The Annals of Probability
2019, Vol. 47, No. 5, 3321–3374
https://doi.org/10.1214/19-AOP1341

LARGEST ENTRIES OF SAMPLE CORRELATION MATRICES FROM EQUI-CORRELATED NORMAL POPULATIONS

BY JIANQING FAN[1] AND TIEFENG JIANG[2]

Princeton University and University of Minnesota

The paper studies the limiting distribution of the largest off-diagonal entry of the sample correlation matrices of high-dimensional Gaussian populations with equi-correlation structure. Assume the entries of the population distribution have a common correlation coefficient $\rho > 0$ and both the population dimension p and the sample size n tend to infinity with $\log p = o(n^{\frac{1}{3}})$. As $0 < \rho < 1$, we prove that the largest off-diagonal entry of the sample correlation matrix converges to a Gaussian distribution, and the same is true for the sample covariance matrix as $0 < \rho < 1/2$. This differs substantially from a well-known result for the independent case where $\rho = 0$, in which the above limiting distribution is an extreme-value distribution. We then study the phase transition between these two limiting distributions and identify the regime of ρ where the transition occurs. If ρ is less than, larger than or is equal to the threshold, the corresponding limiting distribution is the extreme-value distribution, the Gaussian distribution and a convolution of the two distributions, respectively. The proofs rely on a subtle use of the Chen–Stein Poisson approximation method, conditioning, a coupling to create independence and a special property of sample correlation matrices. An application is given for a statistical testing problem.

CONTENTS

Received September 2017; revised January 2019.
[1]Supported by NSF Grants DMS-1406266 and DMS-1712591.
[2]Supported by NSF Grant DMS-1406279.
MSC2010 subject classifications. Primary 62H10, 62E20; secondary 60F05.
Key words and phrases. Maximum sample correlation, phase transition, multivariate normal distribution, Gumbel distribution, Chen–Stein Poisson approximation.

1. Introduction. The correlation coefficient matrix is an important statistic in the multivariate analysis. It plays pivotal roles in the statistical analysis of a multivariate normal data. The maximum likelihood estimator is the sample correlation matrix. This paper investigates the limiting distribution of the largest off-diagonal entry of the sample correlation matrix in the high-dimensional setting when the correlation matrix admits a compound symmetry structure, namely, is of equi-correlation.

Let $X_1, \dots, X_n$ be a random sample from a p-dimensional population. We have the data matrix $\mathbf{X} = (X_1, \dots, X_n)'$. Write $\mathbf{X} = (x_{ij})_{n\times p} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(p)})$, then the Pearson correlation coefficient between $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$ is given by

$$\hat{\rho}_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2}\sqrt{\sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}, \quad 1 \le i, j \le p, \tag{1.1}$$

where $\bar{x}_i = n^{-1}\sum_{k=1}^n x_{ki}$. In particular, $\hat{\rho}_{ii} = 1$ for all $1 \le i \le p$. The sample correlation matrix $\hat{\mathbf{R}}$ is then defined by $\hat{\mathbf{R}} = (\hat{\rho}_{ij})_{p\times p}$. In contrast, $\mathbf{X}'\mathbf{X}/n$ is refereed to as the sample covariance matrix. Define the largest magnitude of off-diagonal entries of the sample correlation matrix by

$$L_{0n} = \max_{1\le i<j\le p} |\hat{\rho}_{ij}|. \tag{1.2}$$

Assuming that x_{ij}'s are independent and identically distributed but not necessarily Gaussian-distributed, the asymptotic distribution of L_{0n} have been extensively studied as both p and n tend to infinity. The first result on the topic is due to Jiang [9], who uses the Chen–Stein Poisson approximation method to get the limiting distribution of the L_{0n} as follows.

Assume $E|x_{11}|^{30+\varepsilon} < \infty$ *for some* $\varepsilon > 0$. *Let* $p = p_n$ *and* $n/p \to \gamma \in (0, \infty)$ *as* $n \to \infty$, *then*

$$P\bigl(nL_{0n}^2 - 4\log n + \log\log n \le t\bigr) \to \exp\biggl(-\frac{\gamma^2}{\sqrt{8\pi}}e^{-t/2}\biggr)$$

for any $t \in \mathbb{R}$, *or equivalently*,

$$P\bigl(nL_{0n}^2 - 4\log p + \log\log p \le t\bigr) \to \exp\biggl(-\frac{1}{\sqrt{8\pi}}e^{-t/2}\biggr). \tag{1.3}$$

Zhou [20] relaxes the moment condition to that $\lim_{x\to\infty} x^6 P(|x_{11}x_{12}| > x) = 0$ and that $\limsup_{n\to\infty} p/n < \infty$. Li and Rosalsky [14] consider the strong limit of L_{0n} under some more relaxed assumption. Li et al. [12, 13] have further improved the assumption of the result, under the assumption that p/n bounded away from zero or infinity. They actually obtain some necessary and sufficient conditions under which (1.3) holds. As $p/n \to \infty$, Liu et al. [16] establish similar results to (1.3) under the assumption $p = O(n^\alpha)$ where α is a constant. Cai and Jiang [3] consider the ultra-high-dimensional case where p can be as large as e^{n^α} for some $0 < \alpha \le 1$ and they extend the result to dependent case. Can and Jiang [4] derive the limiting distribution of L_{0n} under the assumption that the population has a spherical distribution. In fact, a phase transition phenomenon occurs at three different regimes: $\log p/n \to 0$, $\log p/n \to \alpha \in (0, \infty)$ and $\log p/n \to \infty$. By using the limiting distribution of L_{0n}, Cai et al. [2] work on the asymptotic behavior of the pairwise geodesic distances among n random points that are independently and uniformly distributed on the unit sphere in the p-dimensional spaces. The same phase transition phenomenon is also understood. Without the Gaussian assumption, Shao and Zhou [19] obtain similar results to (1.3) as $\log p = o(n^\alpha)$ for some $0 < \alpha \le 1$.

Assuming the p entries of $\mathbf{x}$ are independent, most of the aforementioned work mainly focus on the improvement of the moment assumption on x_{11} from the data matrix $\mathbf{X} = (x_{ij})_{n\times p}$ as well as relaxing the range of p relative to n. The question of how dependence impacts on the limiting distribution of the largest correlations remains largely unknown.

In this paper, we will consider a case that all the entries of $\mathbf{x}$ are very dependent. In fact, we assume $\mathbf{x} \sim N_p(\mu, \mathbf{\Sigma})$, where $N_p(\mu, \mathbf{\Sigma})$ stand for a p-variate normal population with the correlation matrix $\mathbf{R} = (\rho_{ij})_{p\times p}$, and the corresponding correlation matrix $\mathbf{R}$ has the compound symmetry structure, which is also referred to as the intraclass covariance or equi-correlation structure in literature, that is,

$$\mathbf{R} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}. \tag{1.4}$$

It is easy to see that $\mathbf{R}$ is positive definite if and only if $1 > \rho > -1/(p-1)$. Since we will be in the scenario that $p = p_n \to \infty$, we will always assume $\rho \ge 0$ later.

When $\rho > 0$, the sample correlations $\hat{\rho}_{ij}$, $1 \le i < j \le p$ are highly dependent and new technical challenges arise in deriving the limiting distribution of the maximum value of these entries. In addition, we found somewhat surprisingly that such a limiting distribution is Gaussian. This is in sharp contrast to the independence case ($\rho = 0$) in which the limiting distribution is a Gumbel distribution. Where does the phase transition occur? In what way the limiting distribution changes over the regime of correlation ρ? We provide sharp asymptotic results to describe these regimes of ρ and their associated limiting distributions of the maximum correlation.

Related to our study is the maximum spurious correlation between each variable in X and an independent variable Y in which the variables in X are correlated. Fan et al. [8] derive the asymptotic distribution of such a maximum spurious correlation using Gaussian approximation techniques of Chernozhukov et al. [6]. Unless the correlation matrix of X is of a specific form, such a limiting distribution can not be analytically derived and they require a multiplier bootstrap method to estimate the limiting distribution. Their setting relates to our case with the last row of off-diagonal correlation equal to zero and only computes the maximum sample correlation in the last row, albeit these sample correlations are also dependent due to the dependence of X.

Some notation will be used in the paper. The symbol $\xrightarrow{d}$ means convergence in distribution, $\xi \stackrel{d}{=} \eta$ implies that ξ and η have the same distribution. For two nonrandom sequences a_n and b_n, $b_n = o(a_n)$ means $b_n/a_n \to 0$, and $b_n = O(a_n)$ means $\limsup_{n\to\infty} |b_n/a_n| < \infty$. For a random sequence ξ_n and a nonrandom sequence a_n, $\xi_n = o_p(a_n)$ means $\xi_n/a_n \to 0$ in probability as $n \to \infty$, and $\xi_n = O_p(a_n)$ means $\lim_{C\to\infty} \limsup_{n\to\infty} P(|\xi_n/a_n| > C) = 0$. In addition, we denote C and C_1 positive constants independent of n or p, and their values may be different from line to line.

The rest of the paper is organized as follows. Section 2 gives the main results, discussions and an application. The proofs are relegated to Section 3, where we develop necessary technical tools for our quests.

2. Main results and discussions. Let $X_1, \ldots, X_n$ be a random sample from the population $N_p(\mu, \mathbf{\Sigma})$ with the population correlation matrix $\mathbf{R}$ defined as in (1.4). The data matrix is given by $\mathbf{X} = (X_1, \ldots, X_n)' = (x_{ij})_{n\times p}$.

We will study the following two statistics in this paper:

$$J_n = \max_{1\le i<j\le p} \frac{1}{n}\sum_{k=1}^{n} x_{ki}x_{kj} \quad \text{and} \quad L_n = \max_{1\le i<j\le p} \hat{\rho}_{ij}, \tag{2.1}$$

where $\hat{\rho}_{ij}$ is defined as in (1.1). The first statistic is the maximum off-diagonal entry of normalized sample covariances when $\mu = 0$, whereas the second statistic is the maximum off-diagonal entry of the sample correlations. The purpose of having the normalization in J_n is such that J_n and L_n have the same scale. To make our analysis thorough, we allow ρ to depend on n. It will be seen from Corollaries 2.1 and 2.2 later on that J_n and L_n behave differently as ρ is a constant.

2.1. *Main results.* We first consider the limiting distribution for the statistic J_n.

THEOREM 2.1. *Let $\rho_n \ge 0$ for each $n \ge 1$ and $\sup_{n\ge 1} \rho_n < 1/2$. Assume $\mu = \mathbf{0}$ and $\mathbf{\Sigma} = \mathbf{R}$, where $\mathbf{R}$ is given by (1.4). Suppose $p = p_n \to \infty$ and $\log p = o(n^{1/3})$ as $n \to \infty$. Set*

$$\mu_1 = \sqrt{n}\rho_n + \left(2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}}\right)\sqrt{1-\rho_n^2}.$$

The following holds as $n \to \infty$:

(i) *If $\rho_n\sqrt{\log p} \to 0$, then*

$$4\sqrt{\log p}(\sqrt{n}J_n - \mu_1) \xrightarrow{d} \phi,$$

where ϕ has distribution function $F(x) = e^{-Ke^{-x/2}}$, $x \in \mathbb{R}$ with $K = \frac{1}{4\sqrt{2\pi}}$.

(ii) *If* $\rho_n\sqrt{\log p} \to \lambda \in (0, \infty)$, *then*

$$\frac{\sqrt{n}J_n - \mu_1}{\sqrt{2}\rho_n} \xrightarrow{d} \xi + \lambda_0\phi,$$

where $\xi \sim N(0, 1)$, $\lambda_0 = \frac{1}{4\sqrt{2}\lambda}$, ϕ *is as in* (*i*) *and* ϕ *is independent of* ξ.

(iii) *If* $\rho_n\sqrt{\log p} \to \infty$, *then*

$$\frac{\sqrt{n}J_n - \mu_1}{\sqrt{2}\rho_n} \xrightarrow{d} N(0, 1).$$

The above theorem has the following implication.

COROLLARY 2.1. *Let* $\rho \in (0, 1/2)$ *be fixed*, $\mu = \mathbf{0}$ *and* $\mathbf{\Sigma} = \mathbf{R}$, *where* $\mathbf{R}$ *is as in* (1.4). *Suppose* $p = p_n \to \infty$ *and* $\log p = o(n^{1/3})$ *as* $n \to \infty$. *Then*

$$\frac{\sqrt{n}J_n - \mu_1}{\sqrt{2}\rho} \xrightarrow{d} N(0, 1)$$

as $n \to \infty$, *where* $\mu_1 = \sqrt{n}\rho + 2\sqrt{(1 - \rho^2)\log p}$.

For the largest entry of the sample correlation matrix $\hat{\mathbf{R}}$, we have the following.

THEOREM 2.2. *Let* $\rho_n \ge 0$ *for each* $n \ge 1$ *and* $\sup_{n\ge1}\rho_n < 1$. *Assume* $\mathbf{\Sigma} = \mathbf{R}$, *where* $\mathbf{R}$ *is as in* (1.4). *Let* $p = p_n \to \infty$ *satisfying* $\log p = o(n^{1/3})$ *as* $n \to \infty$. *Set*

$$\mu_2 = \sqrt{n-1}\rho_n + (1 - \rho_n) \cdot \sqrt{1 + 2\rho_n - \rho_n^2} \cdot \left(2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}}\right).$$

The following holds as $n \to \infty$:

(i) *If* $\rho_n\sqrt{\log p} \to 0$, *then*

$$4\sqrt{\log p}(\sqrt{n-1}L_n - \mu_2) \xrightarrow{d} \phi,$$

where ϕ *has the distribution function* $F(x) = e^{-Ke^{-x/2}}$, $x \in \mathbb{R}$ *with* $K = \frac{1}{4\sqrt{2\pi}}$.

(ii) *If* $\rho_n\sqrt{\log p} \to \lambda \in (0, \infty)$, *then*

$$\frac{\sqrt{n-1}L_n - \mu_2}{\sqrt{2}\rho_n} \xrightarrow{d} \xi + \lambda_0\phi,$$

where $\xi \sim N(0, 1)$, $\lambda_0 = \frac{1}{4\sqrt{2}\lambda}$ *and* ϕ *is the same as in* (i) *and* ϕ *is independent of* ξ.

(iii) *If* $\rho_n\sqrt{\log p} \to \infty$, *then*

$$\frac{\sqrt{n-1}L_n - \mu_2}{\sqrt{2}\rho_n(1 - \rho_n)} \xrightarrow{d} N(0, 1).$$

If ρ is close to zero, presumably the behavior of L_n is close to an extreme-value distribution as in (1.3); if ρ is relatively large, L_n is asymptotically the normal distribution as stated in Theorem 2.2. Part (ii) of the above theorem actually gives the phase transition between the two cases. The following is an easy consequence of Theorem 2.2.

COROLLARY 2.2. *Let* $\rho \in (0, 1)$ *be fixed and* $\mathbf{\Sigma} = \mathbf{R}$, *where* $\mathbf{R}$ *is as in* (1.4). *Suppose* $p = p_n \to \infty$ *and* $\log p = o(n^{1/3})$ *as* $n \to \infty$. *Then*, $(\sqrt{n-1}L_n - \mu_2)/\sigma_2 \xrightarrow{d} N(0, 1)$ *as* $n \to \infty$, *where*

$$\mu_2 = \rho\sqrt{n-1} + 2(1-\rho) \cdot \sqrt{1 + 2\rho - \rho^2} \cdot \sqrt{\log p} \quad \text{and} \quad \sigma_2 = \sqrt{2}\rho(1-\rho).$$

The above two results are totally different from Jiang [9], Zhou [20], Liu et al. [16], Li et al. [12, 13], Cai and Jiang [3], Can and Jiang [4], Cai et al. [2], Shao and Zhou [19]. They all end up with the Gumbel distribution by arguing that $\hat{\rho}_{ij}$'s are roughly independent random variables. In Theorems 2.1 and 2.2, the appearance of ρ creates a strong dependency among the terms $\sum_{k=1}^{n} x_{ki}x_{kj}$, $1 \le i < j \le p$, in the definition of J_n from (2.1). This is also true for the terms $\hat{\rho}_{ij}$, $1 \le i < j \le p$. The occurrence of ρ makes the situation so delicate that, if ρ is a constant, the limiting distributions of J_n and L_n are no longer the Gumbel distribution, they are the normal distribution instead.

For J_n (similarly for L_n), a key difference between the case $\rho = 0$ and the case $\rho > 0$ is explained as follows. When $\rho > 0$, each term of the denominator in (1.1) can no longer be regarded as roughly $\sqrt{n}$ as that in the case $\rho = 0$. In particular, if $\rho > 0$ is a constant, the dependence really matters, and the difference can be seen from Corollary 2.2 by comparing the means and the variances.

2.2. *Discussions*. The paper investigates the limiting distributions of the largest off-diagonal entry of sample covariance/correlation matrices generated by a random sample from a high-dimensional normal distribution. We assume the normal distribution has the structure of equi-correlation (1.4). Under the assumption that $p \to \infty$ and $\log p = o(n^{1/3})$, the asymptotic distributions of the largest off-diagonal entries of both matrices are established. Their behaviors depend on the value of ρ. The limits are the normal distribution if ρ is reasonably large; the limiting distributions are the extreme-value distribution if ρ is very small. We also figure out the regime to differentiate the two scenarios. In particular, for ρ in the regime, the limiting distribution is the convolution of the Gaussian distribution and the extreme-value distribution.

We make a few remarks as follows.

REMARK 2.1. For the sample correlation matrix $\hat{\mathbf{R}}$, we get the limiting distribution of its largest off-diagonal entry for each $\rho \in [0, 1)$. The same result holds for the sample covariance matrix but under the more stringent restriction $0 \le \rho < 1/2$, which is required in Lemma 3.11. This difference will be easily understood by the fact that the sample correlation matrix can be regarded as a type of self-normalized statistics. It is known that self-normalized statistics are more "tamed" (see, e.g., Shao and Wang [18]), and hence the range of ρ is more relaxed in the case of the sample correlation matrix than that in the case of the sample covariance matrix. We do not know whether or not Theorem 2.1 is still true for the case $\rho \in [1/2, 1)$. It is an interesting project for future.

REMARK 2.2. Under the Gaussian assumption and that for the equi-correlation $\mathbf{R}$ in (1.4), the decomposition structure of (3.21), that is,

$$X_1 = \sqrt{\rho}(\xi, \dots, \xi)' + \sqrt{1-\rho}(\xi_1, \dots, \xi_p)', \tag{2.2}$$

where $\xi, \xi_1, \dots, \xi_p$ are independent standard Gaussian random variables, and plays a key role in the proofs. Now let us remove the Gaussian assumption. Instead, we assume the decomposition (2.2) continues to hold with $\xi, \xi_1, \dots, \xi_p$ relaxed to be i.i.d. random variables. Then Theorems 2.1 and 2.2 may also hold.

REMARK 2.3. The paper deals with the equi-correlation matrix. If $\mathbf{R} = (r_{ij})$ has another special structure, one may like to work on $\max_{1 \le i < j \le p} \hat{\rho}_{ij}$ or $\max_{1 \le i < j \le p} \hat{\rho}_{ij}/\rho_{ij}$. It seems that, to get good properties for these two quantities, $\mathbf{R}$ cannot be arbitrary.

REMARK 2.4. Recall the notation J_n in (2.1). Let $\hat{J}_n = n^{-1}\sum_{k=1}^n x_{k1}x_{k2}$ be the sample mean of i.i.d. random variables. Assuming ρ is fixed. It is easy to check that $(\sqrt{n}\hat{J}_n - \sqrt{n}\rho)/\sqrt{1+\rho^2} \xrightarrow{d} N(0,1)$. Corollary 2.1 asserts that

$$\frac{\sqrt{n}J_n - \sqrt{n}\rho - 2\sqrt{(1-\rho^2)\log p}}{\sqrt{2}\rho} \xrightarrow{d} N(0,1). \tag{2.3}$$

Comparing the two limiting results, we see J_n has a larger mean and a smaller standard deviation than those of $\hat{J}_n$, respectively. However, the two means are on the same scale and so are the two standard deviations. This concludes that the correlation, instead of independence, dominates J_n such that J_n, as the maximum of many random variables, is more or less like a single term $\hat{J}_n$.

REMARK 2.5. Assuming $\rho = 0$, Jiang [10] obtains the limiting spectral distribution of the sample correlation matrix $\hat{\mathbf{R}}$. When $n/p \to c \in (0, \infty)$, the author proves that the empirical spectral distribution of $\hat{\mathbf{R}}$ asymptotically obeys the Marchenko–Pastur law. If $0 < \rho < 1$, by using the decomposition (3.21), it can be shown easily that the spectral distribution of the sample covariance matrix also takes the Marchenko–Pastur law as its limit. A similar result is expected for correlation matrix $\hat{\mathbf{R}}$ for the case $\rho > 0$ by employing the approximation method from Jiang [10].

