DETERMINANT OF SAMPLE CORRELATION MATRIX WITH APPLICATION

By Tiefeng Jiang[1]

University of Minnesota

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be independent random vectors of a common p-dimensional normal distribution with population correlation matrix $\mathbf{R}_n$. The sample correlation matrix $\hat{\mathbf{R}}_n = (\hat{r}_{ij})_{p \times p}$ is generated from $\mathbf{x}_1, \ldots, \mathbf{x}_n$ such that $\hat{r}_{ij}$ is the Pearson correlation coefficient between the ith column and the jth column of the data matrix $(\mathbf{x}_1, \ldots, \mathbf{x}_n)'$. The matrix $\hat{\mathbf{R}}_n$ is a popular object in multivariate analysis and it has many connections to other problems. We derive a central limit theorem (CLT) for the logarithm of the determinant of $\hat{\mathbf{R}}_n$ for a big class of $\mathbf{R}_n$. The expressions of mean and the variance in the CLT are not obvious, and they are not known before. In particular, the CLT holds if p/n has a nonzero limit and the smallest eigenvalue of $\mathbf{R}_n$ is larger than $1/2$. Besides, a formula of the moments of $|\hat{\mathbf{R}}_n|$ and a new method of showing weak convergence are introduced. We apply the CLT to a high-dimensional statistical test.

1. Introduction. We first give a background of the sample correlation matrix, then state our main result. Two new tools are introduced and the method of the proof is elaborated by using them. At last an application is presented.

1.1. *Main results.* Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be a sequence of independent random vectors from a common distribution $N_p(\mu, \Sigma)$, that is, a p-dimensional normal distribution with mean vector μ and covariance matrix $\Sigma = (\sigma_{ij})_{p \times p}$. We always assume $\sigma_{ii} > 0$ for each i to avoid trivial cases. In order to take limit in n, the dimension p is assumed to depend on n and is written by p_n. Sometimes we write p for p_n and Σ for Σ_n to ease notation. The corresponding correlation matrix $\mathbf{R}_n = (r_{ij})_{p \times p}$ is defined by

$$(1.1) \qquad r_{ii} = 1 \quad \text{and} \quad r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

for all $1 \le i \ne j \le p$. Write $\mathbf{X} = (x_{ij})_{n \times p} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)'$. Let $\hat{r}_{ij}$ denote the Pearson correlation coefficient between $(x_{1i}, \ldots, x_{ni})'$ and $(x_{1j}, \ldots, x_{nj})'$, given by

$$(1.2) \qquad \hat{r}_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \cdot \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}},$$

where $\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ki}$ and $\bar{x}_j = \frac{1}{n}\sum_{k=1}^{n} x_{kj}$. Then the *sample correlation matrix* obtained from the random sample $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is defined by

$$(1.3) \qquad\qquad \hat{\mathbf{R}}_n := (\hat{r}_{ij})_{p \times p}.$$

A natural requirement for nonsingularity of $\hat{\mathbf{R}}_n$ is $n > p$. In this paper, we will prove that $\log |\hat{\mathbf{R}}_n|$, the logarithm of the determinant of $\hat{\mathbf{R}}_n$, satisfies the central limit theorem (CLT) under suitable assumption on Σ. Our theorem holds for a big class of $\mathbf{R}_n$ containing $\mathbf{I}$, in particular, for $\mathbf{R}_n$ generated by $\mathbf{x}_1$ as in (1.1) such that the p entries of $\mathbf{x}_1$ are far from independent.

In the next section, we will review the literature on $\hat{\mathbf{R}}_n$ and narrate our motivation to investigate $\log |\hat{\mathbf{R}}_n|$.

