > a_n,\, Z_{14} > sa_n\Big) \le O\Big(p^{-2-2s^2+\epsilon}\Big) \tag{123}
$$

as $n \to \infty$. In particular,

$$
P\Big(Z_{12} > a_n,\, Z_{14} > \frac{a_n}{\sqrt{3}}\Big) \le O\Big(p^{-\frac{8}{3}+\epsilon}\Big) \tag{124}
$$

as $n \to \infty$ for any $\epsilon > 0$.

Now we bound the last term in (122). Note that $|u_{11}u_{12}| \le (u_{11}^2 + u_{12}^2)/2$, it follows that $Ee^{|u_{11}u_{12}|/2} < \infty$ by independence and $E\exp(N(0,1)^2/4) < \infty$. Since $\{u_{k1}, u_{k2};\ 1 \le k \le n\}$ are i.i.d. with mean zero and variance one, and $y_n := a_n/\sqrt{n\log p} \to 2$ as $n \to \infty$, taking $\alpha = 1$ in Lemma 6.8, we get

$$P\Big(Z_{12} > a_n\Big) = P\Big(\frac{1}{\sqrt{n\log p}}|\sum_{k=1}^n u_{k1}u_{k2}| > \frac{a_n}{\sqrt{n\log p}}\Big)$$

$$\sim 2\cdot \frac{p^{-y_n^2/2}(\log p)^{-1/2}}{2\sqrt{2\pi}} \sim \frac{e^{-y/2}}{\sqrt{2\pi}}\cdot\frac{1}{p^2} \tag{125}$$

as $n \to \infty$. Similarly, for any $t > 0$,

$$P\Big(Z_{12} > ta_n\Big) = O\Big(p^{-2t^2+\epsilon}\Big) \tag{126}$$

as $n \to \infty$ (this can also be derived from (i) of Lemma 6.4). In particular,

$$P\Big(Z_{54} > \frac{a_n}{\sqrt{3}}\Big) = P\Big(Z_{12} > \frac{a_n}{\sqrt{3}}\Big) = O\Big(p^{-\frac{2}{3}+\epsilon}\Big) \tag{127}$$

as $n \to \infty$ for any $\epsilon > 0$. Combining (125) and (127), we know that the last term in (122) is bounded by $O(p^{-\frac{8}{3}+\epsilon})$ as $n \to \infty$ for any $\epsilon > 0$. This together with (122) and (124) concludes the lemma. ■

Proof of Lemma 6.11. Fix $\delta \in (0,1)$. Take independent standard normals $\{u_{k5}, u_{k6};\ 1 \le k \le n\}$ that are also independent of $\{u_{ki};\ 1 \le i \le 4;\ 1 \le k \le n\}$. Then, since $\{u_{k1}, u_{k2}, u_{k5}, u_{k6};\ 1 \le k \le n\}$ are i.i.d. standard normals, by checking covariance matrix Σ_4, we know

$$(u_{11}, u_{12}, u_{13}, u_{14}) \stackrel{d}{=} (u_{11}, u_{12}, r_1u_{11} + r_1'u_{15}, r_2u_{12} + r_2'u_{16}) \tag{128}$$

where $r_1' = \sqrt{1-r_1^2}$ and $r_2' = \sqrt{1-r_2^2}$. Define $Z_{ij} = |\sum_{k=1}^n u_{ki}u_{kj}|$ for $1 \le i, j \le 6$. Then

$$|\sum_{k=1}^n (r_1u_{k1} + r_1'u_{k5})(r_2u_{k2} + r_2'u_{k6})|$$

$$\le |r_1r_2|Z_{12} + |r_1r_2'|Z_{16} + |r_1'r_2|Z_{25} + |r_1'r_2'|Z_{56}$$

$$\le (1-\delta)^2 Z_{12} + 3\max\{Z_{16}, Z_{25}, Z_{56}\} \tag{129}$$

for all $|r_1|, |r_2| \le 1-\delta$. Let $\alpha = (1 + (1-\delta)^2)/2$, $\beta = \alpha/(1-\delta)^2$ and $\gamma = (1-\alpha)/3$. Then

$$\beta > 1 \ \text{ and } \ \gamma > 0. \tag{130}$$

Easily, if $Z_{12} \le \beta a_n$, $\max\{Z_{16}, Z_{25}, Z_{56}\} \le \gamma a_n$, then from (129) we know that the left hand side of (129) is controlled by a_n. Consequently, by (128) and the i.i.d. property,

$$\begin{aligned}
& P(Z_{12} > a_n,\, Z_{34} > a_n) \\
= \;& P\Big(Z_{12} > a_n,\, |\sum_{k=1}^{n}(r_1 u_{k1} + r_1' u_{k5})(r_2 u_{k2} + r_2' u_{k6})| > a_n\Big) \\
\le \;& P(Z_{12} > a_n, Z_{12} > \beta a_n) + \sum_{i\in\{1,2,5\}} P(Z_{12} > a_n, Z_{i6} > \gamma a_n) \\
= \;& P(Z_{12} > \beta a_n) + 2P(Z_{12} > a_n, Z_{16} > \gamma a_n) \\
& \qquad + P(Z_{12} > a_n)\cdot P(Z_{56} > \gamma a_n) \qquad (131)
\end{aligned}$$

where "$2P(Z_{12} > a_n, Z_{16} > \gamma a_n)$" comes from that $(Z_{12}, Z_{16}) \overset{d}{=} (Z_{12}, Z_{26})$. Keep in mind that $(Z_{12}, Z_{16}) \overset{d}{=} (Z_{12}, Z_{14})$ and $Z_{56} \overset{d}{=} Z_{12}$. Recall (130), applying (123) and (126) to the three terms in the sum on the right hand side of (131), we conclude (72). ■

Proof of Lemma 6.12. Reviewing notation $\Omega_3 = \Omega_j$ for $j = 3$ defined below (65), the current case is that $d_1 \le d_3 \le d_2 \le d_4$ with $d = (d_1, d_2)$ and $d' = (d_3, d_4)$. Of course, by definition, $d_1 < d_2$ and $d_3 < d_4$. To save notation, define the "neighborhood" of d_i as follows:

$$N_i = \Big\{d \in \{1, \cdots, p\};\, |d - d_i| < \tau\Big\} \qquad (132)$$

for $i = 1, 2, 3, 4$.