REMARK 2.6. Review the limiting distribution in Corollary 2.2 is normal. This does not mean the convergence rate is similar to that of the standard CLT. The convergence rate of CLT is much faster than that of maximum of i.i.d. Gaussian random variables which converges to the Gumbel distribution. This can be seen easily from the following example. Let $\{x_i; 1 \le i \le n\}$ be $N(0, 1)$-distributed random variables with $\mathrm{Cov}(x_i, x_j) = \rho > 0$ for all $i \ne j$. Then we can write $x_i = \sqrt{\rho}\eta + \sqrt{1-\rho}\eta_i$ for each i, where $\eta, \eta_1, \dots, \eta_n$ are i.i.d. $N(0, 1)$-distributed random variables. Obviously, $\max_{1\le i\le n} x_i = \sqrt{\rho}\eta + \sqrt{1-\rho}\max_{1\le i\le n}\eta_i$. It is known that

$$\max_{1\le i\le n}\eta_i = \sqrt{2\log n} - \frac{\log\log n + \log(4\pi)}{2\sqrt{2\log n}} + \frac{U_n}{\sqrt{2\log n}}, \tag{2.4}$$

where $U_n \xrightarrow{d} e^{-e^{-x}}$ (see, e.g., Leadbetter et al. [11]). This concludes that

$$\max_{1\le i\le n} x_i - \sqrt{2(1-\rho)\log n} = \sqrt{\rho}\eta + \varepsilon_n,$$

where $\varepsilon_n \sim \frac{\sqrt{1-\rho}}{\sqrt{8}} \cdot \frac{\log\log n}{\sqrt{\log n}}$ in probability. So, $\max_{1\le i\le n} x_i - \sqrt{2(1-\rho)\log n} \xrightarrow{d} N(0, \rho)$ with speed $\frac{\log\log n}{\sqrt{\log n}}$, which is even slower than the convergence rate $\frac{1}{\sqrt{\log n}}$ appeared in the coefficient of U_n from (2.4).

2.3. *An application to a high-dimensional test.* Let $X_1, \dots, X_n$ be a random sample from the population $N_p(\mu, \mathbf{\Sigma})$. We are interested in testing whether $\mathbf{\Sigma}$ is diagonal. A natural nonparametric test is to use the test statistic L_n, which is powerful for sparse alternatives. The null distribution of such a test statistic corresponds to the limiting distribution for case $\rho = 0$ in regime (i) of Theorem 2.2. A question arises naturally how powerful it is under the dense alternatives. The specific alternative of interest is

$$H_0 : \rho = 0 \quad \text{v.s.} \quad H_1 : \rho = \rho_1,$$

where $\rho_1 \in (0, 1)$ is given.

Assume the dimension p and sample size n are all very large such that $\log p = o(n^{1/3})$. By part (i) of Theorem 2.2, under H_0,

$$4\sqrt{\log p}(\sqrt{n-1}L_n - \mu_{20}) \xrightarrow{d} \phi,$$

where $\mu_{20} = 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}}$ and ϕ has distribution function $F(x) = e^{-Ke^{-x/2}}$, $x \in \mathbb{R}$ and $K = \frac{1}{4\sqrt{2\pi}}$. For $0 < \alpha < 1$, denote q_α the $(1-\alpha)$-quantile of the distribution $F(x)$, that is,

$$q_\alpha = -\log(32\pi) - 2\log\log(1-\alpha)^{-1}. \tag{2.5}$$

Then a rejection region of the asymptotic size-α test is given by

$$\mathfrak{X}_0 = \left\{ \sqrt{n-1} L_n \geq 2\sqrt{\log p} + \frac{q_\alpha - \log\log p}{4\sqrt{\log p}} \right\}. \tag{2.6}$$

Using part (i) of Theorem 2.2 again, when $\rho_1 = o(1/\sqrt{\log p})$, the asymptotic power is still α, like a random guess, as the asymptotic distribution under such a contiguous alternative hypothesis is the same as that under the null hypothesis. Now the power starts to emerge when $\rho_1 = \lambda/\sqrt{\log p}$ for $\lambda \in (0, \infty)$ as in case (ii) of Theorem 2.2. In this case, μ_2 can be calculated as

$$\mu_{22} := \lambda \sqrt{\frac{n-1}{\log p}} + \left[1 - \frac{2\lambda^2}{\log p} + O\left(\frac{1}{(\log p)^{3/2}} \right) \right] \mu_{20}.$$

The power function is

$$\begin{aligned} \beta(\rho_1) &= P\{\sqrt{n-1} L_n \geq \mu_{20} + q_\alpha/(4\sqrt{\log p}) | \rho_1\} \\ &= P\{4\sqrt{\log p}(\sqrt{n-1} L_n - \mu_{22}) \geq q_\alpha - 4\lambda\sqrt{n-1} + 16\lambda^2 + o(1) | \rho_1\}. \end{aligned}$$

According to part (ii) of Theorem 2.2, the power tends to 1 for each fixed λ. By using a similar argument, it is easy to show that the power in case (iii) has also asymptotic power 1.

3. Proofs. The proofs of Theorems 2.1 and 2.2 are quite involved. We break them into Sections 3.1, 3.2, 3.3 and 3.4. In Section 3.1, we provide some preliminary results; In Section 3.2, we prove the main results; in Section 3.3 and Section 3.4, we prove some key technical steps used in the proofs of Theorem 2.1 and Theorem 2.2, respectively.

3.1. *Some technical tools.* In this section, we will collect and prove some technical tools toward the proofs of Theorems 2.1 and 2.2.

We first present Lemma 3.2, a special property of the sample correlation matrix $\hat{\mathbf{R}}$ as defined below (1.1). An auxiliary fact has to be derived first.

LEMMA 3.1. *Let $X_1, \ldots, X_n$ be i.i.d. random vectors and $X_1 \sim N_p(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a $p \times p$ positive semidefinite matrix. Set $\mathbf{X} = (X_1, \ldots, X_n)'$. Then, for any $n \times n$ orthogonal matrix $\mathcal{O}$, we have $\mathcal{O}\mathbf{X} \stackrel{d}{=} \mathbf{X}$.*

PROOF. Let $Y_1, \ldots, Y_n$ be i.i.d. and $Y_1 \sim N_p(\mathbf{0}, \mathbf{I}_p)$. Then X_i and $\mathbf{\Sigma}^{1/2} Y_i$ have the same distribution for each i. By independence,

$$\mathbf{X} = (X_1, \ldots, X_n)' \stackrel{d}{=} (Y_1, \ldots, Y_n)' \mathbf{\Sigma}^{1/2}. \tag{3.1}$$

As a consequence,

$$\mathcal{O}\mathbf{X} \stackrel{d}{=} \mathcal{O}(Y_1, \ldots, Y_n)' \mathbf{\Sigma}^{1/2}$$

for any $n \times n$ orthogonal matrix $\mathcal{O}$. Write $(Y_1, \dots, Y_n)' = (y_{ij})_{n\times p}$. Then y_{ij}'s are i.i.d. $N(0,1)$-distributed random variables. Hence $\mathcal{O}(Y_1, \dots, Y_n)' \stackrel{d}{=} (Y_1, \dots, Y_n)'$ by the orthogonal invariance of independent Gaussian random variables. From (3.1), it follows that

$$\mathcal{O}\mathbf{X} \stackrel{d}{=} (Y_1, \dots, Y_n)'\mathbf{\Sigma}^{1/2} \stackrel{d}{=} \mathbf{X}. \qquad \square$$

The following lemma provides a simple expression for the sample correlation matrix.

LEMMA 3.2. *Let $X_1, \dots, X_n$ be i.i.d. random vectors and $X_1 \sim N_p(\mu, \mathbf{\Sigma})$ where $\mu \in \mathbb{R}^p$ and $\mathbf{\Sigma}$ is a positive definite matrix. Let $\hat{\rho}_{ij}$ be as in* (1.1). *Suppose $Y_1, \dots, Y_{n-1}$ are i.i.d. and $Y_1 \sim N_p(\mathbf{0}, \mathbf{\Sigma})$. Write $(Y_1, \dots, Y_{n-1})' = (V_1, \dots, V_p)_{(n-1)\times p}$. Then*

$$(\hat{\rho}_{ij})_{p\times p} \stackrel{d}{=} \left(\frac{V_i'V_j}{\|V_i\| \cdot \|V_j\|}\right)_{p\times p}.$$

PROOF. Since $\hat{\rho}_{ij}$ is invariant under translation and scaling of the vectors $X_1, \dots, X_n$, we assume $\mu = \mathbf{0}$ without loss of generality.

Denotes $\mathbf{1} = (1, 1, \dots, 1)' \in \mathbb{R}^{n\times 1}$ and $\mathbf{A}_{n\times n} = \mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}'$. Trivially, $\mathbf{A}$ is an idempotent matrix with $\operatorname{tr}(\mathbf{A}) = n-1$, then there exists an $n \times n$ orthogonal matrix $\mathcal{O}$ such that

$$\mathbf{A} = \mathcal{O}' \begin{pmatrix} \mathbf{I}_{n-1} & 0 \\ 0 & 0 \end{pmatrix} \mathcal{O}.$$

Write

$$\begin{pmatrix} x_{1j} - \bar{x}_j \\ x_{2j} - \bar{x}_j \\ \vdots \\ x_{nj} - \bar{x}_j \end{pmatrix} = \mathbf{A} \begin{pmatrix} cx_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} \tag{3.2}$$

for each $1 \le j \le p$. Write $\mathbf{X} = (X_1, \dots, X_n)' = (x_{ij})_{n\times p}$. Then

$$\mathbf{H} := \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \cdots & x_{np} - \bar{x}_p \end{pmatrix}$$

$$= \mathcal{O}' \begin{pmatrix} \mathbf{I}_{n-1} & 0 \\ 0 & 0 \end{pmatrix} \mathcal{O}\mathbf{X} \stackrel{d}{=} \mathcal{O}' \begin{pmatrix} \mathbf{I}_{n-1} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{X}$$

by Lemma 3.1. Then

$$\tilde{\mathbf{X}} := \begin{pmatrix} \mathbf{I}_{n-1} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{X} = \begin{pmatrix} (x_{ij})_{(n-1)\times p} \\ \mathbf{0} \end{pmatrix},$$

where $\mathbf{0}$ above is a p-dimensional row vector with all entries equal to zero. Therefore, $\mathbf{H} \stackrel{d}{=} \mathcal{O}'\tilde{\mathbf{X}}$, and hence

$$\mathbf{H}'\mathbf{H} \stackrel{d}{=} \tilde{\mathbf{X}}'\tilde{\mathbf{X}} = (x_{ij})'_{(n-1)\times p}(x_{ij})_{(n-1)\times p}.$$

Define $(x_{ij})_{(n-1)\times p} = (V_1, \dots, V_p)_{(n-1)\times p}$. The above implies

$$\mathbf{H}'\mathbf{H} \stackrel{d}{=} (V_i'V_j)_{p\times p}. \tag{3.3}$$

For a positive definite matrix $\mathbf{M} = (m_{ij})_{p\times p}$, define $h(\mathbf{M})$ to be a $p \times p$ matrix such that its (i, j)-entry is equal to $m_{ij}m_{ii}^{-1/2}m_{jj}^{-1/2}$. Let $\mathcal{M}_{p\times p}$ be the set of all $p \times p$ positive definite matrices. Then $h : \mathcal{M}_{p\times p} \to \mathcal{M}_{p\times p}$ is continuous map and, therefore, is Borel-measurable map. From (3.3), we conclude $h(\mathbf{H}'\mathbf{H}) \stackrel{d}{=} h((V_i'V_j)_{p\times p})$. The desired conclusion then follows. □

The next one is the Chen–Stein Poisson approximation method, which is a special case of Arratia, Goldstein and Gordon [1], Theorem 1.

LEMMA 3.3. *Let $\{\eta_\alpha, \alpha \in I\}$ be random variables on an index set I and $\{B_\alpha, \alpha \in I\}$ be a set of subsets of I, that is, for each $\alpha \in I$, $B_\alpha \subset I$. For any $t \in \mathbb{R}$, set $\lambda = \sum_{\alpha\in I} P(\eta_\alpha > t)$, Then we have*

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

where

$$b_1 = \sum_{\alpha\in I}\sum_{\beta\in B_\alpha} P(\eta_\alpha > t)P(\eta_\beta > t),$$

$$b_2 = \sum_{\alpha\in I}\sum_{\alpha\neq\beta\in B_\alpha} P(\eta_\alpha > t, \eta_\beta > t),$$

$$b_3 = \sum_{\alpha\in I}\big|P\{\eta_\alpha > t|\sigma(\eta_\beta, \beta \notin B_\alpha)\} - P(\eta_\alpha > t)\big|,$$

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

The next lemma is on the moderation deviation of the partial sum of i.i.d. random variables. It can be seen, for instance, from Linnik [15].

LEMMA 3.4. *Suppose $\{\zeta, \zeta_1, \zeta_2, \dots\}$ is a sequence of i.i.d. random variables with $E\zeta_1 = 0$ and $E\zeta_1^2 = 1$. Define $S_n = \sum_{i=1}^n \zeta_i$. If $Ee^{t_0|\zeta|^\alpha} < \infty$ for some $0 < \alpha \le 1$ and $t_0 > 0$, then*

$$\lim_{n\to\infty} \frac{1}{x_n^2}\log P\left(\frac{S_n}{\sqrt{n}} \ge x_n\right) = -\frac{1}{2}$$

for any $x_n \to \infty$, $x_n = o(n^{\frac{\alpha}{2(2-\alpha)}})$.

The following lemma studies the moderation deviation of the partial sum of the independent but not necessarily identically distributed random variables; see Chen et al. [5], Proposition 4.5.

LEMMA 3.5. *Let η_i, $1 \le i \le n$ be independent random variables with $E\eta_i = 0$ and $Ee^{h_n|\eta_i|} < \infty$ for some $h_n > 0$ and $1 \le i \le n$. Assume that $\sum_{i=1}^n E\eta_i^2 = 1$. Then*

$$\frac{P(\sum_{i=1}^n \eta_i \ge x)}{1-\Phi(x)} = 1 + C_n(1+x^3)\gamma e^{4x^3\gamma}$$

for all $0 \le x \le h_n$ and $\gamma = \sum_{i=1}^n E(|\eta_i|^3 e^{x|\eta_i|})$, where $\sup_{n\ge1}|C_n| \le C$ and C is an absolute constant.

In our framework, η_i above is a quadratic form of two independent normal variables for each i. We first need to control $E(|\eta_i|^3 e^{x|\eta_i|})$.

LEMMA 3.6. *Let U and V be i.i.d. $N(0,1)$-distributed random variables. Let a, b, c, d, e, f be real numbers. Set $\eta = aU^2 + bUV + cV^2 + dU + eV + f$. Then*

$$E(|\eta|^3 e^{x|\eta|}) \le C \cdot (|a|^3 + |b|^3 + |c|^3 + |d|^3 + |e|^3 + |f|^3) \cdot e^{2(d^2+e^2)x^2+|f|x}$$

as $0 < x \le \frac{1}{12(|a|+|b|+|c|)}$, where C is constant not depending on a, b, c, d, e or f.

PROOF. First, use $|UV| \le U^2 + V^2$ to see

$$|\eta| \le (|a|+|b|)U^2 + (|b|+|c|)V^2 + |dU+eV| + |f|. \tag{3.4}$$

In particular,

$$\begin{aligned} E|\eta|^9 &\le 4^8 \cdot E[(|a|+|b|)^9 U^{18} + (|b|+|c|)^9 V^{18} + |dU+eV|^9 + |f|^9] \\ &\le C_1[(|a|+|b|)^9 + (|b|+|c|)^9 + (d^2+e^2)^{9/2} + |f|^9] \\ &\le C_1[(|a|+|b|+|c|)^9 + (|d|+|e|)^9 + |f|^9], \end{aligned}$$

where C_1 is a constant not depending on a, b, c, d, e or f. We also use the facts $E(U^{18}+V^{18}) < \infty$ and $dU + eV \overset{d}{=} \sqrt{d^2+e^2}\,U$. It follows that

$$(E|\eta|^9)^{1/3} \le C_1^{1/3}(|a|+|b|+|c|+|d|+|e|+|f|)^3. \tag{3.5}$$

From (3.4),

$$\begin{aligned} E(|\eta|^3 e^{x|\eta|}) \le (E|\eta|^9)^{1/3} &\cdot [E\exp(3x(|a|+|b|)U^2 + 3x(|b|+|c|)V^2)]^{1/3} \\ &\cdot [E\exp(3x|dU+eV|)]^{1/3} \cdot e^{x|f|}. \end{aligned}$$

First,

$$Ee^{3x\cdot|dU+eV|} = Ee^{3x\sqrt{d^2+e^2}|U|}$$
$$\leq Ee^{3x\sqrt{d^2+e^2}U} + Ee^{-3x\sqrt{d^2+e^2}U} = 2e^{9x^2(d^2+e^2)/2}$$

by using the identity $Ee^{tN(0,1)} = e^{t^2/2}$ for all $t \in \mathbb{R}$. Second, setting $\alpha = 3x(|a| + |b|)$ and $\beta = 3x(|b| + |c|)$, and reviewing $Ee^{sU^2} = (1-2s)^{-1/2}$ for all $s < \frac{1}{2}$, we have

$$E\exp\bigl(3x\bigl(|a|+|b|\bigr)U^2 + 3x\bigl(|b|+|c|\bigr)V^2\bigr) = (1-2\alpha)^{-1}\cdot(1-2\beta)^{-1} \leq 4$$

if $\alpha \leq \frac{1}{4}$ and $\beta \leq \frac{1}{4}$ by independence. Finally, combining the above, we see

$$E\bigl(|\eta|^3 e^{x|\eta|}\bigr) \leq C\cdot\bigl(|a|+|b|+|c|+|d|+|e|+|f|\bigr)^3 e^{2(d^2+e^2)x^2+|f|x}$$

as $0 < x \leq \frac{1}{12(|a|+|b|+|c|)}$. The conclusion then comes from an inequality on convex function $f(x) := x^3$ for $x \geq 0$. $\square$

In our setting, the parameter γ from Lemma 3.5 needs a special care. This will be done below with the help of Lemma 3.6.

LEMMA 3.7. *Let $\{\xi_k; k \geq 1\}$ be i.i.d. $N(0,1)$-distributed random variables. Set $\tau = E(|\xi_1|^3) + 1$. Assume $p = p_n$ satisfies that $p \to \infty$ and $\log p = o(n^{1/3})$. Let $\{y_n > 0; n \geq 1\}$ be real numbers such that $y_n = O(\log p)$. Then,*

$$P\left(\frac{1}{n}\sum_{k=1}^{n}(1+|\xi_k|^3)e^{y_n\xi_k^2/n} \geq 2\tau\right) \leq \exp\left(-\frac{1}{4}n^{1/2}(\log n)^{-2}\right)$$

as n is sufficiently large.

PROOF. By assumption, we assume $y_n \leq N_0 \log p$ for all $n \geq 1$, where $N_0 > 0$ is a constant. For $\varepsilon > 0$, set $\Theta_\varepsilon = \{\max_{1\leq k\leq n}\xi_k^2 \leq \varepsilon n/y_n\}$. By the inequality $P(N(0,1) \geq y) \leq \frac{1}{\sqrt{2\pi}y}e^{-y^2/2} \leq e^{-y^2/2}$ for all $y \geq 1$, there exists a constant $n_1 \geq 1$ such that

$$P\bigl(\Theta_\varepsilon^c\bigr) \leq nP\bigl(|\xi_1| > (\varepsilon n/y_n)^{1/2}\bigr) \leq n\cdot\exp\left(-\frac{\varepsilon}{2}\cdot\frac{n}{y_n}\right)$$

as $n \geq n_1$, which is again bounded by

$$n\cdot\exp\left(-\frac{\varepsilon}{2N_0}\cdot\frac{n}{\log p}\right) \leq n\cdot e^{-n^{2/3}}$$

as $n \geq n_\varepsilon \geq n_1$, where $n_\varepsilon \geq 1$ is an integer depending on ε. It follows that

$$P\left(\frac{1}{n}\sum_{k=1}^{n}(1+|\xi_k|^3)e^{y_n\xi_k^2/n} \geq 2\tau\right) \leq P\left(\frac{1}{n}\sum_{k=1}^{n}(1+|\xi_k|^3)e^{\varepsilon} \geq 2\tau\right) + n\cdot e^{-n^{2/3}}$$

as $n \geq n_\varepsilon$. Take $\varepsilon = \log(4/3)$. Then $2e^{-\varepsilon}\tau = 3\tau/2$. Consequently,

$$P\left(\frac{1}{n}\sum_{k=1}^{n}(1+|\xi_k|^3)e^{\varepsilon} \geq 2\tau\right) = P(\frac{1}{n}\sum_{k=1}^{n}\left(|\xi_k|^3 - E(|\xi_k|^3) \geq \frac{1}{2}\tau\right) \leq P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\zeta_k \geq x_n\right)$$

as n is sufficiently large, where $\zeta_k = (|\xi_k|^3 - E(|\xi_k|^3))/\sqrt{\mathrm{Var}(\xi_1^3)}$ and $x_n = n^{1/4}/\log n$. Set $\sigma = \sqrt{\mathrm{Var}(\xi_1^3)}$. Observe that $\sigma^{2/3}\cdot|\zeta_k|^{2/3} \leq |\xi_k|^2 + (E|\xi_k|^3)^{2/3}$. This implies $E\exp(\sigma^{2/3}|\zeta_k|^{2/3}/4) < \infty$ since $\xi_k \sim N(0,1)$. Take $\alpha = 2/3$ in Lemma 3.4 to see

$$P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\zeta_k \geq x_n\right) \leq \exp\left(-\frac{1}{4}n^{1/2}(\log n)^{-2}\right)$$

as n is sufficiently large. In summary,

$$P\left(\frac{1}{n}\sum_{k=1}^{n}(1+|\xi_k|^3)e^{y_n\xi_k^2/n} \geq 2\tau\right) \leq \exp\left(-\frac{1}{4}n^{1/2}(\log n)^{-2}\right) + n\cdot e^{-n^{2/3}}.$$

This implies the desired inequality. □

The following result provides us with an equivalent expression on a limit theorem. It will be applied to the proofs of Propositions 3.1 and 3.2 later, in which $F(x)$ is an extreme-value distribution.