The sample correlation matrix $\hat{\mathbf{R}}_n$ is a popular target in multivariate analysis, a subject in statistics; see, for example, the classical books by Anderson (1958), Muirhead (1982) and Eaton (1983). Particularly, $|\hat{\mathbf{R}}_n|$ is the likelihood ratio test statistic for testing that the p entries of $\mathbf{X}_1$ are independent, but not necessarily identically distributed. If $\mathbf{R}_n = \mathbf{I}$, some are understood about $\hat{\mathbf{R}}_n$. For example, the density of $|\hat{\mathbf{R}}_n|$ is given by

$$\text{Constant} \cdot |R_n|^{(n-p-2)/2} dR_n;$$

see, for example, Theorem 5.1.3 in Muirhead (1982). Unfortunately, the density of the eigenvalues of $\hat{\mathbf{R}}_n$ is not known because $\hat{\mathbf{R}}_n$ does not have the orthogonal invariance that standard random matrices (e.g., the Hermite ensemble, the Laguerre ensemble and the Jacobi ensemble) usually have. As a consequence, unlike typical random matrices, the research on $\hat{\mathbf{R}}_n$ cannot rely on the density of its eigenvalues. This makes any efforts more involved.

On the other hand, the largest entries of $\hat{\mathbf{R}}_n$ is related to other statistical testing problems, random packing on spheres in $\mathbb{R}^p$ and the seventh most challenging problem in the twenty-first century by Smale [Jiang (2004a) and Cai, Fan and Jiang (2013)]. For example, Smale (2000) asks a packing of n points on the unit sphere such that the product of all pairwise distances among the n points almost attains its largest value. There are also investigations on the least moment condition to guarantee the limit law of the largest element of $\hat{\mathbf{R}}_n$; see, for example, Li and Rosalsky (2006), Zhou (2007) and Li, Liu and Rosalsky (2010).

Again, under $\mathbf{R}_n = \mathbf{I}$, some of the behaviors of the eigenvalues of $\hat{\mathbf{R}}_n$ are known. The empirical distribution of the eigenvalues of $\hat{\mathbf{R}}_n$ satisfies the Marčenko–Pastur law Jiang, 2004b; the largest eigenvalue of $\hat{\mathbf{R}}_n$ asymptotically satisfies the Tracy–Widom law [Bao, Pan and Zhou (2012)]; the quantity $\log |\hat{\mathbf{R}}_n|$ satisfies the CLT [Jiang and Yang (2013) and Jiang and Qi (2015)].

To our knowledge, little is understood on $\hat{\mathbf{R}}_n$ as $\mathbf{R}_n \neq \mathbf{I}$. In particular, there is no CLT for $\log |\hat{\mathbf{R}}_n|$ as $\mathbf{R}_n \neq \mathbf{I}$.

Our motivations in this paper to study the CLT of $\log|\hat{\mathbf{R}}_n|$ for $\boldsymbol{\Sigma} \neq \mathbf{I}$ are as follows. First, we plan to understand its behavior in general. Second, there is some recent interest to study the determinants of various random matrices. For instance, Tao and Vu (2012) and Nguyen and Vu (2014) derive the CLT for the determinants of Wigner matrices; Cai, Liang and Zhou (2015) derive the CLT for sample covariance matrices. So it is natural to look into $|\hat{\mathbf{R}}_n|$. Lastly, when we study the power for the hypothesis test H_0, the entries of $\mathbf{X}_1$ are completely independent; we need to derive the asymptotic distribution of $\log|\hat{\mathbf{R}}_n|$ under $\mathbf{R}_n \neq \mathbf{I}$. This problem occurs in the high-dimensional statistics, which together with machine learning forms the most important tools to study big data. We will further elaborate the test in Section 1.3.

Before introducing our main results, we need some notation. For matrix $\mathbf{M} = (m_{ij})_{p \times p}$, set $\|\mathbf{M}\|_\infty = \max_{1 \le i, j \le p} |m_{ij}|$ and $\|\mathbf{M}\|_2 = (\sum_{i,j} |m_{ij}|^2)^{1/2}$. We denote by $\|\mathbf{M}\|$ the spectral norm of $\mathbf{M}$, equivalently, the largest singular value of $\mathbf{M}$. The notation $|\mathbf{M}|$ stands for the determinant of $\mathbf{M}$. If $\mathbf{M}$ is symmetric, its smallest eigenvalue is denoted by $\lambda_{\min}(\mathbf{M})$. In particular, if $\mathbf{R}_n$ is positive definite, then $\lambda_{\min}(\mathbf{R}_n) \in (0, 1]$ due to the fact $r_{ii} = 1$ (see also Lemma 2.3). Keep in mind that $\mathbf{R}_n$ is nonrandom and $\hat{\mathbf{R}}_n$ is a random matrix constructed from data. We adopt the convention $\frac{0}{0} = 1$.