Given $d_1 < d_2$, there are two possibilities for d_4: (a) $d_4 - d_2 > \tau$ and (b) $0 \le d_4 - d_2 \le \tau$. There are four possibilities for d_3: (A) $d_3 \in N_2\backslash N_1$; (B) $d_3 \in N_1\backslash N_2$; (C) $d_3 \in N_1 \cap N_2$; (D) $d_3 \notin N_1 \cup N_2$. There are eight combinations for the locations of (d_3, d_4) in total. However, by (64) the combination (a) & (D) is excluded. Our analysis next will exhaust all of the seven possibilities.

Case (a) & (A). Let $\Omega_{a,A}$ be the subset of $(d, d') \in \Omega_3$ satisfying restrictions (a) and (A), and others such as $\Omega_{b,C}$ are similarly defined. Thus,

$$\sum_{(d,d')\in\Omega_3} P(Z_d > a_n, Z_{d'} > a_n) \le \sum_{\theta,\Theta}\sum_{(d,d')\in\Omega_{\theta,\Theta}} P(Z_d > a_n, Z_{d'} > a_n) \qquad (133)$$

where θ runs over set $\{a, b\}$ and Θ runs over set $\{A, B, C, D\}$ but $(\theta, \Theta) \ne (a, D)$.

Easily, $|\Omega_{a,A}| \le \tau p^3$ and the covariance matrix of $(w_{d_2}, w_{d_1}, w_{d_3}, w_{d_4})$ (see (67)) is

$$\begin{pmatrix} 1 & 0 & \gamma & 0 \\ 0 & 1 & 0 & 0 \\ \gamma & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad |\gamma| \le 1.$$

Take $\epsilon = 1/2$ in Lemma 6.9 to have $P(Z_d > a_n, Z_{d'} > a_n) \equiv \rho_n = o(p^{-7/2})$ for all $(d, d') \in \Omega_{a,A}$. Thus

$$\sum_{(d,d')\in R} P(Z_d > a_n, Z_{d'} > a_n) = |R| \cdot \rho_n \to 0 \tag{134}$$

as $n \to \infty$ for $R = \Omega_{a,A}$.

Case (a) & (B). Notice $|\Omega_{a,B}| \le \tau p^3$ and $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ has the same covariance matrix as that in Lemma 6.9. By the lemma we then have (134) for $R = \Omega_{a,B}$.

Case (a) & (C). Notice $|\Omega_{a,C}| \le \tau^2 p^2$ and the covariance matrix of $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ is the same as that in Lemma 6.10. By the lemma, we know (134) holds for $R = \Omega_{a,C}$.

Case (b) & (A). In this case, $|\Omega_{b,A}| \le \tau^2 p^2$ and the covariance matrix of $(w_{d_3}, w_{d_4}, w_{d_2}, w_{d_1})$ is the same as that in Lemma 6.10. By the lemma and using the fact that

$$P(Z_d > a_n, Z_{d'} > a_n) = P(Z_{(d_3,d_4)} > a_n, Z_{(d_2,d_1)} > a_n)$$

we see (134) holds with $R = \Omega_{b,A}$.

Case (b) & (B). In this case, $|\Omega_{b,B}| \le \tau^2 p^2$ and the covariance matrix of $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ is the same as that in Lemma 6.11. By the lemma, we know (134) holds for $R = \Omega_{b,B}$.

Case (b) & (C). We assign positions for d_1, d_3, d_2, d_4 step by step: there are at most p positions for d_1 and at most k positions for each of d_3, d_2 and d_4. Thus, $|\Omega_{b,C}| \le \tau^3 p$. By (125),

$$P(Z_d > a_n, Z_{d'} > a_n) \le P(Z_d > a_n) = P\Big(|\sum_{i=1}^{n} \xi_i \eta_i| > a_n\Big) = O\Big(\frac{1}{p^2}\Big)$$

as $n \to \infty$, where $\{\xi_i, \eta_i;\, i \ge 1\}$ are i.i.d. standard normals. Therefore, (134) holds with $R = \Omega_{b,C}$.

Case (b) & (D). In this case, $|\Omega_{b,C}| \leq \tau p^3$ and the covariance matrix of $(w_{d_4}, w_{d_3}, w_{d_2}, w_{d_1})$ is the same as that in Lemma 6.9. By the lemma and noting the fact that

$$P(Z_d > a_n, Z_{d'} > a_n) = P(Z_{(d_4,d_3)} > a_n, Z_{(d_2,d_1)} > a_n)$$

we see (134) holds with $R = \Omega_{b,D}$.

We obtain (72) by combining (134) for all the cases considered above with (133). ■

Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104.
E-mail: tcai@wharton.upenn.edu
URL: http://www-stat.wharton.upenn.edu/ tcai

School of Statistics
University of Minnesota
224 Church Street
Minneapolis, MN55455
E-mail: tjiang@stat.umn.edu
URL: http://www.stat.umn.edu/ tjiang