LEMMA 3.8. *Let M_n be a random variable for each $n \geq 1$ satisfying*

$$\lim_{n\to\infty} P(M_n \leq \sqrt{4\log p - \log\log p + x}) = F(x)$$

for any $x \in \mathbb{R}$, where $F(x)$ is a continuous distribution function on $\mathbb{R}$. Then

$$M_n = 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{1}{4\sqrt{\log p}}U_n,$$

where U_n converges weakly to a probability measure with distribution function $F(x)$.

PROOF. Easily, $(1+t)^{1/2} = 1 + \frac{1}{2}t + r(t)$ where $\sup_{|t|<\varepsilon} |r(t)| \le t^2$ for some $\varepsilon > 0$. Fix $x_0 \in \mathbb{R}$. Let $A_0 > 0$ be given. For any $x \in [x_0 - A_0, x_0 + A_0]$,

$$\begin{aligned}\sqrt{4\log p - \log\log p + x} &= 2\sqrt{\log p}\left(1 - \frac{\log\log p}{4\log p} + \frac{x}{4\log p}\right)^{1/2} \\ &= 2\sqrt{\log p}\left[1 - \frac{\log\log p}{8\log p} + \frac{x}{8\log p} + r(p,x)\right],\end{aligned}$$

where

$$\sup_{|x-x_0|\le A_0} |r(p,x)| \le \sup_{|x-x_0|\le A_0} \left(\frac{\log\log p}{4\log p} - \frac{x}{4\log p}\right)^2 \le \frac{(\log\log p)^2}{15(\log p)^2}$$

as n is large enough. By the given condition,

$$\lim_{n\to\infty} P\left(M_n \le 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{x}{4\sqrt{\log p}} + s(p,x)\right) = F(x) \tag{3.6}$$

as $n \to \infty$, where $s(p,x) := 2r(p,x)\sqrt{\log p}$ and

$$\sup_{|x-x_0|\le A_0} |s(p,x)| \le \frac{(\log\log p)^2}{7(\log p)^{3/2}} \tag{3.7}$$

as n is sufficiently large. Define

$$U_n = 4\sqrt{\log p}\left(M_n - 2\sqrt{\log p} + \frac{\log\log p}{4\sqrt{\log p}}\right). \tag{3.8}$$

Then (3.6) implies that

$$\lim_{n\to\infty} P\big(U_n \le x + t(p,x)\big) = F(x), \tag{3.9}$$

where $t(p,x) := 4s(p,x)\sqrt{\log p}$. Easily, from (3.7),

$$\sup_{|x-x_0|\le A_0} |t(p,x)| \le \frac{(\log\log p)^2}{\log p}$$

as n is sufficiently large. Therefore, for any $\delta > 0$,

$$P(U_n \le x - \delta) \le P\big(U_n \le x + t(p,x)\big) \le P(U_n \le x + \delta)$$

as n is sufficiently large. From (3.9),

$$\limsup_{n\to\infty} P(U_n \le x - \delta) \le F(x) \le \liminf_{n\to\infty} P(U_n \le x + \delta)$$

for any $x \in [x_0 - A_0, x_0 + A_0]$. For $\delta \in (0, A_0)$, taking $x = x_0 + \delta$ and $x = x_0 - \delta$, respectively, we have

$$\begin{aligned}&\limsup_{n\to\infty} P(U_n \le x_0) \le F(x_0 + \delta);\\ &\liminf_{n\to\infty} P(U_n \le x_0) \ge F(x_0 - \delta).\end{aligned}$$

Letting $\delta \downarrow 0$, we obtain $\lim_{n\to\infty} P(U_n \le x_0) = F(x_0)$. Since $x_0 \in \mathbb{R}$ is arbitrary, this limiting result together with (3.8) concludes the proof. □

The next lemma is a coupling result, enabling us to prove that the two random variables appearing in parts (ii) of Theorems 2.1 and 2.2 are asymptotically independent. Now we assume $\{\xi_k, \xi_{ki}; 1 \le k \le n, 1 \le i < j \le p\}$ are i.i.d. random variables with the distribution $N(0, 1)$.

LEMMA 3.9. *Suppose $p = p_n \to \infty$ as $n \to \infty$. For all $1 \le i < j \le p$, let $C_{nij} = n^{-1/2} \sum_{k=1}^n \xi_k(\xi_{ki} + \xi_{kj})$. For any real numbers $\{\lambda_n; n \ge 1\}$ and any set of random variables $\{H_{ij}; 1 \le i < j \le p\}$, we have*

$$\max_{1\le i<j\le p}\{H_{i,j} + \lambda_n C_{nij}\} = \max_{1\le i<j\le p}\left\{H_{i,j} + \lambda_n \cdot \frac{\sqrt{n}}{\|\xi\|} C_{nij}\right\} + O_p\left(\frac{\lambda_n\sqrt{\log p}}{\sqrt{n}}\right)$$

as $n \to \infty$. The above statement also holds if "C_{nij}" is replaced by "C_{mij}" with $m = n - 1$.

PROOF. Recall $\|\xi\| = (\xi_1^2 + \cdots + \xi_n^2)^{1/2}$. Then

$$\begin{aligned}\left|\frac{\sqrt{n}}{\|\xi\|} - 1\right| &= \frac{|\|\xi\|^2 - n|}{\|\xi\| + \sqrt{n}} \cdot \frac{1}{\|\xi\|} \\ &\le \frac{1}{\sqrt{n}} \cdot \frac{\sqrt{n}}{\|\xi\|} \cdot \left|\frac{1}{\sqrt{n}} \sum_{k=1}^n (\xi_k^2 - 1)\right| = O_p(n^{-1/2})\end{aligned} \tag{3.10}$$

as $n \to \infty$ since $\sqrt{n}/\|\xi\| \to 1$ in probability and $n^{-1/2}\sum_{k=1}^n(\xi_k^2 - 1)$ converges to $N(0, 2)$ weakly. For any real numbers $\{\lambda_n; n \ge 1\}$ and any set of random variables $\{H_{ij}; 1 \le i < j \le p\}$, by a triangle inequality,

$$\begin{aligned}&\left|\max_{1\le i<j\le p}\{H_{i,j} + \lambda_n C_{nij}\} - \max_{1\le i<j\le p}\left\{H_{i,j} + \lambda_n \cdot \frac{\sqrt{n}}{\|\xi\|} C_{nij}\right\}\right| \\ &\quad\le \lambda_n \cdot \left|\frac{\sqrt{n}}{\|\xi\|} - 1\right| \cdot \max_{1\le i<j\le p} |C_{nij}|.\end{aligned} \tag{3.11}$$

Note that

$$\max_{1\le i<j\le p} |C_{nij}| \le 2 \cdot \max_{1\le i\le p} \frac{1}{\sqrt{n}}\left|\sum_{k=1}^n \xi_k\xi_{ki}\right|.$$

Observe that $E(\xi_1\xi_{11}) = 0$, that $\mathrm{Var}(\xi_1\xi_{11}) = 1$ and that $E\exp(|\xi_1\xi_{11}|/2) < \infty$. By Lemma 3.4 and the assumption that $\log p = o(n^{1/3})$, we have

$$\begin{aligned}P\Big(\max_{1\le i<j\le p} |C_{nij}| \ge 2A\sqrt{\log p}\Big) &\le p \cdot P\left(\frac{1}{\sqrt{n}}\left|\sum_{k=1}^n \xi_k\xi_{k1}\right| \ge A\sqrt{\log p}\right) \\ &\le p \cdot e^{-A^2(\log p)/3} \to 0\end{aligned} \tag{3.12}$$

as long as $A > \sqrt{3}$. So $\max_{1\le i<j\le p} |C_{nij}| = O_p(\sqrt{\log p})$. This together with (3.10) and (3.11) implies the desired result. Reviewing the arguments above, we see the assertion is still true if "n" is replaced by "m." □

3.2. *Proofs of Theorems* 2.1 *and* 2.2. The key elements in our proofs are a special property for sample correlation matrices under Gaussian assumptions, the Chen–Stein Poisson approximation method, conditioning arguments and a coupling to create independence. To make it clear, we take L_n from Theorem 2.2 to elaborate this point through a few steps:

(a) The special property (Lemma 3.2) for sample correlation matrices allows us to remove $\bar{x}_i$ and $\bar{x}_j$ from the expression $\hat{\rho}_{ij}$ in (1.1); so we get an easier form of the target to work with.

(b) With some efforts, we are able to write

$$L_n = \alpha_n + \beta_n Q_n + \gamma_n R_n, \tag{3.13}$$

where α_n, β_n, γ_n are constants, Q_n goes to $N(0, 1)$, R_n (the quantity M'_n from Proposition 3.2) is the maximum of sums of independent but nonidentically distributed random variables; see (3.46).

(c) We use the Chen–Stein Poisson approximation method to work on R_n. However, due to the strong dependency, we are not able to apply the method directly. In particular, the methods for deriving the limiting distribution of R_n under the assumption $\rho = 0$ in all earlier literature are no longer valid. We will use a conditioning trick. In fact, conditioning on certain event, we obtain the asymptotic distribution of R_n by the Chen–Stein method. After taking the expectation of the conditional probability, we finally derive the limiting distribution of R_n (Proposition 3.2).

(d) We construct R'_n such that it is independent of Q_n in (3.13) by Lemma 3.9 and R'_n has the same asymptotic distribution as that of R_n. Furthermore, we show that the difference between L_n and $L'_n := \alpha_n + \beta_n Q_n + \gamma_n R'_n$ is negligible. So, basically speaking, L_n is reduced to a linear combination of two independent random variables such that one goes to the normal distribution and another goes to the extreme-value distribution.

Now we start to prove the main results. The following notation will be used throughout the rest of the paper. The random variables

$$\{\xi_k, \xi'_k, \xi_{ki}; k, i = 1, 2, \dots\} \text{ are i.i.d. as } N(0, 1). \tag{3.14}$$

Given $\rho_n \in [0, 1)$ for each $n \ge 1$, set

$$a_n = \sqrt{\frac{1-\rho_n}{1+\rho_n}} \quad \text{and} \quad b_n = \sqrt{\frac{\rho_n}{1+\rho_n}}. \tag{3.15}$$

For $x \in \mathbb{R}$ and integer $p \ge 1$, set

$$s_p = \sqrt{4\log p - \log\log p + x}. \tag{3.16}$$

In our theorems, we assume $p \to \infty$, so s_p is well defined as p is large. This clarification will not be repeated in the future. Let ξ_i's be as in (3.14), and write

$$\xi = (\xi_1, \ldots, \xi_n)' \quad \text{and} \quad \|\xi\| = (\xi_1^2 + \cdots + \xi_n^2)^{1/2}. \tag{3.17}$$

To prove Theorem 2.1, we need further notation as follows:

$$\eta_{kij} = a_n \xi_{ki} \xi_{kj} + b_n \xi_k (\xi_{ki} + \xi_{kj}); \tag{3.18}$$

$$M_{nij} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{kij} \tag{3.19}$$

for all $1 \le i < j \le p$. The quantity $\max_{1 \le i < j \le p} M_{nij}$ will serves as a key building block to understand J_n, the largest entry of a sample covariance matrix. It will be used in the proof of Theorem 2.1.

PROPOSITION 3.1. *Let $\rho_n \ge 0$ for each $n \ge 1$ and $\sup_{n \ge 1} \rho_n < 1/2$. Let s_p be as in (3.16). Set $M_n = \max_{1 \le i < j \le p} M_{nij}$. If $p = p_n \to \infty$ and $\log p = o(n^{1/3})$, then*

$$\lim_{n \to \infty} P(M_n \le s_p) = e^{-Ke^{-x/2}}$$

for any $x \in \mathbb{R}$, where $K = \frac{1}{4\sqrt{2\pi}}$.

The proof of Proposition 3.1 will be presented in Section 3.3. Let J_n be as in (2.1). Define

$$W_n = nJ_n = \max_{1 \le i < j \le p} \sum_{k=1}^{n} x_{ki} x_{kj}, \quad n \ge 2. \tag{3.20}$$

PROOF OF THEOREM 2.1. By assumption, $\mu = \mathbf{0}$. Let $\{\xi_k, \xi_{ki}, k, i = 1, 2, \ldots\}$ and $\|\xi\|$ be as in (3.14)–(3.17). Write

$$x_{ki} = \sqrt{\rho_n}\xi_k + \sqrt{1 - \rho_n}\xi_{ki}, \quad 1 \le k \le n, 1 \le i \le p. \tag{3.21}$$

It is easy to check the n rows of the matrix $(x_{ij})_{n \times p}$ are i.i.d. random vectors, $x_{1i} \sim N(0, 1)$ for each $1 \le i \le p$ and $\mathrm{Cov}(x_{1i}, x_{1j}) = \rho_n$ for $1 \le i < j \le p$. That is, each row follows $N_p(\mathbf{0}, \mathbf{R})$. As a result, $\mathbf{X}$ and $(x_{ij})_{n \times p}$ have the same distribution. So we assume $\mathbf{X} = (x_{ij})_{n \times p}$ in the next. Denote

$$\begin{aligned} A_n &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k^2, \\ B_{nij} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_{ki} \xi_{kj}, \\ C_{nij} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k (\xi_{ki} + \xi_{kj}) \end{aligned} \tag{3.22}$$

for all $1 \le i \le j \le p$. Then it follows from the expression (3.21) that

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} x_{ki} x_{kj} = \rho_n A_n + (1-\rho_n) B_{nij} + \sqrt{\rho_n(1-\rho_n)} C_{nij}. \tag{3.23}$$

First, by the central limit theorem, we are able to write

$$A_n = \sqrt{n} + \sqrt{2} U_{n1},$$

where $U_{n1} := \frac{1}{\sqrt{2n}} \sum_{k=1}^{n} (\xi_k^2 - 1) \xrightarrow{d} N(0,1)$. Define

$$\begin{aligned} M_{nij} &:= \frac{(1-\rho_n) B_{nij} + \sqrt{\rho_n(1-\rho_n)} C_{nij}}{\sqrt{1-\rho_n^2}} \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} [a_n \xi_{ki} \xi_{kj} + b_n \xi_k (\xi_{ki} + \xi_{kj})], \end{aligned}$$

where $a_n = \sqrt{\frac{1-\rho_n}{1+\rho_n}}$ and $b_n = \sqrt{\frac{\rho_n}{1+\rho_n}}$. Denote $M_n = \max_{1 \le i < j \le p} M_{nij}$. From these notation, we have

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} x_{ki} x_{kj} = \rho_n \sqrt{n} + \sqrt{2} \rho_n U_{n1} + \sqrt{1-\rho_n^2} M_{nij}, \tag{3.24}$$

and hence

$$\max_{1 \le i < j \le p} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} x_{ki} x_{kj} = \sqrt{n} \rho_n + \sqrt{2} \rho_n U_{n1} + \sqrt{1-\rho_n^2} M_n. \tag{3.25}$$

Review the notation $\xi = (\xi_1, \dots, \xi_n)'$. Define

$$\begin{aligned} \tilde{M}_n &= \max_{1 \le i < j \le p} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left[a_n \xi_{ki} \xi_{kj} + b_n \frac{\sqrt{n}}{\|\xi\|} \xi_k (\xi_{ki} + \xi_{kj}) \right] \\ &= \max_{1 \le i < j \le p} \left\{ H_{i,j} + b_n \cdot \frac{\sqrt{n}}{\|\xi\|} C_{nij} \right\}, \end{aligned}$$

where $H_{i,j} = n^{-1/2} \sum_{k=1}^{n} a_n \xi_{ki} \xi_{kj}$. By Lemma 3.9 and the fact that $0 \le b_n \le 1$,

$$M_n = \tilde{M}_n + O_p\left(\frac{\sqrt{\log p}}{\sqrt{n}} \right). \tag{3.26}$$

This together with (3.25) implies that

$$\begin{aligned} &\max_{1 \le i < j \le p} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} x_{ki} x_{kj} \\ &\quad = \sqrt{n} \rho_n + \sqrt{2} \rho_n U_{n1} + \sqrt{1-\rho_n^2} \tilde{M}_n + O_p\left(\frac{\sqrt{\log p}}{\sqrt{n}} \right). \end{aligned} \tag{3.27}$$

By Eaton [7], Proposition 7.3, or Muirhead [17], Theorem 1.5.6, we know $\|\xi\|$ and $\frac{\xi}{\|\xi\|}$ are independent. Then $U_{n1} = \frac{1}{\sqrt{2n}}(\|\xi\|^2 - n)$ and $\tilde{M}_n$, which is a function of $\frac{\xi}{\|\xi\|}$ and ξ_{ki}'s, are also independent. This is a crucial observation in the following argument.

Now, it follows from Lemma 3.8 and Proposition 3.1 that

$$M_n = 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{1}{4\sqrt{\log p}} U_{n2}, \tag{3.28}$$

where $U_{n2} \xrightarrow{d} \eta$ with distribution function $F_\eta(x) = e^{-\frac{1}{4\sqrt{2\pi}} e^{-\frac{x}{2}}}$ for all $x \in \mathbb{R}$. From (3.26),

$$\tilde{M}_n = 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{1}{4\sqrt{\log p}} U_{n2} + O_p\left(\frac{\sqrt{\log p}}{\sqrt{n}}\right).$$

Then

$$\tilde{U}_{n2} := 4\sqrt{\log p} \cdot \left(\tilde{M}_n - 2\sqrt{\log p} + \frac{\log\log p}{4\sqrt{\log p}}\right) \xrightarrow{d} \eta. \tag{3.29}$$

Since U_{n1} and $\tilde{M}_n$ are independent, U_{n1} and $\tilde{U}_{n2}$ are independent. Reviewing the definition of W_n as in (3.20). Solve $\tilde{M}_n$ from the first identity in (3.29) and then plug it into (3.27) to see

$$\begin{aligned}\frac{1}{\sqrt{n}} W_n - \mu_1 &= \sqrt{2}\rho_n U_{n1} + \frac{\sqrt{1-\rho_n^2}}{4\sqrt{\log p}} \tilde{U}_{n2} + O_p\left(\frac{\sqrt{\log p}}{\sqrt{n}}\right) \\ &= \sqrt{2}\rho_n U_{n1} + \frac{\sqrt{1-\rho_n^2}}{4\sqrt{\log p}} \tilde{U}_{n2} + o_p\left(\frac{1}{\sqrt{\log p}}\right)\end{aligned}$$

by the assumption that $\log p = o(n^{1/3})$, where

$$\mu_1 = \sqrt{n}\rho_n + \left(2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}}\right)\sqrt{1-\rho_n^2}.$$

We now derive the three conclusions by the above relation.

Case (i): $\rho_n\sqrt{\log p} \to 0$. From the Slutsky lemma, it follows that

$$\frac{4\sqrt{\log p}}{\sqrt{1-\rho_n^2}}(n^{-1/2} W_n - \mu_1) \xrightarrow{d} \phi,$$

where ϕ is the extreme-value distribution $F(x) = e^{-Ke^{-\frac{x}{2}}}$ with $K = \frac{1}{4\sqrt{2\pi}}$. The conclusion then follows by the assumption $\rho_n \to 0$ and the Slutsky lemma again.

Case (ii): $\rho_n\sqrt{\log p} \to \lambda \in (0, +\infty)$. By using the independence between U_{n1} and $\tilde{U}_{n2}$ and the Slutsky lemma again, we have

$$\frac{n^{-1/2}W_n - \mu_1}{\sqrt{2}\rho_n} \xrightarrow{d} \xi + \lambda_0\phi,$$

where $\xi \sim N(0, 1)$, $\lambda_0 = \frac{1}{4\sqrt{2}\lambda}$ and ϕ is as in case (i) and ϕ is independent of ξ.