In this paper, we derive the following CLT for the sample correlation matrix.

THEOREM 1. *Assume* $p := p_n$ *satisfy that* $n > p + 4$ *and* $p \to \infty$. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be i.i.d. from* $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *and* $\hat{\mathbf{R}}_n$ *be as in* (1.3). *Suppose that* $\inf_{n \ge 6} \lambda_{\min}(\mathbf{R}_n) > \frac{1}{2}$. *Set*

$$\mu_n = \left(p - n + \frac{3}{2}\right) \log\left(1 - \frac{p}{n-1}\right) - \frac{n-2}{n-1}p + \log|\mathbf{R}_n|;$$

$$\sigma_n^2 = -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right] + \frac{2}{n-1}\operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2].$$

Then $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n$ *converges weakly to* $N(0, 1)$ *as* $n \to \infty$ *provided one of the following holds*:

  (i) $\inf_{n \ge 6} \frac{p_n}{n} > 0$;
  (ii) $\inf_{n \ge 6} \frac{1}{n}\operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2] > 0$;
  (iii) $\sup_{n \ge 6} \frac{p_n \|\mathbf{R}_n - \mathbf{I}\|_\infty}{\|\mathbf{R}_n - \mathbf{I}\|_2} < \infty$.

It is not intuitive to see or guess why μ_n and σ_n^2 have the expressions in Theorem 1. They come purely from computation.

Condition (i) is particularly valid if $\frac{p_n}{n} \to c \in (0, 1]$. Reading σ_n^2, conditions (ii) says the entries of $\mathbf{R}_n$ and/or p_n are large enough. Naturally, $\|\mathbf{R}_n - \mathbf{I}\|_2 \le p\|\mathbf{R}_n - \mathbf{I}\|_\infty$, so condition (iii) is equivalent to that $\|\mathbf{R}_n - \mathbf{I}\|_2 \le p\|\mathbf{R}_n - \mathbf{I}\|_\infty \le$

$K \|\mathbf{R}_n - \mathbf{I}\|_2$ for all $n \geq 6$, where K is a constant not depending on n. This condition says that almost all of the entries of $\mathbf{R}_n - \mathbf{I}$ are at the same magnitude.

The condition $\inf_{n \geq 6} \lambda_{\min}(\mathbf{R}_n) > \frac{1}{2}$ in Theorem 1 is used in Lemma 2.16. The number "$\frac{1}{2}$" comes essentially from the Gaussian density $\text{const} \cdot e^{-\frac{1}{2}x^2}$. Literally, it is related to "$\frac{1}{2}m$" in (1.5). It will be interesting to see if the limiting distribution is still the normal distribution when $\lambda_{\min}(\mathbf{R}_n) \leq \frac{1}{2}$.

Take $\mathbf{R}_n = \mathbf{I}$ in Theorem 1, then $p_n \|\mathbf{R}_n - \mathbf{I}\|_\infty = \|\mathbf{R}_n - \mathbf{I}\|_2 = 0$. So (iii) above holds and $\log |\mathbf{R}_n| = \text{tr}[(\mathbf{R}_n - \mathbf{I})^2] = 0$; the conclusion becomes Corollary 3 from Jiang and Qi (2015). Now let us look at a case where $\mathbf{R}_n$ is of "compound symmetry structure," that is, all of the off-diagonal entries of $\mathbf{R}_n$ are equal to $a \in [0, 1/2)$, then $\mathbf{R}_n$ has eigenvalues $(1-a) + ap, 1-a, \ldots, 1-a$. Hence $\lambda_{\min}(\mathbf{R}_n) = 1 - a > \frac{1}{2}$. Obviously, $\|\mathbf{R}_n - \mathbf{I}\|_\infty = a$ and $\|\mathbf{R}_n - \mathbf{I}\|_2 = a\sqrt{p(p-1)}$ for all $n \geq 6$. So condition (iii) from Theorem 1 holds. We then have the following conclusion.