Case (iii): $\rho_n\sqrt{\log p} \to \infty$. By the Slutsky lemma,

$$\frac{n^{-1/2}W_n - \mu_1}{\sqrt{2}\rho_n} \xrightarrow{d} N(0, 1).$$

The proof is completed by using (3.20). $\square$

Now we proceed to prove Theorem 2.2. Recall the notation from (3.14) and (3.15). Define $\sigma_{n1}^2 = (1-\rho_n)^2 + 2\rho_n a_n^2$,

$$a'_n = \frac{a_n}{\sigma_{n1}} \quad \text{and} \quad b'_n = \frac{(1-\rho_n)b_n}{\sigma_{n1}} \tag{3.30}$$

for $1 \le k \le m := n-1$. With these notation, we further define

$$\eta'_{kij} = a'_n\left[\xi_{ki}\xi_{kj} - \frac{\rho_n}{2}(\xi_{ki}^2 + \xi_{kj}^2 - 2)\right] + b'_n\xi_k(\xi_{ki} + \xi_{kj}),$$

$$M'_{nij} = \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{kij},$$

for $k = 1, 2, \dots, m$.

PROPOSITION 3.2. *Set* $M'_n = \max_{1\le i<j\le p} M'_{nij}$. *Let* $\rho_n \in (0, 1)$ *for each* $n \ge 1$. *Let* s_p *be as in* (3.16). *If* $p \to \infty$ *and* $\log p = o(n^{1/3})$, *then*

$$\lim_{n\to\infty} P(M'_n \le s_p) = e^{-Ke^{-x/2}}$$

for any $x \in \mathbb{R}$, *where* $K = \frac{1}{4\sqrt{2\pi}}$.

The proof of the above conclusion is arranged at Section 3.4. The major contribution of L_n in Theorem 2.2 comes from the left-hand side of (3.31) next, which is asymptotically the sum of two random variables well understood earlier.

LEMMA 3.10. *Let* $\rho_n \in [0, 1)$ *for all* $n \ge 1$. *Recall the notation in* (3.14) *and* (3.15). *Define* $x_{ki} = \sqrt{\rho_n}\xi_k + \sqrt{1-\rho_n}\xi_{ki}$ *for* $1 \le k \le m$ *and* $1 \le i \le p$, *where*

$m = n-1$. *Assume* $\log p = o(n^{\frac{1}{3}})$ *as* $n \to \infty$. *Then*

$$\begin{aligned}
&\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m} x_{ki}x_{kj}\right)\cdot\left[1-\frac{1}{4m}\sum_{k=1}^{m}\left(x_{ki}^2+x_{kj}^2\right)\right] \\
&\quad = \frac{1}{2}\rho_n\sqrt{m}+\frac{1}{\sqrt{2}}\rho_n(1-\rho_n)U_{m1}+\frac{1}{2}\sqrt{1-\rho_n^2}\cdot\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\psi_{kij}+\Delta_{nij},
\end{aligned} \tag{3.31}$$

where $U_{m1} = \frac{1}{\sqrt{2m}}\sum_{k=1}^{m}(\xi_k^2-1)$,

$$\psi_{kij} = a_n\left[\xi_{ki}\xi_{kj} - \frac{\rho_n}{2}\left(\xi_{ki}^2+\xi_{kj}^2-2\right)\right] + (1-\rho_n)b_n\xi_k(\xi_{ki}+\xi_{kj})$$

and, *as* $n \to \infty$,

$$\max_{1\le i<j\le p}|\Delta_{nij}| = O_p\left(\frac{\log p}{\sqrt{m}}\right). \tag{3.32}$$

PROOF. Define

$$M_{mij} = \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta_{kij}, \quad 1\le i\le j\le p,$$

where $\eta_{kij} = a_n\xi_{ki}\xi_{kj} + b_n\xi_k(\xi_{ki}+\xi_{kj})$, $a_n = \sqrt{\frac{1-\rho_n}{1+\rho_n}}$ and $b_n = \sqrt{\frac{\rho_n}{1+\rho_n}}$. From (3.24), we have

$$\frac{1}{\sqrt{m}}\sum_{k=1}^{m}x_{ki}x_{kj} = \rho_n\sqrt{m}+\sqrt{2}\rho_n U_{m1}+\sqrt{1-\rho_n^2}M_{mij},$$

where $U_{m1} = \frac{1}{\sqrt{2m}}\sum_{k=1}^{m}(\xi_k^2-1)$. In particular,

$$M_{mii} = \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\left(a_n\xi_{ki}^2+2b_n\xi_k\xi_{ki}\right).$$

We can write

$$\begin{aligned}
&\frac{1}{4m}\sum_{k=1}^{m}\left(x_{ki}^2+x_{kj}^2\right) \\
&\quad = \frac{1}{4\sqrt{m}}(2\rho_n\sqrt{m}+2\sqrt{2}\rho_n U_{m1}) + \frac{1}{4\sqrt{m}}\sqrt{1-\rho_n^2}(M_{mii}+M_{mjj}) \\
&\quad = \frac{1}{2}\rho_n + \frac{1}{2}a_n\sqrt{1-\rho_n^2} + \frac{\rho_n}{\sqrt{2m}}U_{m1} + \frac{1}{4\sqrt{m}}\sqrt{1-\rho_n^2}T_{mij} \\
&\quad = \frac{1}{2} + \frac{\rho_n}{\sqrt{2m}}U_{m1} + \frac{1}{4\sqrt{m}}\sqrt{1-\rho_n^2}T_{mij},
\end{aligned}$$

where

$$T_{mij} = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \big[a_n(\xi_{ki}^2 + \xi_{kj}^2 - 2) + 2b_n \xi_k(\xi_{ki} + \xi_{kj})\big]. \tag{3.33}$$

So the product on the left-hand side of (3.31) is equal to

$$\begin{aligned} &\Big(\rho_n\sqrt{m} + \sqrt{2}\rho_n U_{m1} + \sqrt{1-\rho_n^2} M_{mij}\Big) \\ &\quad \cdot \left(\frac{1}{2} - \frac{\rho_n}{\sqrt{2m}} U_{m1} - \frac{1}{4\sqrt{m}}\sqrt{1-\rho_n^2} T_{mij}\right) \\ &= \frac{1}{2}\rho_n\sqrt{m} + \frac{1}{\sqrt{2}}\rho_n(1-\rho_n)U_{m1} + \frac{1}{2}\sqrt{1-\rho_n^2} M_{mij} \\ &\quad - \frac{\rho_n}{4}\sqrt{1-\rho_n^2} T_{mij} + \Delta_{nij}, \end{aligned}$$

where

$$\begin{aligned} \Delta_{nij} = &-\frac{\rho_n^2}{\sqrt{m}} U_{m1}^2 - \frac{\rho_n\sqrt{1-\rho_n^2}}{\sqrt{2m}}(U_{m1}M_{mij}) \\ &- \frac{1}{4\sqrt{m}}\sqrt{1-\rho_n^2} T_{mij}\Big(\sqrt{2}\rho_n U_{m1} + \sqrt{1-\rho_n^2} M_{mij}\Big). \end{aligned}$$

Observe that

$$\frac{1}{2}\sqrt{1-\rho_n^2} M_{mij} - \frac{\rho_n}{4}\sqrt{1-\rho_n^2} T_{mij} = \frac{1}{2}\sqrt{1-\rho_n^2} \cdot \frac{1}{\sqrt{m}} \sum_{k=1}^{n} \psi_{kij},$$

where

$$\psi_{kij} = a_n\left[\xi_{ki}\xi_{kj} - \frac{\rho_n}{2}(\xi_{ki}^2 + \xi_{kj}^2 - 2)\right] + (1-\rho_n) b_n \xi_k(\xi_{ki} + \xi_{kj}).$$

Use the trivial bound $1 - \rho_n^2 \le 1$ to see

$$\max_{1 \le i < j \le p} |\Delta_{nij}| \le \frac{\rho_n^2}{\sqrt{m}} U_{m1}^2 + \frac{\rho_n}{\sqrt{m}}(M_m + T_m)|U_{m1}| + \frac{1}{\sqrt{m}}(M_m T_m), \tag{3.34}$$

where

$$M_m = \max_{1 \le i < j \le p} |M_{mij}| \quad \text{and} \quad T_m = \max_{1 \le i < j \le p} |T_{mij}|.$$

From Proposition 3.1, we know

$$\frac{M_m}{\sqrt{\log p}} \to 2 \tag{3.35}$$

in probability. Now, from (3.33) we have

$$
\begin{aligned}
T_m &\le \max_{1\le i<j\le p} \frac{1}{\sqrt{m}}\left|\sum_{k=1}^m (\xi_{ki}^2+\xi_{kj}^2-2)\right| \\
&\quad + \max_{1\le i<j\le p} \frac{1}{\sqrt{m}}\left|\sum_{k=1}^m \xi_k(\xi_{ki}+\xi_{kj})\right| \\
&\le 2\max_{1\le i\le p} \frac{1}{\sqrt{m}}\left|\sum_{k=1}^m (\xi_{ki}^2-1)\right| + 2\max_{1\le i<j\le p} \frac{1}{\sqrt{m}}\left|\sum_{k=1}^m \xi_k\xi_{ki}\right| \\
&:= 2I_n + 2I_n'.
\end{aligned} \tag{3.36}
$$

Let $\zeta_k = (\xi_{k1}^2-1)/\sqrt{2}$ for $1\le k\le m$. Then $E\zeta_k=0$, $\mathrm{Var}(\zeta_k)=1$ and $Ee^{|\zeta_1|/2}<\infty$. By Lemma 3.4, from assumption $\sqrt{\log p}=o(n^{1/3})$, we see that

$$
\begin{aligned}
P(I_n \ge 2A_2\sqrt{\log p}) &\le p\cdot P\left(\frac{1}{\sqrt{m}}\left|\sum_{k=1}^m \zeta_k\right| \ge A_2\sqrt{\log p}\right) \\
&\le p\cdot e^{-A_2^2(\log p)/3} \to 0
\end{aligned} \tag{3.37}
$$

as long as $A_2>\sqrt{3}$. So $I_n = O_p(\sqrt{\log p})$. Furthermore, notice $E(\xi_1\xi_{11})=0$, $\mathrm{Var}(\xi_1\xi_{11})=1$ and $E\exp(\frac{1}{2}|\xi_1\xi_{11}|)<\infty$. By the same argument as obtaining (3.37), we have $I_n' = O_p(\sqrt{\log p})$. In summary, $T_m = O_p(\sqrt{\log p})$. This together with (3.35) and the fact $U_{m1}\to N(0,1)$ implies that

$$
\max_{1\le i<j\le p} |\Delta_{nij}| = O_p\left(\frac{\log p}{\sqrt{m}}\right)
$$

by using (3.34). We then get (3.32). □

PROOF OF THEOREM 2.2. As explained at the beginning of the proof of Lemma 3.2, without loss of generality, we assume $\mu=\mathbf{0}$.

Let $\{\xi_k, \xi_{ki}, k, i = 1, 2, \dots\}$ be as in (3.14) and (3.15). As before, $p=p_n$. Define

$$
x_{ki} = \sqrt{\rho_n}\xi_k + \sqrt{1-\rho_n}\xi_{ki}, \quad 1\le k\le n-1, 1\le i\le p.
$$

Review the beginning of the proof of Theorem 2.1, we know that the $n-1$ rows of the matrix $(x_{ij})_{(n-1)\times p}$ are i.i.d. random vectors, each of which follows $N_p(\mathbf{0},\mathbf{R})$. Write $(x_{ij})_{(n-1)\times p} = (V_1,\dots,V_p)$ such that, for each $1\le j\le p$, $V_j = (x_{1j},\dots,x_{n-1,j})'$. By Lemma 3.2, we have

$$
\sqrt{n-1}\max_{1\le i<j\le p}\hat{\rho}_{ij} \stackrel{d}{=} \max_{1\le i<j\le p} \frac{\frac{1}{\sqrt{n-1}}\sum_{k=1}^{n-1} x_{ki}x_{kj}}{\sqrt{\frac{1}{n-1}\sum_{k=1}^{n-1} x_{ki}^2}\sqrt{\frac{1}{n-1}\sum_{k=1}^{n-1} x_{kj}^2}}.
$$

Denote $m = n-1$, $h_i = \sqrt{m^{-1}\sum_{k=1}^m x_{ki}^2}$ and

$$\Lambda_{nij} = \frac{\frac{1}{\sqrt{m}}\sum_{k=1}^m x_{ki}x_{kj}}{h_i h_j}. \tag{3.38}$$

So it suffices to prove the statements (i), (ii) and (iii) with "$\sqrt{n-1}L_n$" replaced by "$\max_{1\le i<j\le p}\Lambda_{nij}$". Our proof consists of a few of steps.

Step 1: *Reduction of L_n to a simple form*. Write $h_i^{-1} = (1+\zeta_{ni}/\sqrt{m})^{-1/2}$, where $\zeta_{ni} := m^{-1/2}\sum_{k=1}^m (x_{ki}^2 - 1)$. By the Taylor expansion, there exists $\delta \in (0,1)$ such that $(1+x)^{-1/2} = 1 - x/2 + \phi(x)$, where $|\phi(x)| \le x^2$ for all $x \in [-\delta, \delta]$. It follows that

$$\begin{aligned}\frac{1}{h_i h_j} &= \left[1 - \frac{\zeta_{ni}}{2\sqrt{m}} + \phi\left(\frac{\zeta_{ni}}{\sqrt{m}}\right)\right]\cdot\left[1 - \frac{\zeta_{nj}}{2\sqrt{m}} + \phi\left(\frac{\zeta_{nj}}{\sqrt{m}}\right)\right] \\ &= 1 - \frac{\zeta_{ni}}{2\sqrt{m}} - \frac{\zeta_{nj}}{2\sqrt{m}} + \varepsilon_{ij},\end{aligned} \tag{3.39}$$

where

$$\begin{aligned}\varepsilon_{ij} = \frac{\zeta_{ni}\zeta_{nj}}{4m} &+ \left(1 - \frac{\zeta_{ni}}{2\sqrt{m}}\right)\cdot\phi\left(\frac{\zeta_{nj}}{\sqrt{m}}\right) \\ &+ \left(1 - \frac{\zeta_{nj}}{2\sqrt{m}}\right)\cdot\phi\left(\frac{\zeta_{ni}}{\sqrt{m}}\right) + \phi\left(\frac{\zeta_{ni}}{\sqrt{m}}\right)\phi\left(\frac{\zeta_{nj}}{\sqrt{m}}\right).\end{aligned}$$

Obviously, if $|\zeta_{ni}/\sqrt{m}| < \delta$ and $|\zeta_{nj}/\sqrt{m}| < \delta$, then $\max_{k=i,j}|1-\zeta_{nk}/(2\sqrt{m})| < 2$ because $\delta \in (0,1)$, and hence

$$|\varepsilon_{ij}| \le \frac{|\zeta_{ni}|\cdot|\zeta_{nj}|}{4m} + \frac{2\zeta_{ni}^2}{m} + \frac{2\zeta_{nj}^2}{m} + \frac{\zeta_{ni}^2}{m}\cdot\frac{\zeta_{nj}^2}{m} \le \frac{4(\zeta_{ni}^2 + \zeta_{nj}^2)}{m}.$$

This gives that

$$\max_{1\le i<j\le p}|\varepsilon_{ij}| \le \frac{8}{m}\cdot\max_{1\le i\le m}\zeta_{ni}^2 \tag{3.40}$$

provided $\max_{1\le i\le p}|\zeta_{ni}/\sqrt{m}| < \delta$. Let $\zeta_k = (\xi_{k1}^2 - 1)/\sqrt{2}$ for $1 \le k \le m$. Then $E\zeta_k = 0$, $\mathrm{Var}(\zeta_k) = 1$ and $Ee^{|\zeta_1|/2} < \infty$. By assumption, $(x_{1i}, x_{2i}, \ldots, x_{mi}) \stackrel{d}{=} (\xi_1, \xi_2, \ldots, \xi_k)$ for each $1 \le i \le p$. Set

$$\Omega_n = \left\{\max_{1\le i\le p}|\zeta_{ni}| < 3\sqrt{\log p}\right\}.$$

Then it follows from (3.37) that

$$\lim_{n\to\infty} P(\Omega_n) = 1. \tag{3.41}$$

Now we see from (3.38) and (3.39) that

$$
\begin{aligned}
\Lambda_{nij} &= \left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m} x_{ki}x_{kj}\right)\cdot\left(1-\frac{\zeta_{ni}}{2\sqrt{m}}-\frac{\zeta_{nj}}{2\sqrt{m}}+\varepsilon_{ij}\right)\\
&= \left(\frac{2}{\sqrt{m}}\sum_{k=1}^{m} x_{ki}x_{kj}\right)\cdot\left[1-\frac{1}{4m}\left(\sum_{k=1}^{m} x_{ki}^2+\sum_{k=1}^{m} x_{kj}^2\right)\right]+\varepsilon'_{ij},
\end{aligned}
\tag{3.42}
$$

where

$$
\varepsilon'_{ij} := \left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m} x_{ki}x_{kj}\right)\cdot\varepsilon_{ij}.
$$

Proceed to estimate $\max_{1\le i<j\le p} m^{-1/2}|\sum_{k=1}^{m} x_{ki}x_{kj}|$. To this end, (3.2) implies that

$$
\begin{aligned}
&\max_{1\le i<j\le p}\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m} x_{ki}x_{kj}\right|\\
&\quad\le \sqrt{m}+|U_{m1}|+\max_{1\le i<j\le p}\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m}\left[a_n\xi_{ki}\xi_{kj}+b_n\xi_k(\xi_{ki}+\xi_{kj})\right]\right|\\
&\quad\le \sqrt{m}+|U_{m1}|+\max_{1\le i<j\le p}\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m}\xi_{ki}\xi_{kj}\right|+2\max_{1\le i\le p}\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m}\xi_k\xi_{ki}\right|,
\end{aligned}
$$

where $U_{m1}=(2m)^{-1/2}\sum_{k=1}^{m}(\xi_k^2-1)$. Observe that the last two maxima above have the same distribution. By the estimate of I'_n from (3.36), each of them has size $O_p(\sqrt{\log p})$. Using the assumption $\log p = o(n^{1/3})$, we see

$$
\Upsilon_n := \max_{1\le i<j\le p}\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m} x_{ki}x_{kj}\right| = O_p(\sqrt{n})
$$

as $n\to\infty$. Therefore, by (3.40),

$$
\max_{1\le i<j\le p}|\varepsilon'_{ij}| \le \Upsilon_n\cdot\max_{1\le i<j\le p}|\varepsilon_{ij}| \le \frac{8}{m}\cdot\Upsilon_n\cdot\max_{1\le i\le m}\zeta_{ni}^2
$$

provided $\max_{1\le i<j\le p}|\frac{\zeta_{ni}}{\sqrt{m}}|<\delta$. By assumption, $\frac{3\sqrt{\log p}}{\sqrt{m}}\to 0$. This enables us to see

$$
I_{\Omega_n}\cdot\max_{1\le i<j\le p}|\varepsilon'_{ij}| = \frac{8}{m}\cdot O_p(\sqrt{n})\cdot(3\sqrt{\log p})^2 = O\left(\frac{\log p}{\sqrt{n}}\right).
\tag{3.43}
$$

By Lemma 3.10 and (3.42),

$$\begin{aligned}\Lambda_{nij} &= \rho_n\sqrt{m} + \sqrt{2}\rho_n(1-\rho_n)U_{m1} \\ &\quad + \sqrt{1-\rho_n^2}\cdot\frac{1}{\sqrt{m}}\sum_{k=1}^{n}\psi_{kij} + 2\Delta_{nij} + \varepsilon'_{ij} \\ &= \rho_n\sqrt{m} + \sqrt{2}\rho_n(1-\rho_n)U_{m1} \\ &\quad + \sigma_{n1}\sqrt{1-\rho_n^2}\cdot\frac{1}{\sigma_{n1}\sqrt{m}}\sum_{k=1}^{n}\psi_{kij} + \varepsilon''_{ij},\end{aligned} \tag{3.44}$$

where ψ_{kij} and Δ_{nij} are defined as in the lemma, $\varepsilon''_{ij} = 2\Delta_{nij} + \varepsilon'_{ij}$ and $\sigma_{n1}^2 = (1-\rho_n)^2 + 2\rho_n a_n^2$. Easily

$$I_{\Omega_n}\cdot\max_{1\le i<j\le p}|\varepsilon''_{ij}| \le 2\cdot\max_{1\le i<j\le p}|\Delta_{nij}| + I_{\Omega_n}\cdot\max_{1\le i<j\le p}|\varepsilon'_{ij}| = O_p\left(\frac{\log p}{\sqrt{n}}\right)$$

by (3.32) and (3.43). Let $f(i,j)$ and $g(i,j)$ be real functions defined on $\{(i,j); 1\le i<j\le m\}$. It is easy to see that

$$\left|\max_{1\le i<j\le p} f(i,j) - \max_{1\le i<j\le p} g(i,j)\right| \le \max_{1\le i<j\le p}|f(i,j)-g(i,j)|.$$

Therefore, from (3.44) we have

$$\begin{aligned}&I_{\Omega_n}\cdot\max_{1\le i<j\le p}\Lambda_{nij} \\ &\quad = I_{\Omega_n}\cdot[\rho_n\sqrt{m} + \sqrt{2}\rho_n(1-\rho_n)U_{m1}] \\ &\qquad + \sigma_{n1}\sqrt{1-\rho_n^2}\cdot\max_{1\le i<j\le p}\left\{\frac{1}{\sigma_{n1}\sqrt{m}}\cdot\sum_{k=1}^{n}\psi_{kij}\right\}\cdot I_{\Omega_n} + O_p\left(\frac{\log p}{\sqrt{n}}\right).\end{aligned} \tag{3.45}$$

Observe that the last maximum is exactly M'_n appeared in Proposition 3.2. Writing $I_{\Omega_n} = 1 - I_{\Omega_n^c}$, we eventually get

$$\begin{aligned}\max_{1\le i<j\le p}\Lambda_{nij} &= \rho_n\sqrt{m} + \sqrt{2}\rho_n(1-\rho_n)U_{m1} \\ &\quad + \sigma_{n1}\sqrt{1-\rho_n^2}M'_n + O_p\left(\frac{\log p}{\sqrt{m}}\right) + I_{\Omega_n^c}\cdot\Psi_n,\end{aligned} \tag{3.46}$$

for some random variable Ψ_n.

Step 2: *Asymptotic independence between* U_{m1} *and* M'_n. Review the definition of ψ_{kij} in Lemma 3.10. Set

$$\tilde{\eta}'_{kij} = a_n\left[\xi_{ki}\xi_{kj} - \frac{\rho_n}{2}(\xi_{ki}^2 + \xi_{kj}^2 - 2)\right] + (1-\rho_n)b_n\frac{\sqrt{m}}{\|\xi\|}\xi_k(\xi_{ki}+\xi_{kj}),$$

$$\tilde{M}'_n = \max_{1\le i<j\le p}\left\{\frac{1}{\sigma_{n1}\sqrt{m}}\cdot\sum_{k=1}^{m}\tilde{\eta}'_{kij}\right\}.$$

From the definitions of a_n' and b_n' in (3.30), it is easy to see that $(1+\rho_n^2)a_n'^2 + 2b_n'^2 = 1$. Since $b_n' = \frac{(1-\rho_n)b_n}{\sigma_{n1}}$, we get $|\frac{(1-\rho_n)b_n}{\sigma_{n1}}| \le \frac{1}{2}$. By Lemma 3.9,

$$\tilde{M}_n' - M_n' = O_p\left(\frac{(1-\rho_n)b_n}{\sigma_{n1}} \cdot \frac{\sqrt{\log p}}{\sqrt{n}}\right) = O_p\left(\frac{\sqrt{\log p}}{\sqrt{n}}\right). \tag{3.47}$$

By Lemma 3.8 and Proposition 3.2,

$$M_n' = 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{1}{4\sqrt{\log p}} V_n,$$

where $V_n \xrightarrow{d} \phi$ with distribution function $F(x) = e^{-Ke^{-\frac{x}{2}}}$ for all $x \in \mathbb{R}$, where $K = \frac{1}{4\sqrt{2\pi}}$. The above two assertions tell us that

$$\tilde{M}_n' = 2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{1}{4\sqrt{\log p}} V_n + o_p\left(\frac{1}{\sqrt{\log p}}\right).$$

Then

$$\tilde{U}_{n2} := 4\sqrt{\log p} \cdot \left(\tilde{M}_n' - 2\sqrt{\log p} + \frac{\log\log p}{4\sqrt{\log p}}\right) \xrightarrow{d} \phi. \tag{3.48}$$

Since $U_{m1} = (2m)^{-1/2}\sum_{k=1}^m (\xi_k^2 - 1)$ and $\tilde{M}_n'$ are independent by the same argument as that after (3.27), U_{m1} and $\tilde{U}_{n2}$ are independent. Evidently,

$$\begin{aligned}\sigma_{n1}\sqrt{1-\rho_n^2} &= \left((1-\rho_n)^2 + 2\rho_n \cdot \frac{1-\rho_n}{1+\rho_n}\right)^{1/2} \cdot \sqrt{1-\rho_n^2} \\ &= (1-\rho_n) \cdot \sqrt{1+2\rho_n-\rho_n^2}.\end{aligned}$$

In particular, $\sigma_{n1}\sqrt{1-\rho_n^2} \le 2$. Combining (3.46), (3.47) and (3.48), we obtain

$$\begin{aligned}&\max_{1\le i<j\le p} \Lambda_{nij} - \rho_n\sqrt{m} - \sqrt{2}\rho_n(1-\rho_n)U_{m1} \\ &= \sigma_{n1}\sqrt{1-\rho_n^2}\tilde{M}_n' + O_p\left(\frac{\log p}{\sqrt{n}}\right) + I_{\Omega_n^c} \cdot \Psi_n \\ &= \sigma_{n1}\sqrt{1-\rho_n^2}\left(2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}} + \frac{\tilde{U}_{n2}}{4\sqrt{\log p}}\right) + O_p\left(\frac{\log p}{\sqrt{n}}\right) + I_{\Omega_n^c} \cdot \Psi_n.\end{aligned}$$

Set

$$\mu_2 = \rho_n\sqrt{m} + (1-\rho_n) \cdot \sqrt{1+2\rho_n-\rho_n^2} \cdot \left(2\sqrt{\log p} - \frac{\log\log p}{4\sqrt{\log p}}\right).$$

Then

$$\begin{aligned}\max_{1\le i<j\le p} \Lambda_{nij} - \mu_2 &= \sqrt{2}\rho_n(1-\rho_n)U_{m1} + (1-\rho_n) \cdot \sqrt{1+2\rho_n-\rho_n^2} \cdot \frac{\tilde{U}_{n2}}{4\sqrt{\log p}} \\ &\quad + o_p\left(\frac{1}{\sqrt{\log p}}\right) + I_{\Omega_n^c} \cdot \Psi_n,\end{aligned}$$

where the equality $O_p(\log p/\sqrt{m}) = o_p(1/\sqrt{\log p})$ holds due to the assumption $\log p = o(n^{1/3})$. Notice that $P(|I_{\Omega_n^c} \cdot \Psi_n| \cdot \sqrt{\log p} \geq \varepsilon) \leq P(\Omega_n^c) \to 0$ for any $\varepsilon > 0$ by (3.41), hence $I_{\Omega_n^c} \cdot \Psi_n = o_p(1/\sqrt{\log p})$. It follows that

$$\max_{1\leq i<j\leq p} \Lambda_{nij} - \mu_2 = \sqrt{2}\rho_n(1-\rho_n)U_{m1} + (1-\rho_n)\cdot\sqrt{1+2\rho_n-\rho_n^2}\cdot\frac{\tilde{U}_{n2}}{4\sqrt{\log p}} + o_p\left(\frac{1}{\sqrt{\log p}}\right).$$

Step 3: *Derivation of conclusions* (i), (ii) *and* (iii). Recall the assumption that $\rho_n \geq 0$ for each $n \geq 1$ and $\sup_{n\geq 1}\rho_n < 1$.

Case (i): $\rho_n\sqrt{\log p} \to 0$. For this case, by the Slutsky lemma,

$$\frac{4}{(1-\rho_n)\cdot\sqrt{1+2\rho_n-\rho_n^2}}\cdot\sqrt{\log p}\cdot\Big(\max_{1\leq i<j\leq p}\Lambda_{nij}-\mu_2\Big) \xrightarrow{d} \phi,$$

where ϕ has distribution function $F(x) = e^{-Ke^{-\frac{x}{2}}}$ with $K = \frac{1}{4\sqrt{2\pi}}$. The conclusion follows by the assumption $\rho_n \to 0$ and the Slutsky lemma again.

Case (ii): $\rho_n\sqrt{\log p} \to \lambda \in (0,\infty)$. By the Slutsky lemma and independence,

$$\frac{\max_{1\leq i<j\leq p}\Lambda_{nij}-\mu_2}{\sqrt{2}\rho_n(1-\rho_n)} \xrightarrow{d} \xi + \lambda_0\phi,$$

where $\xi \sim N(0,1)$, $\lambda_0 = \frac{1}{4\sqrt{2}\lambda}$ and ϕ is the same as in case (i) and ϕ is independent of ξ. The conclusion is yielded by the assumption $\rho_n \to 0$ and the Slutsky lemma again.

Case (iii): $\rho_n\sqrt{\log p} \to \infty$. In this situation, by the Slutsky lemma,

$$\frac{\max_{1\leq i<j\leq p}\Lambda_{nij}-\mu_2}{\sqrt{2}\rho_n(1-\rho_n)} = U_{m1} + o_p(1) \xrightarrow{d} N(0,1).$$

The proof is complete. □

3.3. *The proof of Proposition* 3.1. In this section, we will use the Chen–Stein Poisson approximation method (Lemma 3.3) with conditioning argument to get the asymptotical distribution of M_n defined in Proposition 3.1. The conditioning procedure is employed because Lemma 3.3 asks more independence than what the structure behind M_n has.

LEMMA 3.11. *Recall the notation in* (3.14)–(3.16). *Assume* $\rho_n \geq 0$ *for all* $n \geq 1$ *and* $\sup_{n\geq 1}\rho_n < 1/2$. *Define* $Z_n = n^{-1/2}b_n\sum_{k=1}^n \xi_k\xi_k'$. *If* $p = p_n \to \infty$ *and* $\log p = o(n^{1/3})$ *as* $n \to \infty$, *then there exists a constant* $\delta \in (0,1)$ *such that*

$$E\exp\left[-\frac{1+\rho_n}{1+\delta}(Z_n - s_p)^2\right] = o\left(\frac{1}{p^3}\right)$$

as $n \to \infty$.

PROOF. If $\rho_n = 0$ for some $n \ge 1$, then $Z_n = 0$ and the expectation in the lemma is identical to $\exp(-(4\log p - \log\log p + x)/(1+\delta))$, which, by taking $\delta \in (0, 1)$ small enough, is bounded by $p^{-3.5}$ as n is sufficiently large. Therefore, to prove the lemma, w.l.o.g., we assume $\rho_n > 0$ for all $n \ge 1$.

First, we show

$$Ee^{-\alpha_1(\xi_1-\beta_1)^2} = \frac{1}{\sqrt{2\alpha_1+1}}\exp\left(-\frac{\alpha_1\beta_1^2}{2\alpha_1+1}\right) \tag{3.49}$$

for any $\alpha_1 > 0$ and $\beta_1 \in \mathbb{R}$. In fact

$$Ee^{-\alpha_1(\xi_1-\beta_1)^2} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\alpha_1(x-\beta_1)^2-\frac{x^2}{2}}\,dx.$$

Write

$$-\alpha_1(x-\beta_1)^2 - \frac{x^2}{2} = -\left(\sqrt{\alpha_1+\frac{1}{2}}x - \frac{\alpha_1\beta_1}{\sqrt{\alpha_1+\frac{1}{2}}}\right)^2 - \frac{\alpha_1\beta_1^2}{2\alpha_1+1}.$$

Now, define y such that

$$\frac{y}{\sqrt{2}} = \sqrt{\alpha_1+\frac{1}{2}}x - \frac{\alpha_1\beta_1}{\sqrt{\alpha_1+\frac{1}{2}}}.$$

It follows that

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\alpha_1(x-\beta_1)^2-\frac{x^2}{2}}\,dx = e^{-\frac{\alpha_1\beta_1^2}{2\alpha_1+1}}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}\,dy\cdot\frac{1}{\sqrt{2\alpha_1+1}}.$$

Thus, (3.49) holds.

Recall the notation (3.17). By Eaton [7], Proposition 7.3, or Muirhead [17], Theorem 1.5.6, we know $\|\xi\|$ and $\xi/\|\xi\|$ are independent. Also, $\|\xi\|^{-1}\sum_{k=1}^n \xi_k\xi_k' \sim N(0, 1)$ by independence. Consequently,

$$Z_n \stackrel{d}{=} b_n\cdot\frac{\|\xi\|}{\sqrt{n}}\cdot\xi_1'.$$

In particular, $\|\xi\|/\sqrt{n}$ and ξ_1' are independent. Let $\tau = \frac{1+\delta}{1+\rho_n}$. Observe

$$E\exp\left[-\frac{(Z_n-s_p)^2}{\tau}\right] = Ee^{-\alpha_1(\xi_1'-\beta_1)^2} = E\big[E_1 e^{-\alpha_1(\xi_1'-\beta_1)^2}\big], \tag{3.50}$$

where E_1 stands for the conditional expectation given $\|\xi\|$,

$$\alpha_1 = \frac{b_n^2\|\xi\|^2}{n\tau} \quad\text{and}\quad \beta_1 = \frac{\sqrt{n}s_p}{b_n\|\xi\|}.$$

By using (3.49), we obtain

$$E_1 e^{-\alpha_1(\xi_1'-\beta_1)^2} \le \exp\left(-\frac{s_p^2}{\tau + 2b_n^2\frac{\|\xi\|^2}{n}}\right) \le \exp\left\{-\frac{s_p^2}{(1+\delta)[(1+\rho_n)^{-1}+2b_n^2]}\right\} \tag{3.51}$$

if $\|\xi\|^2/n < 1+\delta$. Observe that $(1+\rho_n)^{-1} + 2b_n^2 = \frac{1+2\rho_n}{1+\rho_n} \le \frac{1+2\rho}{1+\rho} < 4/3$ for all $n \ge 1$, where $\rho := \sup_{n\ge 1}\rho_n < 1/2$ by assumption. Take $\delta \in (0,1)$ such that $\theta := (1+\delta)\frac{1+2\rho}{1+\rho} < 4/3$. Hence, given $\|\xi\|^2/n < 1+\delta$,

$$E_1 e^{-\alpha_1(\xi_1'-\beta_1)^2} \le \frac{(\log p)^{5/\theta}}{p^{4/\theta}}$$

as n is sufficiently large. By the large deviations for i.i.d. random variables, there exists a constant $C > 0$ depending on τ only such that $P(\frac{\|\xi\|^2}{n} \ge 1+\delta) < e^{-nC}$ for all $n \ge 1$. Combining the above inequality, (3.50) and (3.51), we arrive at

$$\begin{aligned} &E\exp\left[-\frac{(Z_n - s_p)^2}{\tau}\right] \\ &\quad = E\left[E_1 e^{-\alpha_1(\xi_1'-\beta_1)^2} I\left(\frac{\|\xi\|^2}{n} < 1+\delta\right)\right] + P\left(\frac{\|\xi\|^2}{n} \ge 1+\delta\right) \\ &\quad \le \frac{(\log p)^{5/\theta}}{p^{4/\theta}} + e^{-nC} = o\left(\frac{1}{p^3}\right), \end{aligned}$$

where the last equality follows from the assumption $\log p = o(n^{1/3})$. □

PROOF OF PROPOSITION 3.1. In the next section, we will assume p is large enough such that $s_p > 0$. Set $I = \{(i,j); 1 \le i < j \le p\}$. For $\alpha = (i,j) \in I$, define $X_\alpha = M_{nij}$ and

$$B_\alpha = \{(k,l) \in I; \text{either } k \in \{i,j\} \text{ or } l \in \{i,j\}, \text{but } (k,l) \ne \alpha\}.$$

Let P_2 and E_2 stand for the conditional probability and the conditional expectation given $\{\xi_k; 1 \le k \le n\}$, respectively. The crucial point is that, given $\{\xi_k; 1 \le k \le n\}$, random variable X_α is independent of $\{X_\beta; \beta \notin B_\alpha\}$. Since $\{X_\alpha, \alpha \in I\}$ are identically distributed under P_2, by Lemma 3.3, we have

$$\left|P_2\left(\max_{\alpha\in I} X_\alpha \le s_p\right) - e^{-\lambda_{p1}}\right| \le w_1 + w_2, \tag{3.52}$$

where

$$\lambda_{p1} = \frac{p(p-1)}{2} P_2\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right)$$

and

$$w_1 = \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} P_2(X_\alpha > s_p) P_2(X_\beta > s_p)$$
$$\leq \frac{p(p-1)}{2} \cdot (2p) \cdot P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)^2$$

and

$$w_2 = \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} P_2(X_\alpha > s_p, X_\beta > s_p)$$
$$\leq \frac{p(p-1)}{2} \cdot (2p) \cdot P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p, \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k13} > s_p\right).$$

Note that $P(\max_{\alpha \in I} X_\alpha \leq s_p) = E P_2(\max_{\alpha \in I} X_\alpha \leq s_p)$. From (3.52),

$$\left| P\left(\max_{\alpha \in I} X_\alpha \leq s_p\right) - E e^{-\lambda_{p1}} \right| \leq E \left| P_2\left(\max_{\alpha \in I} X_\alpha \leq s_p\right) - e^{-\lambda_{p1}} \right|$$
$$\leq E w_1 + E w_2.$$

Now,

$$E e^{-\lambda_{p1}} = E \exp\left[-\frac{p(p-1)}{2} P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)\right];$$
$$E w_1 \leq p^3 \cdot E\left[P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)^2 \right];$$
$$E w_2 \leq p^3 \cdot P\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p, \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k13} > s_p\right).$$

The following three lemmas say that $E e^{-\lambda_{p1}} \to \exp(-\frac{1}{4\sqrt{2\pi}} e^{-x/2})$, $E w_1 \to 0$ and $E w_2 \to 0$. The proof is then complete. □

LEMMA 3.12. *Let the assumptions in Proposition 3.1 hold. Review that P_2 stands for the conditional probability given $\{\xi_k; 1 \leq k \leq n\}$. Then*

$$E \exp\left[-\frac{p(p-1)}{2} P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)\right] \to \exp\left(-\frac{1}{4\sqrt{2\pi}} e^{-x/2}\right)$$

as $n \to \infty$ for all $x \in \mathbb{R}$.

LEMMA 3.13. *Let the assumptions in Proposition 3.1 hold. Review that P_2 stands for the conditional probability given $\{\xi_k; 1 \le k \le n\}$. Then*

$$E\left[P_2\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right)^2\right] = o\left(\frac{1}{p^3}\right)$$

as $n \to \infty$.

LEMMA 3.14. *Let the assumptions in Proposition 3.1 hold. Then*

$$P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p, \frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k13} > s_p\right) = o\left(\frac{1}{p^3}\right)$$

as $n \to \infty$.

Now we start to prove the three results one by one.

PROOF OF LEMMA 3.12. Write

$$\sum_{k=1}^n \eta_{k12} = \sum_{k=1}^n [a_n\xi_{k1}\xi_{k2} + b_n\xi_k(\xi_{k1}+\xi_{k2})]. \tag{3.53}$$

Given $\{\xi_k; 1 \le k \le n\}$, it is the sum of independent random variables with mean $E_2[a_n\xi_{k1}\xi_{k2} + b_n\xi_k(\xi_{k1} + \xi_{k2})] = 0$ and variance $\text{Var}_2[a_n\xi_{k1}\xi_{k2} + b_n\xi_k(\xi_{k1} + \xi_{k2})]^2 = a_n^2 + 2b_n^2\xi_k^2$. Thus,

$$\text{Var}_2\left(\sum_{k=1}^n \eta_{k12}\right) = na_n^2 + 2b_n^2\sum_{k=1}^n \xi_k^2. \tag{3.54}$$

Define

$$F_n = \left\{\max_{1\le k\le n}|\xi_k| \le \sqrt{n} \text{ and } \frac{6}{7} \le \frac{1}{n}\sum_{k=1}^n \xi_k^2 \le \frac{15}{14}\right\}.$$

Set $\tau = E(|\xi_1|^3) + 1$. For $v > 0$, set

$$G_n(v) = \left\{\frac{1}{n}\sum_{k=1}^n (1 + |\xi_k|^3)e^{v\xi_k^2(\log p)/n} \le 2\tau\right\}.$$

The parameter v will be chosen later. By the fact $P(|N(0,1)| \ge x) \le \frac{2}{\sqrt{2\pi}x}e^{-x^2/2}$ for all $x > 0$, the large deviations for i.i.d. random variables and Lemma 3.7, we have

$$P\big((F_n \cap G_n(v))^c\big) \le nP\big(|\xi_1| \ge \sqrt{n}\big) + P\left(\frac{1}{n}\sum_{k=1}^n \xi_k^2 \in \left[\frac{6}{7}, \frac{15}{14}\right]^c\right)$$
$$+ P\left(\frac{1}{n}\sum_{k=1}^n (1+|\xi_k|^3)e^{v\xi_k^2(\log p)/n} > 2\tau\right)$$
$$\le 3\exp\left(-\frac{1}{4}n^{1/2}(\log n)^{-2}\right), \tag{3.55}$$

as $n \ge n_v$, where $n_v \ge 1$ is a constant depending on v. Define $\sigma_{n0}^2 = a_n^2 + 2b_n^2(\frac{1}{n}\sum_{k=1}^n \xi_k^2)$. Then, on F_n,

$$\frac{1}{2} = \frac{1}{2}(a_n^2 + 2b_n^2) \le \sigma_{n0}^2 \le a_n^2 + \frac{15}{14}(2b_n^2) \le \frac{8}{7}, \tag{3.56}$$

where the last inequality follows from the identity $a_n^2 + 2b_n^2 = 1$.

Next, we will use Lemma 3.5 to get a precise estimate on $P_2(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p)$. To do so, Lemma 3.6 will be applied to control γ defined in Lemma 3.5.