COROLLARY 1. *Assume* $p := p_n$ *satisfy that* $n > p + 4$ *and* $p \to \infty$. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be i.i.d. from* $N_p(\mu, \Sigma)$ *and* $\hat{\mathbf{R}}_n$ *be as in* (1.3). *Assume the off-diagonal entries of* $\mathbf{R}_n$ *are all equal to* $a_n \geq 0$ *for each* n *and satisfy* $\sup_{n \geq 4} a_n < 1/2$. *Then* $(\log |\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n$ *goes to* $N(0, 1)$ *weakly as* $n \to \infty$, *where* μ_n *and* σ_n *are as in Theorem 1 with*

$$|\mathbf{R}_n| = (1 + a_n(p-1))(1 - a_n)^{p-1} \quad and \quad \text{tr}[(\mathbf{R}_n - \mathbf{I})^2] = p(p-1)a_n^2.$$

Now we apply Theorem 1 to the correlation matrix from $AR(1)$ and a banded correlation matrix. The notation $AR(1)$ stands for the autoregressive process of order one [see, e.g., Brockwell and Davis (2002)]. Both are very popular models in statistics.

COROLLARY 2. *Assuming the set-up in Theorem 1 with* $\inf_{n \geq 6} \frac{p_n}{n} > 0$, *the following is true*:

  (i) *Let* ρ *be a constant with* $|\rho| < \frac{1}{5}$. *If* $\mathbf{R}_n = (r_{ij})_{p \times p}$ *with* $r_{ij} = \rho^{|i-j|}$ *for all* $1 \leq i, j \leq p$, *then the CLT in Theorem 1 holds.*
  (ii) *Let* $k \geq 1$ *be a constant integer. Suppose* $\mathbf{R}_n = (r_{ij})_{p \times p}$ *satisfies* $r_{ij} = 0$ *for* $|j - i| > k$ *and* $\sup_{n \geq 6} \max_{i \neq j} |r_{ij}| < \frac{1}{4k}$. *Then the CLT in Theorem 1 holds.*

Corollary 2 will be checked at the end of Section 2.6. Besides the three special matrices of $\mathbf{R}_n$ studied in the above, there are a lot of other patterned matrices including the Toeplitz matrices, the Hankel matrices and the symmetric circulant matrices; see, for example, Brockwell and Davis (2002). One needs to check if the smallest eigenvalue of $\mathbf{R}_n$ is larger than $1/2$. If so, the CLT holds.

The proof of Theorem 1 is relatively lengthy. It needs some new tools. The tools and the method of the proof are introduced Section 1.2. In addition, we obtain some interesting matrix inequalities as byproducts. They are stated in Section 2.2.

1.2. *New tools and strategy of proofs.* Under $\Sigma = \mathbf{I}_p$, Jiang and Yang (2013) and Jiang and Qi (2015) prove that $L_n := (\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n$ converges weakly to $N(0,1)$; see Theorem 1 by taking $\Sigma = \mathbf{I}_p$. The argument is showing $\lim_{n\to\infty} Ee^{tL_n} = e^{t^2/2}$ for $t \in (-t_0, t_0)$. The proof may fail if the limit holds only for t in an half interval $[0, t_0]$ or $[-t_0, 0]$. Of course, the exact expression of $E[|\hat{\mathbf{R}}_n|^t]$ is crucial. So we prefer to get a closed form of $E[|\hat{\mathbf{R}}_n|^s]$ as $\Sigma \neq \mathbf{I}_p$. To see what we are able to obtain, define

$$(1.4) \qquad \Gamma_p(z) := \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(z - \frac{1}{2}(i-1)\right)$$

for complex number z with $\mathrm{Re}(z) > \frac{1}{2}(p-1)$; see, for example, page 62 from Muirhead (1982). Throughout the paper, the notation $\mathrm{Beta}(a, b)$ stands for the beta distribution with probability density $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$ for $0 \leq x \leq 1$, where $a > 0$ and $b > 0$ are two parameters.