Reviewing (3.53), we take $a = a_n/\sqrt{n}\sigma_{n0}$, $b = b_n\xi_k/\sqrt{n}\sigma_{n0}$. Set $\eta_k = a\xi_{k1}\xi_{k2} + b(\xi_{k1} + \xi_{k2})$. Then it follows from (3.54) that

$$E_2\eta_k = 0 \quad \text{and} \quad \sum_{k=1}^n \mathrm{Var}_2(\eta_k) = 1 \tag{3.57}$$

for each k. Furthermore, by (3.56) we have

$$|a| \le \frac{2}{\sqrt{n}} \quad \text{and} \quad |b| \le \frac{2|\xi_k|}{\sqrt{n}} \le 2 \tag{3.58}$$

on F_n. Then, on F_n, use the Hölder inequality, the facts that $2|\xi_{11}\xi_{12}| \le \xi_{11}^2 + \xi_{12}^2$ and $\xi_{11} + \xi_{12} \sim \sqrt{2}N(0,1)$, and independence to see

$$Ee^{h|\eta_k|} \le E\exp\left(\frac{2h}{\sqrt{n}}|\xi_{11}\xi_{12}| + 2h|\xi_{11} + \xi_{12}|\right)$$
$$\le \left[E\exp\left(\frac{2h}{\sqrt{n}}(\xi_{11}^2 + \xi_{12}^2)\right)\right]^{1/2} \cdot \big[E\exp\big(4\sqrt{2}hN(0,1)\big)\big]^{1/2}$$
$$= E\exp\left(\frac{2h}{\sqrt{n}}N(0,1)^2\right) \cdot e^{16h^2} < \infty, \tag{3.59}$$

for all h, k, n satisfying $0 < h \le h_n := \frac{1}{8}\sqrt{n}$ and $1 \le k \le n$. Now, on F_n, by Lemma 3.6 and (3.58) we have

$$E_2\big(|\eta_k|^3 e^{x|\eta_k|}\big) \le \frac{C}{n^{3/2}}(1+|\xi_k|^3)e^{4b^2x^2} \le \frac{C}{n^{3/2}}(1+|\xi_k|^3)e^{16x^2\xi_k^2/n} \tag{3.60}$$

for all $x \in (0, \frac{1}{12|a|})$. Observe that $(0, \frac{\sqrt{n}}{24}) \subset (0, \frac{1}{12|a|})$ on F_n by (3.58). Thus, (3.60) particularly holds for all $x \in (0, \sqrt{n}/24)$. Now take $x_0 = s_p/\sigma_{n0}$. Then

$$x_0 \le 2s_p < \sqrt{n}/24 \tag{3.61}$$

on F_n by the assumption $\log p = o(n^{1/3})$. We then have

$$\begin{aligned}\gamma &:= \sum_{k=1}^n E_2\big(|\eta_k|^3 e^{x_0|\eta_k|}\big)\\ &\le \frac{C}{n^{3/2}}\sum_{k=1}^n\big(1+|\xi_k|^3\big)e^{16x_0^2\xi_k^2/n}\\ &\le \frac{C}{n^{3/2}}\sum_{k=1}^n\big(1+|\xi_k|^3\big)e^{256\xi_k^2(\log p)/n}\end{aligned}$$

on F_n. Thus, $\gamma \le 2C\tau/\sqrt{n}$ on $F_n \cap G_n(256) := H_n$. The inequality in (3.55) implies

$$P\big(H_n^c\big) \le 3\exp\big(-n^{1/2}(\log n)^{-2}/4\big) \tag{3.62}$$

as n is sufficiently large. From (3.57), (3.59) and Lemma 3.5, we conclude

$$\begin{aligned}P_2\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right) &= P_2\left(\sum_{k=1}^n \eta_k > x_0\right)\\ &= \big[1-\Phi(x_0)\big]\cdot\big[1+O(1)\big(1+x_0^3\big)\gamma e^{4x_0^3\gamma}\big]\end{aligned} \tag{3.63}$$

on H_n since $x_0 < h_n = \sqrt{n}/8$ by (3.61). Finally, $x_0^3\gamma = O(s_p^3 n^{-1/2}) \to 0$ on H_n by the assumption $\log p = o(n^{1/3})$. Reviewing (3.56), we have $s_p/2 \le s_p/\sigma_{n0} \le 2s_p$ on H_n. Hence, from the formula $P(N(0,1) \ge x) = \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}(1+o(1))$ as $x \to \infty$ we obtain that, on H_n,

$$\begin{aligned}P_2\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right) &= \left[1-\Phi\left(\frac{s_p}{\sigma_{n0}}\right)\right]\cdot\left[1+O\left(\frac{\log^{3/2} p}{\sqrt{n}}\right)\right]\\ &= \frac{\sigma_{n0}}{\sqrt{2\pi}s_p}\cdot e^{-s_p^2/(2\sigma_{n0}^2)}\cdot\big(1+o(1)\big)\end{aligned} \tag{3.64}$$

as $n \to \infty$, where the last term "$o(1)$" does not depend on ξ_k's.

To prove the lemma, it is enough to show

$$\frac{p^2}{2}\cdot P_2\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right) \to \frac{1}{4\sqrt{2\pi}}e^{-x/2}$$

in probability as $n \to \infty$. Since $P(H_n) \to 1$, to complete the proof, it suffices to check

$$\frac{p^2}{2} \cdot P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right) \cdot I_{H_n} \to \frac{1}{4\sqrt{2\pi}} e^{-x/2}$$

in probability as $n \to \infty$. Now $\sigma_{n0} \to 1$ in probability as $n \to \infty$ and $s_p \sim 2\sqrt{\log p}$, comparing this with (3.64), it suffices to show

$$\frac{p^2}{4\sqrt{2\pi \log p}} \cdot e^{-s_p^2/(2\sigma_{n0}^2)} \cdot I_{H_n} \to \frac{1}{4\sqrt{2\pi}} e^{-x/2} \tag{3.65}$$

in probability. By the central limit theorem for i.i.d. random variables, $\sigma_{n0}^2 = 1 + O_p(1/\sqrt{n})$. Hence $\sigma_{n0}^{-2} = 1 + O_p(1/\sqrt{n})$. It follows that

$$\begin{aligned}\frac{s_p^2}{2\sigma_{n0}^2} &= \left(2\log p - \frac{1}{2}\log\log p + \frac{1}{2}x\right) \cdot \left[1 + O_p\left(\frac{1}{\sqrt{n}}\right)\right] \\ &= 2\log p - \frac{1}{2}\log\log p + \frac{1}{2}x + o_p(1)\end{aligned}$$

by the condition $\log p/n^{1/3} \to 0$. This implies (3.65). □

PROOF OF LEMMA 3.13. Review the proof of Lemma 3.12. Let H_n be defined as above (3.62). By (3.63), there exists a constant $n_1 \geq 1$ not depending on ξ_k's such that

$$P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right) \cdot I_{H_n} \leq 2\left[1 - \Phi\left(\frac{s_p}{\sigma_{n0}}\right)\right] \cdot I_{H_n}$$

as $n \geq n_1$ since $x_0 = s_p/\sigma_{n0}$. Recall the inequality $1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}$ for all $x > 0$. Then, from (3.56) we have

$$\left[P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)\right]^2 \cdot I_{H_n} \leq C \cdot \frac{\sigma_{n0}^2}{s_p^2} \cdot e^{-s_p^2/\sigma_{n0}^2} \cdot I_{H_n} \leq \frac{C}{\log p} \cdot e^{-7s_p^2/8}$$

as $p \geq n_2$, where n_2 is a constant not depending on ξ_k's. Therefore, combining this with (3.62), we see

$$\begin{aligned}E\left[P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)^2\right] &\leq E\left[P_2\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \eta_{k12} > s_p\right)^2 \cdot I_{H_n}\right] + P(H_n^c) \\ &\leq p^{-3.4} + 3\exp\left(-\frac{1}{4} n^{1/2} (\log n)^{-2}\right)\end{aligned}$$

as n is sufficiently large since $\log p = o(n^{1/3})$. This proves the lemma. □

PROOF OF LEMMA 3.14. Let P_3 and E_3 stand for the conditional probability and the conditional expectation given $\{\xi_k, \xi_{k1}; 1 \le k \le n\}$, respectively. By independence,

$$\begin{aligned} & P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p, \frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k13} > s_p\right) \\ & \quad = E\left[P_3\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right)^2\right]. \end{aligned} \tag{3.66}$$

Recall the notation in (3.18) and (3.19). Write $\sum_{k=1}^n \eta_{k12} = (b_n \sum_{k=1}^n \xi_k \xi_{k1}) + \sum_{k=1}^n (a_n \xi_{k1} + b_n \xi_k)\xi_{k2}$. Then, given $\{\xi_k, \xi_{k1}; 1 \le k \le n\}$, we have from independence that

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} \sim N(\mu_{0n}, \sigma_{0n}^2), \tag{3.67}$$

where

$$\mu_{0n} = \frac{b_n}{\sqrt{n}}\sum_{k=1}^n \xi_k \xi_{k1} \quad \text{and} \quad \sigma_{0n}^2 = \frac{1}{n}\sum_{k=1}^n (a_n \xi_{k1} + b_n \xi_k)^2.$$

Trivially, $b_n^2 = \frac{\rho_n}{1+\rho_n} \le \sup_{n\ge 1} \frac{\rho_n}{1+\rho_n} := \kappa^2 < \frac{1}{3}$ and $a_n^2 + b_n^2 = \frac{1}{1+\rho_n} \in (\frac{1}{2}, 1]$ for all $\rho_n \in [0, 1)$. Define

$$A = \{|\mu_{0n}| < \sqrt{3} s_p/2\} \quad \text{and} \quad B_\delta = \left\{1 - \delta < \frac{\sigma_{0n}^2}{a_n^2 + b_n^2} < 1 + \delta\right\}$$

for $\delta \in (0, 1)$. Observe $a_n \xi_{11} + b_n \xi_1 \stackrel{d}{=} \sqrt{a_n^2 + b_n^2} \cdot \xi_1$ since ξ_{11} and ξ_1 are i.i.d. $N(0, 1)$-distributed random variables. Thus, $\frac{\sigma_{0n}^2}{a_n^2+b_n^2} \stackrel{d}{=} n^{-1}\sum_{k=1}^n \xi_k^2$. Then, by the large deviations for the sum of i.i.d. random variables, we obtain

$$P(B_\delta^c) = P\left(\frac{1}{n}\sum_{k=1}^n \xi_k^2 \in [1 - \delta, 1 + \delta]^c\right) \le e^{-nC_\delta} \tag{3.68}$$

for all $\delta \in (0, 1)$ where $C_\delta > 0$ for each $\delta \in (0, 1)$. Similarly, $\{\xi_k \xi_{k1}; 1 \le k \le n\}$ are i.i.d. with mean zero and variance one. Notice $|\xi_1 \xi_{11}| \le \frac{1}{2}(|\xi_1|^2 + |\xi_{11}|^2)$. Therefore, $E \exp(\frac{1}{2}|\xi_1 \xi_{11}|) < \infty$. From Lemma 3.4 and the fact $s_p \sim 2\sqrt{\log p} = o(n^{1/6})$, we see that, for any $\varepsilon \in (0, 1)$,

$$\begin{aligned} P(A^c) &\le P\left(\frac{1}{\sqrt{n}}\left|\sum_{k=1}^n \xi_k \xi_{k1}\right| \ge \frac{\sqrt{3} s_p}{2\kappa}\right) = 2P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \xi_k \xi_{k1} \ge \frac{\sqrt{3} s_p}{2\kappa}\right) \\ &\le 2\exp\left(-\frac{1-\varepsilon}{2} \cdot \frac{3 s_p^2}{4\kappa^2}\right) \end{aligned}$$

as long as n is large enough. Since $\kappa^2 < 1/3$, we choose $\varepsilon = 1/2 - \kappa^2$. Then $\frac{1-\varepsilon}{2\kappa^2} = \frac{1}{2} + \frac{1}{4\kappa^2} > 1$. This implies that

$$P(A^c) \le 2\exp\left(-\frac{1-\varepsilon}{2\kappa^2}\cdot\frac{3s_p^2}{4}\right) = o\left(\frac{1}{p^3}\right)$$

as $n \to \infty$.

It is easy to see that $s_p - \mu_{0n} \to \infty$ on A. By the inequality $P(N(0,1) \ge y) \le \frac{1}{\sqrt{2\pi}y} e^{-y^2/2} \le \frac{1}{2} e^{-y^2/2}$ for all $y \ge 1$, we have from (3.67) that, on $A \cap B_\delta$,

$$P_3\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} \ge s_p\right) = P_3\big(N(\mu_{0n}, \sigma_{0n}^2) \ge s_p\big) = P_3\left(N(0,1) \ge \frac{s_p - \mu_{0n}}{\sigma_{0n}}\right)$$
$$\le \exp\left(-\frac{1}{2}\frac{(s_p-\mu_{0n})^2}{\sigma_{0n}^2}\right).$$

Note that $\sigma_{0n}^2 < (1+\delta)(a_n^2 + b_n^2) = \frac{1+\delta}{1+\rho_n}$ on B_δ. Therefore, on $A \cap B_\delta$,

$$P_3\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right) \le \exp\left(-\frac{1+\rho_n}{2(1+\delta)}\cdot (s_p - \mu_{0n})^2\right).$$

Review (3.66). We then have

$$\begin{aligned}
&P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p, \frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k13} > s_p\right)\\
&\quad\le E\left[P_3\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n \eta_{k12} > s_p\right)^2 I_{A^c\cup B_\delta^c}\right] + E\exp\left(-\frac{1+\rho_n}{1+\delta}\cdot (s_p-\mu_{0n})^2\right)\\
&\quad\le P(A^c) + P(B_\delta^c) + E\exp\left(-\frac{1+\rho_n}{1+\delta}\cdot (s_p-\mu_{0n})^2\right)\\
&\quad\le o\left(\frac{1}{p^3}\right) + e^{-nC_\delta} + E\exp\left(-\frac{1+\rho_n}{1+\delta}\cdot(\mu_{0n}-s_p)^2\right).
\end{aligned}$$

By Lemma 3.11, choosing $\delta > 0$ small enough, we know the last expectation is identical to $o(1/p^3)$. The desired conclusion follows from the assumption $\log p = o(n^{1/3})$. $\square$

3.4. *The proof of Proposition* 3.2. The major building block of L_n in Theorem 2.2 is M_n', as seen in the statement of Proposition 3.2. Now we prove this proposition. First, we need an analogue of Lemma 3.11. Review the notation from (3.14), (3.15) and (3.30). Set

$$\gamma_k = -\frac{1}{2}\rho_n a_n'(\xi_{k1}^2 - 1) + b_n'\xi_k\xi_{k1} \tag{3.69}$$

for $1 \le k \le m := n-1$ and $V_n = (\gamma_1 + \cdots + \gamma_m)/\sqrt{m}$.

LEMMA 3.15. *Let $\rho_n \in [0, 1)$ be constants. Suppose $p = p_n \to \infty$ and $\log p = o(n^{1/3})$ as $n \to \infty$. Let s_p be as in (3.16). Then there exists $\delta \in (0, 1)$ such that*

$$E\left\{I_{K_n'} \cdot \exp\left[-\frac{1-\delta}{1-\omega_n^2}(V_n - s_p)^2\right]\right\} = o\left(\frac{1}{p^3}\right) \tag{3.70}$$

as $n \to \infty$, where $K_n' := \{0 < V_n < \sqrt{7}\omega_n s_p/2\}$ and $\omega_n := \sqrt{\mathrm{Var}(\gamma_1)}$.

PROOF. First, if $\rho_n = 0$ for some $n \geq 1$, then $\gamma_k = 0$ for all $1 \leq k \leq m$. Hence $V_n = 0$ and the expectation in (3.70) is zero by the definition of K_n'. So it is enough to prove the conclusion by assuming $\rho_n > 0$ for all $n \geq 1$. The proof is divided into a few of steps.

Step 1. *Reduction of K_n' to a smaller set*. From the definitions of a_n' and b_n' in (3.30), it is easy to check that

$$(1 + \rho_n^2)a_n'^2 + 2b_n'^2 = 1. \tag{3.71}$$

Trivially, we have $\omega_n^2 = \rho_n^2 a_n'^2/2 + b_n'^2$. Therefore,

$$\begin{aligned}\omega_n^2 &= \frac{1}{2} - \frac{a_n'^2}{2}\\ &= \frac{1}{2} - \frac{1}{2} \cdot \frac{\frac{1-\rho_n}{1+\rho_n}}{(1-\rho_n)^2 + (2\rho_n)\frac{1-\rho_n}{1+\rho_n}}\\ &= \frac{1}{2}\left(1 - \frac{1}{1 + 2\rho_n - \rho_n^2}\right) < \frac{1}{4}\end{aligned} \tag{3.72}$$

because $1 + 2x - x^2 < 2$ for all $x \in [0, 1)$. In particular,

$$\frac{1}{1-\omega_n^2}\left(1 - \frac{1}{5}\omega_n\right)^2 > \left(1 - \frac{1}{5}\omega_n\right)^2 \geq \left(\frac{19}{20}\right)^2 > 0.8.$$

This implies that

$$\begin{aligned}&E\left\{I\left(0 < V_n \leq \frac{1}{5}\omega_n s_p\right) \cdot \exp\left[-\frac{1-\delta}{1-\omega_n^2}(V_n - s_p)^2\right]\right\}\\ &\quad \leq \exp\left[-\frac{1-\delta}{1-\omega_n^2}\left(1 - \frac{1}{5}\omega_n\right)^2 s_p^2\right]\\ &\quad \leq \exp[-3.2(1-2\delta)\log p] = o\left(\frac{1}{p^3}\right)\end{aligned}$$

as $n \to \infty$ if $\delta > 0$ is small enough. Therefore, it is enough to prove (3.70) with K_n' being replaced by $K_n'' = \{\omega_n s_p/5 < V_n < \sqrt{7}\omega_n s_p/2\}$.

Step 2. *The tail probability of* V_n. By the formula $\omega_n^2 = \frac{1}{2}\rho_n^2 a_n'^2 + b_n'^2$ again,

$$\left(\frac{-\frac{1}{2}\rho_n a_n'}{\omega_n}\right)^2 = \frac{1}{2}\cdot\left(1+\frac{2b_n'^2}{\rho_n^2 a_n'^2}\right)^{-1} \le \frac{1}{2}$$

and

$$\left(\frac{b_n'}{\omega_n}\right)^2 = \left(1+\frac{1}{2}\rho_n^2\frac{a_n'^2}{b_n'^2}\right)^{-1} \le 1.$$

Recall γ_k in (3.69). Set $\gamma_k' = \gamma_k/(\sqrt{m}\omega_n)$ for $1 \le k \le m$. The above implies that

$$\sqrt{m}\cdot|\gamma_k'| \le \frac{1}{2}(\xi_{k1}^2+1) + |\xi_k\xi_{k1}| \le \xi_{k1}^2 + \xi_k^2 + 1.$$

In addition, γ_k''s are i.i.d. with mean zero and satisfy $\sum_{k=1}^m \mathrm{Var}(\gamma_k') = 1$. Also, $Ee^{t|\gamma_k'|} < \infty$ if $0 < t < \sqrt{m}/4$. Observe

$$\gamma := \sum_{i=1}^m E\big(|\gamma_k'|^3 e^{x|\gamma_k'|}\big) \le \frac{1}{\sqrt{m}} E\big[(\xi_{11}^2+\xi_1^2+1)^3 e^{\frac{x}{\sqrt{m}}(\xi_{11}^2+\xi_1^2+1)}\big] \le \frac{C}{\sqrt{m}}$$

for all $0 \le x \le \sqrt{m}/4$, where $C = E[(\xi_{11}^2+\xi_1^2+1)^3 e^{\frac{1}{4}(\xi_{11}^2+\xi_1^2+1)}] < \infty$. By Lemma 3.5,

$$P(V_n \ge x) = P\left(\sum_{i=1}^m \gamma_k' \ge \frac{x}{\omega_n}\right) \le 2\left[1-\Phi\left(\frac{x}{\omega_n}\right)\right] \tag{3.73}$$

provided $(x/\omega_n)^3/\sqrt{m} \to 0$ and $0 \le x/\omega_n \le \sqrt{m}/4$. In particular, by the assumption $\log p = o(n^{1/3})$ and the fact $P(N(0,1) \ge t) \le \frac{1}{\sqrt{2\pi}t}e^{-t^2/2}$ for all $t > 0$ again, we have

$$P(V_n \ge x) \le e^{-x^2/(2\omega_n^2)} \tag{3.74}$$

for all $\frac{1}{5}\omega_n s_p < x < \frac{\sqrt{7}}{2}\omega_n s_p$.

Step 3. *The estimate of the expectation from* (3.70). Let A_1, B_1 and $\alpha_2 > 0$ be constants. Assume $[A_1, B_1] \subset [0, s_p]$. Notice $\frac{de^{-\alpha_2(x-s_p)^2}}{dx} = 2\alpha_2(s_p - x)\cdot e^{-\alpha_2(x-s_p)^2}$, we have

$$\begin{aligned} e^{-\alpha_2(v-s_p)^2} &= e^{-\alpha_2(A_1-s_p)^2} + 2\alpha_2\int_{A_1}^{v}(s_p - x)\cdot e^{-\alpha_2(x-s_p)^2}\,dx \\ &\le e^{-\alpha_2(A_1-s_p)^2} + 2\alpha_2 s_p\int_0^{\infty} e^{-\alpha_2(x-s_p)^2} I(A_1 \le x \le v)\,dx \end{aligned}$$

for any $s_p > v > A_1$. Replacing v with V_n, then multiplying both sides of the above by $I(A_1 < V_n < B_1)$, we get

$$
\begin{aligned}
&e^{-\alpha_2(V_n-s_p)^2} I(A < V_n < B) \\
&\quad \le e^{-\alpha_2(A_1-s_p)^2} + 2\alpha_2 s_p \int_0^\infty e^{-\alpha_2(x-s_p)^2} I(A_1 \le x \le V_n < B_1)\, dx \\
&\quad \le e^{-\alpha_2(A_1-s_p)^2} + 2\alpha_2 s_p \int_{A_1}^{B_1} e^{-\alpha_2(x-s_p)^2} I(V_n \ge x)\, dx.
\end{aligned}
$$

Set $A_1 = \omega_n s_p/5$ and $B_1 = \sqrt{7}\omega_n s_p/2$. By taking expectations on both sides of the above, we obtain from (3.74) that

$$
\begin{aligned}
&E\big[e^{-\alpha_2(V_n-s_p)^2} I(A_1 < V_n < B_1)\big] \\
&\quad \le e^{-\alpha_2(A_1-s_p)^2} + 2\alpha_2 s_p \int_{A_1}^{B_1} e^{-\alpha_2(x-s_p)^2} P(V_n \ge x)\, dx \qquad (3.75) \\
&\quad \le e^{-\alpha_2(A_1-s_p)^2} + 2\alpha_2 s_p \int_{A_1}^{B_1} \exp\biggl(-\alpha_2(x-s_p)^2 - \frac{x^2}{2\omega_n^2}\biggr)\, dx.
\end{aligned}
$$

Now we evaluate the integral. Write

$$
-\alpha_2(x-s_p)^2 - \frac{x^2}{2\omega_n^2} = -\biggl(\sqrt{\alpha_2 + \frac{1}{2\omega_n^2}}\,x - \frac{\alpha_2 s_p}{\sqrt{\alpha_2 + \frac{1}{2\omega_n^2}}}\biggr)^2 - \frac{\alpha_2 s_p^2}{2\alpha_2\omega_n^2 + 1}.
$$

Now, define y such that

$$
\frac{y}{\sqrt{2}} = \sqrt{\alpha_2 + \frac{1}{2\omega_n^2}}\,x - \frac{\alpha_2 s_p}{\sqrt{\alpha_2 + \frac{1}{2\omega_n^2}}}. \qquad (3.76)
$$

It follows that

$$
\begin{aligned}
&2\alpha_2 s_p \int_{A_1}^{B_1} \exp\biggl(-\alpha_2(x-s_p)^2 - \frac{x^2}{2\omega_n^2}\biggr)\, dx \\
&\quad = (2\alpha_2 s_p) \cdot \exp\biggl(-\frac{\alpha_2 s_p^2}{2\alpha_2\omega_n^2 + 1}\biggr) \cdot \int_{A'}^{B'} e^{-y^2/2}\, dy \cdot \frac{1}{\sqrt{2\alpha_2 + \frac{1}{\omega_n^2}}} \\
&\quad \le \frac{\sqrt{8\pi}\alpha_2}{\sqrt{2\alpha_2 + \frac{1}{\omega_n^2}}} s_p \cdot \exp\biggl(-\frac{\alpha_2 s_p^2}{2\alpha_2\omega_n^2 + 1}\biggr) \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2}\, dy,
\end{aligned}
$$

where A' and B' are the corresponding values of y in (3.76) as $x = A_1$ and B_1, respectively. This combining with (3.75) implies

$$E\big[e^{-\alpha_2(V_n-s_p)^2}I(A_1 < V_n < B_1)\big] \le e^{-\alpha_2(A_1-s_p)^2} + \sqrt{4\pi\alpha_2 s_p^2}\cdot\exp\biggl(-\frac{s_p^2}{2\omega_n^2+\alpha_2^{-1}}\biggr),$$

where the inequality $\alpha_2(2\alpha_2+\omega_n^{-2})^{-1/2} \le \sqrt{\alpha_2/2}$ is used. Take $\alpha_2 = \frac{1-\delta}{1-\omega_n^2}$. Then $1-\delta < \alpha_2 < 4/3$ by (3.72). Note

$$2\omega_n^2+\alpha_2^{-1} = 2\omega_n^2 + \frac{1-\omega_n^2}{1-\delta} \le \frac{1+\omega_n^2}{1-\delta} \le \frac{1+\omega^2}{1-\delta},$$

where $\omega^2 := \sup_{n\ge 1}\omega_n^2 \le 1/4$. From (3.72), we know $A_1 \le s_p/10$. This concludes

$$E\biggl\{I_{K_n''}\cdot\exp\biggl[-\frac{1-\delta}{1-\omega_n^2}(V_n-s_p)^2\biggr]\biggr\} \le \exp\biggl[-(1-\delta)\biggl(\frac{1}{10}s_p - s_p\biggr)^2\biggr] + 2\sqrt{6\pi(\log p)}\cdot\exp\biggl(-\frac{1-\delta}{1+\omega^2}s_p^2\biggr).$$

The first term on the right-hand side is $o(p^{-3})$ if $(1-\delta)(\frac{9}{10})^2\cdot 4 > 3$, which is true if $0<\delta<2/27$; the second term is $o(p^{-3})$ as long as $\frac{1-\delta}{1+\omega^2} > 3/4$, which is equivalent to that $0 < \delta < 1-3(1+\omega^2)/4$. The desired conclusion then follows from the fact $\omega^2 \le 1/4$. □

PROOF OF PROPOSITION 3.2. The strategy of the proof is similar to that of Proposition 3.1. However, the technical details are more involved. Let I, s_p and B_α be as in the proof of Proposition 3.1. For $\alpha = (i,j)\in I$, define $X_\alpha = M'_{nij}$. Let P_2 and E_2 stand for the conditional probability and the conditional expectation given $\{\xi_k; 1\le k\le m\}$, respectively. Again, the key observation is that, given $\{\xi_k; 1\le k\le m\}$, random variable X_α is independent of $\{X_\beta; \beta\notin B_\alpha\}$. Since $\{X_\alpha, \alpha\in I\}$ are identically distributed under P_2, by Lemma 3.3, we have

$$\Bigl|P_2\Bigl(\max_{\alpha\in I}X_\alpha \le s_p\Bigr) - e^{-\lambda_{p2}}\Bigr| \le w_1' + w_2', \tag{3.77}$$

where

$$\lambda_{p2} = \frac{p(p-1)}{2}P_2\Biggl(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta'_{k12} > s_p\Biggr)$$

and

$$w_1' = \sum_{\alpha\in I}\sum_{\beta\in B_\alpha}P_2(X_\alpha > s_p)P_2(X_\beta > s_p) \le \frac{p(p-1)}{2}\cdot(2p)\cdot P_2\Biggl(\frac{1}{\sqrt{m}}\sum_{k=1}^m\eta'_{k12} > s_p\Biggr)^2$$

and

$$w_2' = \sum_{\alpha\in I}\sum_{\beta\in B_\alpha} P_2(X_\alpha > s_p, X_\beta > s_p)$$
$$\leq \frac{p(p-1)}{2}\cdot(2p)\cdot P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k12}' > s_p, \frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k13}' > s_p\right).$$

Note that $P(\max_{\alpha\in I} X_\alpha \leq s_p) = EP_2(\max_{\alpha\in I} X_\alpha \leq s_p)$. From (3.77),

$$\left|P\left(\max_{\alpha\in I} X_\alpha \leq s_p\right) - Ee^{-\lambda_{p2}}\right| \leq E\left|P_2\left(\max_{\alpha\in I} X_\alpha \leq s_p\right) - e^{-\lambda_{p2}}\right|$$
$$\leq Ew_1' + Ew_2'.$$

Obviously,

$$Ee^{-\lambda_{p2}} = E\exp\left[-\frac{p(p-1)}{2}P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k12}' > s_p\right)\right];$$
$$Ew_1' \leq p^3\cdot E\left[P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k12}' > s_p\right)^2\right];$$
$$Ew_2' \leq p^3\cdot P\left(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k12}' > s_p, \frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k13}' > s_p\right).$$

The following three lemmas say that $Ee^{-\lambda_{p2}} \to \exp(-\frac{1}{4\sqrt{2\pi}}e^{-x/2})$, $Ew_1' \to 0$ and $Ew_2' \to 0$. The proof is then complete. □

LEMMA 3.16. *Let the assumptions in Proposition 3.2 hold. Review $m = n-1$ and P_2 stands for the conditional probability given $\{\xi_k; 1 \leq k \leq m\}$. Then*

$$E\exp\left[-\frac{p(p-1)}{2}P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k12}' > s_p\right)\right] \to \exp\left(-\frac{1}{4\sqrt{2\pi}}e^{-x/2}\right) \tag{3.78}$$

as $n \to \infty$ for all $x \in \mathbb{R}$.

LEMMA 3.17. *Let the assumptions in Proposition 3.2 hold. Review $m = n-1$ and P_2 stands for the conditional probability given $\{\xi_k; 1 \leq k \leq m\}$. Then*

$$E\left[P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^m \eta_{k12}' > s_p\right)^2\right] = o\left(\frac{1}{p^3}\right) \tag{3.79}$$

as $n \to \infty$.

LEMMA 3.18. *Let the assumptions in Proposition 3.2 hold. Review $m = n-1$. Then*

$$P\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p, \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k13} > s_p\right) = o\left(\frac{1}{p^3}\right)$$

as $n \to \infty$.

Now we start to prove the three statements one by one.

PROOF OF LEMMA 3.16. We will get a sharp estimate on $P_2(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p)$ first by using Lemma 3.5. To carry on this, we have to check the required conditions.

Step 1: *the behaviors of η'_{k12}*. Write

$$\sum_{k=1}^{m}\eta'_{k12} = \sum_{k=1}^{m}\left\{a'_n\left[\xi_{k1}\xi_{k2} - \frac{\rho_n}{2}(\xi_{k1}^2 + \xi_{k2}^2 - 2)\right] + b'_n\xi_k(\xi_{k1} + \xi_{k2})\right\}. \tag{3.80}$$

Given $\{\xi_k; 1 \le k \le m\}$, it is the sum of independent random variables. It is easy to check that

$$E_2\eta'_{k12} = 0, \qquad \mathrm{Var}_2(\eta'_{kij}) = a_n'^2(1+\rho_n^2) + 2b_n'^2\xi_k^2. \tag{3.81}$$

So the conditional variance

$$\mathrm{Var}_2\left(\sum_{k=1}^{m}\eta'_{k12}\right) = m(1+\rho_n^2)a_n'^2 + 2b_n'^2\sum_{k=1}^{m}\xi_k^2. \tag{3.82}$$

Set

$$F_n = \left\{\max_{1\le k\le n}|\xi_k| \le \sqrt{m}\text{and}\frac{6}{7} \le \frac{1}{m}\sum_{k=1}^{m}\xi_k^2 \le \frac{15}{14}\right\}.$$

Recall the notation $\tau = E(|\xi_1|^3) + 1$ defined earlier. For $v > 0$, define

$$G_n(v) = \left\{\frac{1}{m}\sum_{k=1}^{m}(1+|\xi_k|^3)e^{v\xi_k^2(\log p)/m} \le 2\tau\right\}. \tag{3.83}$$

The parameter v will be chosen later. The inequality (3.55) says that

$$P\big((F_n \cap G_n(v))^c\big) \le 3\exp\left(-\frac{1}{4}n^{1/2}(\log n)^{-2}\right) \tag{3.84}$$

as $n \ge n_v$, where $n_v \ge 1$ is a constant depending on v only. Define

$$\sigma_{n2}^2 = (1+\rho_n^2)a_n'^2 + 2b_n'^2\frac{1}{m}\sum_{k=1}^{m}\xi_k^2. \tag{3.85}$$

Note (3.71). Then, on F_n,

$$\frac{1}{2}=\frac{1}{2}\big[1+\rho_n^2)a_n'^2+2b_n'^2\big]\le\sigma_{n2}^2\le\big(1+\rho_n^2)a_n'^2+2b_n'^2\cdot\frac{15}{14}\le\frac{15}{14}. \tag{3.86}$$

Next, we will use Lemma 3.5 to get a precise estimate on $P_2(\frac{1}{\sqrt{m}}\sum_{k=1}^m\eta_{k12}'>s_p)$. To do so, set

$$a'=c'=-\frac{a_n'\rho_n}{2\sqrt{m}\sigma_{n2}};\qquad b'=\frac{a_n'}{\sqrt{m}\sigma_{n2}};$$
$$d'=e'=\frac{b_n'\xi_k}{\sqrt{m}\sigma_{n2}};\qquad f'=\frac{a_n'\rho_n}{\sqrt{m}\sigma_{n2}}$$

and $\eta_k'=a'\xi_{k1}^2+b'\xi_{k1}\xi_{k2}+c'\xi_{k2}^2+d'\xi_{k1}+e'\xi_{k2}+f'$. Then it follows from (3.82) that

$$E_2\eta_k'=0\quad\text{for each }1\le k\le n\quad\text{and}\quad\sum_{k=1}^n\mathrm{Var}_2(\eta_k')=1. \tag{3.87}$$

Furthermore, from (3.85), $\sigma_{n2}\ge\max\{a_n',b_n'\}$ on F_n. Then

$$\max\{|a'|,|b'|,|c'|,|f'|\}\le\frac{1}{\sqrt{m}}\quad\text{and}\quad|d'|=|e'|\le\frac{|\xi_k|}{\sqrt{m}}\le 1 \tag{3.88}$$

on F_n. Hence, on F_n, $|\eta_k'|\le\frac{2}{\sqrt{m}}(\xi_{k1}^2+\xi_{k2}^2)+|\xi_{k1}+\xi_{k2}|+\frac{1}{\sqrt{m}}$. By the fact $\xi_{k1}+\xi_{k2}\sim\sqrt{2}N(0,1)$ and independence,

$$\begin{aligned}Ee^{h|\eta_k'|}&\le e^{h/\sqrt{m}}\cdot E\exp\biggl(\frac{2h}{\sqrt{m}}(\xi_{k1}^2+\xi_{k2}^2)+h|\xi_{11}+\xi_{12}|\biggr)\\&\le e^{1/16}\cdot\biggl[E\exp\biggl(\frac{4h}{\sqrt{m}}(\xi_{k1}^2+\xi_{k2}^2)\biggr)\biggr]^{1/2}\\&\quad\cdot\bigl[E\exp(2\sqrt{2}hN(0,1))\bigr]^{1/2}\\&=2\cdot E\exp\biggl(\frac{4h}{\sqrt{m}}N(0,1)^2\biggr)\cdot e^{2h^2}<\infty\end{aligned} \tag{3.89}$$

for all h, k, n satisfying $0<h\le h_n:=\sqrt{m}/16$ and $1\le k\le m$. Now, on F_n, by Lemma 3.6, (3.86) and (3.88) we have

$$\begin{aligned}E_2\bigl(|\eta_k'|^3e^{x|\eta_k'|}\bigr)&\le\frac{C}{m^{3/2}}(1+|\xi_k|^3)e^{4x^2\xi_k^2/m}\cdot e^{x/\sqrt{m}}\\&\le\frac{C}{m^{3/2}}(1+|\xi_k|^3)e^{4x^2\xi_k^2/m}\end{aligned} \tag{3.90}$$

if $0<x\le\frac{1}{12(|a'|+|b'|+|c'|)}\wedge\sqrt{m}$, where C here and later in the proof is a constant not depending on ξ_k's and may be different from line to line. Observe that

$(0, \frac{\sqrt{m}}{36}) \subset (0, \frac{1}{12(|a'|+|b'|+|c'|)})$ on F_n by (3.88). Thus, (3.90) particularly holds for all $x \in (0, \sqrt{m}/36)$. Now we take $x_1 = s_p/\sigma_{n2}$. Then, by (3.86),

$$x_1 \le 2s_p < \sqrt{m}/36 \tag{3.91}$$

on F_n by the assumption $\log p = o(n^{1/3})$. We then have

$$\begin{aligned}\gamma &:= \sum_{k=1}^{m} E_2\big(|\eta_k'|^3 e^{x_1|\eta_k'|}\big) \le \frac{C}{m^{3/2}} \sum_{k=1}^{m} (1+|\xi_k|^3) e^{4x_1^2\xi_k^2/m} \\ &\le \frac{C}{m^{3/2}} \sum_{k=1}^{m} (1+|\xi_k|^3) e^{64\xi_k^2(\log p)/m}\end{aligned}$$

on F_n as n is sufficiently large. Thus, $\gamma \le 2C\tau/\sqrt{m}$ on $F_n \cap G_n(64) := H_n$ by (3.83). The inequality in (3.84)) implies

$$P(H_n^c) = o\bigg(\frac{1}{p^3}\bigg) \tag{3.92}$$

as n is sufficiently large since $\log p = o(n^{1/3})$ by assumption.

Step 2: *a sharp estimate on* $P_2(m^{-1/2}\sum_{k=1}^m \eta_{k12}' > s_p)$ *by Lemma* 3.5. By (3.91), we see that $x_1 < \sqrt{m}/36 < \sqrt{m}/16 = h_n$. From (3.87), (3.89) and Lemma 3.5, we conclude

$$\begin{aligned}P_2\bigg(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta_{k12}' > s_p\bigg) &= P_2\bigg(\sum_{k=1}^{m}\eta_k' > \frac{s_p}{\sigma_{n2}}\bigg) \\ &= \big[1-\Phi(x_1)\big]\cdot\big[1+O(1)\big(1+x_1^3\big)\gamma e^{4x_1^3\gamma}\big]\end{aligned}$$

on H_n. Just notice $|O(1)|$ is bounded by an absolute constant. Finally, by (3.91), $x_1^3\gamma = O(s_p^3 m^{-1/2}) \to 0$ on H_n. Reviewing (3.86), we have $s_p/2 \le x_1 \le 2s_p$ on H_n. Hence, from the formula $P(N(0,1) \ge x) = \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}(1+o(1))$ as $x \to \infty$ we obtain that, on H_n,

$$\begin{aligned}P_2\bigg(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta_{k12}' > s_p\bigg) &= \bigg[1-\Phi\bigg(\frac{s_p}{\sigma_{n2}}\bigg)\bigg]\cdot\big(1+o(1)\big) \\ &= \frac{\sigma_{n2}}{\sqrt{2\pi}s_p}\cdot e^{-s_p^2/(2\sigma_{n2}^2)}\cdot\big(1+o(1)\big)\end{aligned} \tag{3.93}$$

as $n \to \infty$, where $o(1)$ does not depend on ξ_k's.

Step 3: *proof of* (3.78) *by* (3.93). By the bounded convergence theorem, to prove the lemma, it is enough to show that

$$\frac{p^2}{2}\cdot P_2\bigg(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta_{k12}' > s_p\bigg) \to \frac{1}{4\sqrt{2\pi}}e^{-x/2}$$

in probability as $n \to \infty$. Since $P(H_n) \to 1$, to complete the proof, it is enough to prove

$$\frac{p^2}{2} \cdot P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right) \cdot I_{H_n} \to \frac{1}{4\sqrt{2\pi}}e^{-x/2}$$

in probability as $n \to \infty$. Recalling (3.71) and (3.85), it is trivial to see $\sigma_{n2} \to 1$ in probability as $n \to \infty$. Also, $s_p \sim 2\sqrt{\log p}$, comparing this with (3.93), it is enough to prove

$$\frac{p^2}{4\sqrt{2\pi \log p}} \cdot e^{-s_p^2/(2\sigma_{n2}^2)} \cdot I_{H_n} \to \frac{1}{4\sqrt{2\pi}}e^{-x/2} \tag{3.94}$$

in probability. By the central limit theorem for i.i.d. random variables, we know $\sigma_{n2}^2 = 1 + O_p(\frac{1}{\sqrt{n}})$ from (3.71) and (3.85). Hence $\sigma_{n2}^{-2} = 1 + O_p(\frac{1}{\sqrt{n}})$. This leads to that

$$\begin{aligned}\frac{s_p^2}{2\sigma_{n2}^2} &= \left(2\log p - \frac{1}{2}\log\log p + \frac{1}{2}x\right) \cdot \left[1 + O_p\left(\frac{1}{\sqrt{n}}\right)\right]\\ &= 2\log p - \frac{1}{2}\log\log p + \frac{1}{2}x + o_p(1)\end{aligned}$$

by the condition $(\log p)/n^{1/3} \to 0$. We then get (3.94). □

PROOF OF LEMMA 3.17. Review the proof of Lemma 3.16. Let H_n be defined as above (3.92). By (3.93), there exists a constant $n_1 \geq 1$ not depending on ξ_k's such that

$$P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right) \cdot I_{H_n} \leq \frac{\sigma_{n2}}{\sqrt{2\pi}s_p} \cdot e^{-s_p^2/(2\sigma_{n2}^2)} \cdot I_{H_n}$$

as $n \geq n_1$. Then

$$\left[P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)\right]^2 \cdot I_{H_n} \leq C \cdot \frac{\sigma_{n2}^2}{s_p^2} \cdot e^{-s_p^2/\sigma_{n2}^2} \cdot I_{H_n} \leq \frac{C}{\log p} \cdot e^{-7s_p^2/8}$$

on H_n as $n \geq n_2$ by (3.86), where n_2 is a constant not depending on ξ_k's. Therefore, combining this with (3.92), we see

$$\begin{aligned}E\left[P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)^2\right] &\leq E\left[P_2\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)^2 \cdot I_{H_n}\right] + P(H_n^c)\\ &\leq p^{-3.4} + o\left(\frac{1}{p^3}\right)\end{aligned}$$

as n is sufficiently large. This proves the lemma. □

PROOF OF LEMMA 3.18. Let P_3 and E_3 stand for the conditional probability and the conditional expectation given $\{\xi_k, \xi_{k1}; 1 \le k \le n\}$, respectively. By independence,

$$\begin{aligned} &P\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p, \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k13} > s_p\right) \\ &\quad = E\left[P_3\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)^2\right]. \end{aligned} \tag{3.95}$$

Write

$$\eta'_{k12} = \alpha_n(U_k^2 - 1) + \beta_k U_k + \gamma_k,$$

where $U_k = \xi_{k2}$,

$$\begin{aligned} \alpha_n &= -\frac{1}{2}\rho_n a'_n, \\ \beta_k &= (a'_n\xi_{k1} + b'_n\xi_k), \end{aligned}$$

and γ_k is defined in (3.69). Now,

$$\begin{aligned} &P_3\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right) \\ &\quad = P_3\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}[\alpha_n(U_k^2 - 1) + \beta_k U_k] > s_p - \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\gamma_k\right). \end{aligned} \tag{3.96}$$

We will complete the proof with a couple of steps.