PROPOSITION 1.1. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. with distribution $N_p(\mu, \Sigma)$ and $n = m + 1 > p$. Recall $\mathbf{R}_n$ and $\hat{\mathbf{R}}_n$ are as in (1.1) and (1.3), respectively. Set $\Delta_n = \mathbf{R}_n - \mathbf{I}$. Then*

$$(1.5) \qquad E[|\hat{\mathbf{R}}_n|^t] = \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2}+t)}\right)^p \cdot \frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})}$$
$$\cdot |\mathbf{R}_n|^t \cdot E[|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \ldots, V_p)|^{-(m/2)-t}]$$

for each $t > 0$, where $V_1, \ldots, V_p$ are i.i.d. with $\mathrm{Beta}(t, \frac{m}{2})$-distribution.

If $\Delta_n = \mathbf{0}$, or equivalently, $\mathbf{R}_n = \mathbf{I}$, Proposition 1.1 implies

$$(1.6) \qquad E[|\hat{\mathbf{R}}_n|^t] = \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2}+t)}\right)^p \cdot \frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})}$$

for all $t > 0$; see page 150 from Muirhead (1982) or page 492 from Wilks (1932). Jiang and Yang (2013) and Jiang and Qi (2015) further generalize this identity. Their results show that the formula is also true for a big range of negative value of t by using the Carlson uniqueness theorem of complex functions. We feel that Proposition 1.1 deserves a further understanding for the case $t < 0$.

In the special case that the off-diagonal entries of $\mathbf{R}_n$ are all equal to a as in Corollary 1, it is seen from Lemma 2.4 later that

$$(1.7) \qquad |\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \ldots, V_p)| = \left[\prod_{i=1}^{p}(1 - aV_i)\right] \cdot \left(1 + \sum_{i=1}^{p}\frac{aV_i}{1 - aV_i}\right)$$

for all $a \in [0, 1)$. For a general form of Δ_n, it seems there is not a better expression. Even worse, the right-hand side of (1.5) involves the probability distribution

Beta$(t, \frac{m}{2})$, which forces $t > 0$. We may consider a complex continuation by representing the expectation in terms of integrals in a similar way to the Riemann's zeta function or the Gamma function. However, we do see an advantage on the right-hand side of (1.5): the expectation is taken over a function of i.i.d. random variables. To make use of this fact while considering the restriction "$t > 0$," we develop a new tool.

We say the distribution of a random variable ξ is uniquely determined by its moments $\{E(\xi^p); \ p = 1, 2, \ldots\}$ if the following is true: for any random variable η with $E(\xi^p) = E(\eta^p)$ for all $p = 1, 2, \ldots$, the probability distributions of ξ and η are identical.

PROPOSITION 1.2. *Let* $\{X_n; \ n = 0, 1, \ldots\}$ *be random variables and* $\delta > 0$ *be a constant such that* $E e^{t X_n} < \infty$ *for all* $n \geq 0$ *and* $t \in [0, \delta]$. *Assume* $\sup_{n \geq 0} E(|X_n|^p) < \infty$ *for each* $p \geq 1$. *Assume* $\lim_{n \to \infty} E e^{t X_n} = E e^{t X_0}$ *for all* $t \in [0, \delta]$. *If the distribution of* X_0 *can be determined uniquely by moments* $\{E(X_0^p); \ p = 1, 2, \ldots\}$, *then* X_n *converges weakly to* X_0 *as* $n \to \infty$.

A classical result says that the above lemma holds if "$t \in [0, \delta]$" is replaced by a stronger assumption "$|t| \leq \delta$"; see, for example, Billingsley (1986). It is interesting to see that the weak convergence in Proposition 1.2 is still valid under a "one-sided" condition on moment generating functions if some extra requirements are fulfilled.