Step 1: *the size of* $\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\gamma_k$. Unconditionally, $\{\gamma_k; 1 \le k \le m\}$ are i.i.d. with mean zero and variance $\omega_n^2 = \frac{1}{2}\rho_n^2 a_n'^2 + b_n'^2$ mentioned below (3.71). By (3.72), $\omega_n^2 < 1/4$. From (3.73) and the fact $P(N(0,1) \ge t) \le \frac{1}{\sqrt{2\pi}t}e^{-t^2/2}$ for all $t > 0$,

$$P\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\gamma_k \ge \theta\omega_n s_p\right) \le 2e^{-(\theta s_p)^2/2} \le p^{-2\theta^2}(\log p)^{\theta^2}$$

as n is sufficiently large for all $\theta > 0$. Review the short argument as in getting (3.73), the above inequality also holds if "γ_k" is replaced by "$-\gamma_k$." It follows that

$$P\left(\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m}\gamma_k\right| \ge \theta\omega_n s_p\right) \le 2p^{-2\theta^2}(\log p)^{\theta^2}$$

as n is sufficiently large for all $\theta > 0$. Set

$$\tilde{K}_n = \left\{\frac{1}{\sqrt{m}}\left|\sum_{k=1}^{m}\gamma_k\right| < \frac{\sqrt{7}}{2}\omega_n s_p\right\}.$$

Then

$$P(\tilde{K}_n^c) = o\Big(\frac{1}{p^3}\Big) \tag{3.97}$$

as $n \to \infty$. Let

$$s'_p = s_p - \frac{1}{\sqrt{m}} \sum_{k=1}^m \gamma_k.$$

Set $W_k = \alpha_n(U_k^2 - 1) + \beta_k U_k$ for $1 \le k \le m$. Now we consider

$$P_3\Big(\frac{1}{\sqrt{m}} \sum_{k=1}^m [\alpha_n(U_k^2 - 1) + \beta_k U_k] > s'_p\Big) = P_3\Big(\frac{1}{\sqrt{m}} \sum_{k=1}^m W_k > s'_p\Big).$$

Step 2: *the behaviors of* W_k*'s on typical sets*. Observe

$$\begin{aligned} &E_3(W_k) = 0; \\ &\mathrm{Var}_3(W_k) = 2\alpha_n^2 + \beta_k^2 = \frac{1}{2}(\rho_n a'_n)^2 + (a'_n \xi_{k1} + b'_n \xi_k)^2. \end{aligned} \tag{3.98}$$

It follows that

$$\begin{aligned} \sigma_{n3}^2 &:= \mathrm{Var}_3\Big(\frac{1}{\sqrt{m}} \sum_{k=1}^m W_k\Big) = \frac{1}{2}(\rho_n a'_n)^2 + \frac{1}{m} \sum_{k=1}^m (a'_n \xi_{k1} + b'_n \xi_k)^2 \\ &\stackrel{d}{=} \frac{1}{2}(\rho_n a'_n)^2 + (a_n'^2 + b_n'^2) \frac{1}{m} \sum_{k=1}^m \xi_k^2. \end{aligned} \tag{3.99}$$

Set

$$F_n(\delta) = \Big\{ \max_{1 \le k \le n} |\xi_k| \le \sqrt{n} \text{ and } 1 - \delta \le \frac{1}{n} \sum_{k=1}^n \xi_k^2 \le 1 + \delta \Big\}$$

for $n \ge 1$ and $\delta \in (0, 1)$. By the fact $P(N(0,1) \ge t) \le \frac{1}{\sqrt{2\pi} t} e^{-t^2/2}$ for all $t > 0$ again and (3.68), for any $\delta > 0$, there is a constant $C_\delta > 0$ such that

$$P(F_n(\delta)^c) \le e^{-nC_\delta} \tag{3.100}$$

as n is sufficiently large. Review $m = n - 1$. Under $F_m(\delta)$, it is easy to see from (3.71) that

$$|\sigma_{n3}^2 - \sigma_{03}^2| \le \delta, \tag{3.101}$$

where $\sigma_{03}^2 := (\frac{1}{2}\rho_n^2 + 1) a_n'^2 + b_n'^2$. Evidently,

$$1/2 \le \sigma_{03}^2 \le 1 \tag{3.102}$$

by (3.71). Now, review the notation $\tau = E(|\xi_1|^3) + 1$ defined earlier. For $v > 0$, define

$$G_m(v) = \left\{ \frac{1}{m} \sum_{k=1}^{m} (1 + |\beta_k|^3) e^{v\beta_k^2 (\log p)/m} \le 2\tau \right\}.$$

The parameter v will be chosen later. Now $(\beta_1, \ldots, \beta_m)' \stackrel{d}{=} \sqrt{a_n'^2 + b_n'^2}(\xi_1, \ldots, \xi_m)$. From (3.71) we know $a_n'^2 + b_n'^2 \le 1$. Then, by Lemma 3.7, for all $v > 0$, there exists $n_v > 0$ such that

$$\begin{aligned} P\big(G_m(v)^c\big) &\le P\left(\frac{1}{m} \sum_{k=1}^{m} (1 + |\xi_k|^3) e^{v\xi_k^2 (\log p)/m} > 2\tau \right) \\ &\le \exp\left(-\frac{1}{4} m^{1/2} (\log m)^{-2} \right) \end{aligned}$$

for all $n \ge n_v$. Define $H_n(\delta, v) := F_m(\delta) \cap G_m(v) \cap \tilde{K}_n$. Join the above with (3.97) and (3.100) to see

$$P\big(H_n(\delta, v)^c\big) = o\left(\frac{1}{p^3}\right) \tag{3.103}$$

as $n \to \infty$ for all $\delta \in (0, 1)$ and $v > 0$. By Hölder's inequality,

$$\begin{aligned} E_3 e^{h|W_k|/\sqrt{m}} &\le e^{h|\alpha_n|/\sqrt{m}} \cdot E_3 \exp\big(hm^{-1/2}|\alpha_n| U_k^2 + hm^{-1/2}|\beta_k||U_k|\big) \\ &\le e^{|\alpha_n|h} \cdot \big[E_3 \exp\big(2hm^{-1/2}|\alpha_n|U_1^2\big)\big]^{1/2} \\ &\quad \cdot \big[E_3 \exp\big(2hm^{-1/2}|\beta_k||U_1|\big)\big]^{1/2} \\ &< \infty \end{aligned} \tag{3.104}$$

as long as $0 < h \le \frac{\sqrt{m}}{8|\alpha_n|}$. From (3.71), we see $|\alpha_n| \le 1$. Therefore, (3.104) holds for all $0 < h \le h_n := \sqrt{m}/8$. Furthermore, by taking $a = \alpha_n/\sqrt{m}$, $d = \beta_k/\sqrt{m}$, $f = -\alpha_n/\sqrt{m}$ and $b = c = e = 0$, we have from Lemma 3.6 that

$$\begin{aligned} E_3\left(\frac{|W_k|^3}{\sqrt{m^3}} e^{x|W_k|/\sqrt{m}} \right) &\le \frac{C}{m^{3/2}} \big(|\alpha_n|^3 + |\beta_k|^3\big) \cdot \exp\left(\frac{2\beta_k^2}{m} x^2 \right) \cdot e^{|\alpha_n|x/\sqrt{m}} \\ &\le \frac{Ce}{m^{3/2}} \cdot \big(1 + |\beta_k|^3\big) \cdot \exp\left(\frac{2\beta_k^2}{m} x^2 \right) \end{aligned} \tag{3.105}$$

for all $0 < x \le \sqrt{m}/12$ since $\frac{1}{12|\alpha_n|}\sqrt{m} \ge \sqrt{m}/12$. Now take $x_3 = \frac{s_p'}{\sigma_{n3}}$. The assertions (3.101) and (3.102) imply that $1/4 \le \sigma_{n3}^2 \le 2$ on $F_m(\delta)$ for all $\delta \in (0, 1/4]$. Then $x_3 \le 2s_p'$ on $H_n(\delta, v)$ for all $\delta \in (0, 1/4]$ and all $v > 0$. Moreover, due to the fact $0 \le \omega_n < 1/2$ we see that

$$0 < s_p' \le s_p + \frac{1}{\sqrt{m}} \left| \sum_{k=1}^{m} \gamma_k \right| \le s_p + \frac{\sqrt{7}}{2} \omega_n s_p \le 2s_p \tag{3.106}$$

on $\tilde{K}_n$. This says that $0 < x_3 \le 2s'_p \le 4s_p \le \sqrt{m}/24$ as $n \ge n_{v,\delta} \ge n_v$ for all $\delta \in (0, 1/4]$ and all $v > 0$, where $n_{v,\delta} > 0$ is a constant depending on δ and v. This and (3.105) yield

$$\begin{aligned}\sum_{k=1}^{m} E_3\left(\frac{|W_k|^3}{\sqrt{m^3}} e^{2x_3|W_k|/\sqrt{m}}\right) &\le \frac{C}{m^{3/2}} \sum_{k=1}^{m} (1 + |\beta_k|^3) \cdot \exp\left(128 s_p^2 \cdot \frac{\beta_k^2}{m}\right) \\ &\le \frac{2\tau C}{\sqrt{m}}\end{aligned} \tag{3.107}$$

on $H_n(\delta, 128)$ as $n \ge n_\delta \ge n_{128,\delta}$ for all $\delta \in (0, 1/4]$, where n_δ depends on δ. The last step follows from the definition of $G_m(v)$ and the fact $s_p^2 \le 4 \log p$ as n is sufficiently large.

Step 3: *a bound on* $P_3(\frac{1}{\sqrt{m}} \sum_{k=1}^m \eta'_{k12} > s_p)$. Review (3.96) and the definition of W_k, we see

$$P_3\left(\frac{1}{\sqrt{m}} \sum_{k=1}^{m} \eta'_{k12} > s_p\right) = P_3\left(\frac{1}{\sqrt{m}} \sum_{k=1}^{m} \frac{W_k}{\sigma_{n3}} > x_3\right) \tag{3.108}$$

since $x_3 = s'_p/\sigma_{n3}$. Set $W'_k = W_k/(\sqrt{m}\sigma_{n3})$ for $1 \le k \le m$. Then (3.98) and (3.99) imply

$$EW'_k = 0 \quad \text{and} \quad \mathrm{Var}_3\left(\sum_{k=1}^{m} W'_k\right) = 1$$

for each $1 \le k \le m$. Since $\frac{1}{4} \le \sigma_{n3}^2 \le 2$ on $F_m(\delta)$ for all $\delta \in (0, \frac{1}{4}]$, we see from (3.104) that $E_3 e^{h|W'_k|} \le E_3 e^{2h|W_k|/\sqrt{m}} < \infty$ for all $0 < h \le h_n := \frac{\sqrt{m}}{16}$. Moreover, by (3.107),

$$\begin{aligned}\gamma := \sum_{k=1}^{m} E\left(|W'_k|^3 e^{x_3|W'_k|}\right) &= \sum_{k=1}^{m} E_3\left[\frac{|W_k|^3}{\sigma_{n3}^3\sqrt{m^3}} \exp\left(x_3 \frac{|W_k|}{\sqrt{m}\sigma_{n3}}\right)\right] \\ &\le 8 \sum_{k=1}^{m} E_3\left(\frac{|W_k|^3}{\sqrt{m^3}} e^{2x_3|W_k|/\sqrt{m}}\right) \le \frac{16\tau C}{\sqrt{m}}\end{aligned}$$

on $H_n(\delta, 128)$ for all $n \ge n_\delta$ and $\delta \in (0, 1/4]$. Trivially, $0 < x_3 \le \sqrt{m}/24 < h_n$. The inequality from (3.106) says that $x_3^3 \gamma = O(s_p^3/\sqrt{m}) \to 0$ on $H_n(\delta, 128)$ by the condition $\log p = o(n^{1/3})$. After verifying all conditions required in Lemma 3.5, we conclude

$$P_3\left(\sum_{k=1}^{m} W'_k > x_3\right) \le 2[1 - \Phi(x_3)]$$

on $H_n(\delta, 128)$ for all $n \ge n_\delta$ and $\delta \in (0, 1/4]$. The definition of s'_p and (3.108) yield that

$$P_3\left(\frac{1}{\sqrt{m}} \sum_{k=1}^{m} \eta'_{k12} > s_p\right) \le 2\left[1 - \Phi\left(\frac{s_p}{\sigma_{n3}} - \frac{1}{\sigma_{n3}\sqrt{m}} \sum_{k=1}^{m} \gamma_k\right)\right]$$

on $H_n(\delta, 128)$ for all $n \geq n_\delta$ and $\delta \in (0, 1/4]$. On $\tilde{K}_n$,

$$\frac{1}{\sqrt{m}\sigma_{n3}}\left|\sum_{k=1}^{m}\gamma_k\right| \leq \frac{\sqrt{7}\omega_n s_p}{2\sigma_{n3}} \leq \frac{\sqrt{7}}{4}\cdot\frac{s_p}{\sigma_{n3}}$$

since $0 \leq \omega_n < 1/2$ by (3.72). By the fact $1/4 \leq \sigma_{n3}^2 \leq 2$ on $H_n(\delta, 128)$ with $\delta \in (0, 1/4]$. Therefore, $\frac{s_p}{\sigma_{n3}} - \frac{1}{\sigma_{n3}\sqrt{m}}\sum_{k=1}^m \gamma_k \to \infty$ on $H_n(\delta, 128)$ as $n \to \infty$. Since $P(N(0, 1) \geq x) \leq e^{-x^2/2}$ for $x \geq 1$, we obtain that, given $\delta \in (0, 1/4]$,

$$P_3\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)^2 \leq 4\cdot\exp\left[-\left(\frac{s_p}{\sigma_{n3}} - \frac{1}{\sigma_{n3}\sqrt{m}}\sum_{k=1}^{m}\gamma_k\right)^2\right]$$

on $H_n(\delta, 128)$ as n is sufficiently large. By (3.101) and (3.103), given $\delta \in (0, 1/4]$,

$$\begin{aligned}
&E\left[P_3\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)^2\right] \\
&\quad\leq 4\cdot E\left\{I_{H_n(\delta,128)}\cdot\exp\left[-\frac{1}{\sigma_{n3}^2}\left(s_p - \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\gamma_k\right)^2\right]\right\} + o\left(\frac{1}{p^3}\right) \\
&\quad\leq 4\cdot E\left\{I_{\tilde{K}_n}\cdot\exp\left[-\frac{1}{(\sigma_{03}^2+\delta)}(s_p - V_n)^2\right]\right\} + o\left(\frac{1}{p^3}\right)
\end{aligned}$$

as n is sufficiently large, where $V_n = m^{-1/2}\sum_{k=1}^m \gamma_k$. Now

$$\begin{aligned}
&E\left\{I(V_n \leq 0)\cdot\exp\left[-\frac{1}{(\sigma_{03}^2+\frac{1}{5})}(s_p - V_n)^2\right]\right\} \\
&\quad\leq \exp\left[-\frac{s_p^2}{(\sigma_{03}^2+\frac{1}{5})}\right] = o\left(\frac{1}{p^3}\right)
\end{aligned}$$

since $1/2 \leq \sigma_{03}^2 \leq 1$ by (3.102). Denote

$$K'_n = \left\{0 < \frac{1}{\sqrt{m}}\sum_{k=1}^{m}\gamma_k < \frac{\sqrt{7}}{2}\omega_n s_p\right\}.$$

Then, for given $\delta \in (0, 1/5]$,

$$\begin{aligned}
&E\left[P_3\left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m}\eta'_{k12} > s_p\right)^2\right] \\
&\quad\leq 4\cdot E\left\{I_{K'_n}\cdot\exp\left[-\frac{1}{(\sigma_{03}^2+\delta)}(s_p - V_n)^2\right]\right\} + o\left(\frac{1}{p^3}\right)
\end{aligned}$$

as n is sufficiently large. By (3.71) and (3.72), $\sigma_{03}^2 + \omega_n^2 = 1$. The desired conclusion then follows from Lemma 3.15 and (3.95). □

Acknowledgment. We thank Professor Junshan Xie very much for many discussions.

REFERENCES

[1] ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1989). Two moments suffice for Poisson approximations: The Chen–Stein method. *Ann. Probab.* **17** 9–25. MR0972770

[2] CAI, T., FAN, J. and JIANG, T. (2013). Distributions of angles in random packing on spheres. *J. Mach. Learn. Res.* **14** 1837–1864. MR3104497

[3] CAI, T. T. and JIANG, T. (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *Ann. Statist.* **39** 1496–1525. MR2850210

[4] CAI, T. T. and JIANG, T. (2012). Phase transition in limiting distributions of coherence of high-dimensional random matrices. *J. Multivariate Anal.* **107** 24–39. MR2890430

[5] CHEN, L. H. Y., FANG, X. and SHAO, Q.-M. (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41** 262–293. MR3059199

[6] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. MR3161448

[7] EATON, M. L. (1983). *Multivariate Statistics*: *A Vector Space Approach*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, New York. MR0716321

[8] FAN, J., SHAO, Q.-M. and ZHOU, W.-X. (2018). Are discoveries spurious? Distributions of maximum spurious correlations and their applications. *Ann. Statist.* **46** 989–1017. MR3797994

[9] JIANG, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14** 865–880. MR2052906

[10] JIANG, T. (2004). The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā* **66** 35–48. MR2082906

[11] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. *Springer Series in Statistics*. Springer, New York. MR0691492

[12] LI, D., LIU, W.-D. and ROSALSKY, A. (2010). Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. *Probab. Theory Related Fields* **148** 5–35. MR2653220

[13] LI, D., QI, Y. and ROSALSKY, A. (2012). On Jiang's asymptotic distribution of the largest entry of a sample correlation matrix. *J. Multivariate Anal.* **111** 256–270. MR2944420

[14] LI, D. and ROSALSKY, A. (2006). Some strong limit theorems for the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **16** 423–447. MR2209348

[15] LINNIK, Y. V. (1961). On the probability of large deviations for the sums of independent variables. In *Proc*. 4*th Berkeley Sympos*. *Math*. *Statist*. *and Prob*., *Vol. II* 289–306. Univ. California Press, Berkeley, CA. MR0137142

[16] LIU, W.-D., LIN, Z. and SHAO, Q.-M. (2008). 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** 2337–2366. MR2474539

[17] MUIRHEAD, R. J. (2009). *Aspects of Multivariate Statistical Theory*. *Wiley Series in Probability and Statistics* **197**. Wiley, New York.

[18] SHAO, Q.-M. and WANG, Q. (2013). Self-normalized limit theorems: A survey. *Probab. Surv.* **10** 69–93. MR3161676

[19] SHAO, Q.-M. and ZHOU, W.-X. (2014). Necessary and sufficient conditions for the asymptotic distributions of coherence of ultra-high dimensional random matrices. *Ann. Probab.* **42** 623–648. MR3178469

[20] ZHOU, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.* **359** 5345–5363. MR2327033

DEPARTMENT OF OPERATIONS RESEARCH
AND FINANCIAL ENGINEERING
PRINCETON UNIVERSITY
205 SHERRED HALL
PRINCETON, NEW JERSEY 08544
USA
E-MAIL: jqfan@princeton.edu

SCHOOL OF STATISTICS
UNIVERSITY OF MINNESOTA
313 FORD HALL
224 CHURCH STREET SE
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: jiang040@umn.edu