With the above two conclusions, we now are ready to state the method of the proof of Theorem 1. By Proposition 1.2, it suffices to show, there exists $s_0 > 0$ such that

$$(1.8) \qquad \log E \exp\left(\frac{\log |\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n} s\right) = \log E\big[|\hat{\mathbf{R}}_n|^t\big] - \mu_n t \to \frac{s^2}{2}$$

for all $s \in (0, s_0)$, where $t = \frac{s}{\sigma_n}$, and

$$(1.9) \qquad \sup_{n \geq 6} E\left[\left(\frac{\log |\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n}\right)^{2k}\right] < \infty$$

for each integer $k \geq 1$. To get (1.8), by Proposition 1.1, we need to work on

$$(1.10) \qquad \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2} + t)}\right)^p \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{m}{2})} \quad \text{and}$$

$$I_n := E\big[|\mathbf{I} + \boldsymbol{\Delta}_n \cdot \mathrm{diag}(V_1, \ldots, V_p)|^{-(m/2)-t}\big].$$

The first one is understood well enough by Jiang and Qi (2015), so it suffices to derive an asymptotic formula for I_n. However, it seems hard to evaluate I_n directly.

This can be convinced easily by the explicit case from (1.7). To compute I_n, by setting $Q_n := |\mathbf{I} + \mathbf{\Delta}_n \cdot \text{diag}(V_1, \ldots, V_p)|^{-\frac{m}{2}-t}$, we will show

(1.11) $e^{-h_n} Q_n$ converges in probability to 1;

(1.12) $\{e^{-h_n} Q_n; n \geq 6\}$ is uniformly integrable,

where h_n is an explicit constant depending on t. The two assertions imply $E Q_n \sim e^{h_n}$. Then (1.8) follows. The statements in (1.9), (1.11) and (1.12) are proved in Sections 2.3, 2.4 and 2.5, respectively. A detailed account of the above, which forms the proof of Theorem 1, is presented in Section 2.6.

1.3. *An application to a high-dimensional likelihood ratio test.* Recall $\mathbf{x}_1, \ldots,$ $\mathbf{x}_n$ are i.i.d. random observations with a common p-variate normal distribution $N_p(\mu, \mathbf{\Sigma})$. The according correlation matrix is given in (1.1). In application, there is only one correlation matrix associated with $\mathbf{\Sigma}$. So we will drop the subscript to write it as $\mathbf{R}$ instead of $\mathbf{R}_n$ in this subsection. We test that the p components of $\mathbf{x}_1$ are independent, which is equivalent to

(1.13) $H_0 : \mathbf{R} = \mathbf{I}_p$ vs $H_a : \mathbf{R} \neq \mathbf{I}_p.$

When p is fixed, the chi-square approximation holds under H_0:

(1.14) $-\left(n - 1 - \dfrac{2p + 5}{6}\right) \log |\hat{\mathbf{R}}_n|$ converges to $\chi^2_{p(p-1)/2}$

in distribution as $n \to \infty$; see, for example, Bartlett (1954) or page 40 from Morrison (1967). According to the latter literature, the rejection region of the like-lihood ratio test for (1.13) is

$$|\hat{\mathbf{R}}_n| \leq c_\alpha,$$

where c_α is determined so that the test has significance level of α. For many modern data, the population dimension p is very large relative to the sample size n. The approximation (1.14) is far from accurate. To correct this, Jiang and Yang (2013) and Jiang and Qi (2015) prove the CLT as in Theorem 1 for the case $\mathbf{\Sigma} = \mathbf{I}_p$. For a hypothesis testing problem, we not only need to know the asymptotic distribution of the test statistic under H_0 to make a decision, we also need to carry out another procedure, that is, to minimize (type II) error, or equivalently, make the so-called power as large as possible. So we have to develop the CLT for the case $H_a : \mathbf{R} \neq \mathbf{I}_p$. Theorem 1 provides us with the exact result. Define

$$\mu_{n,0} = \left(p - n + \frac{3}{2}\right) \log\left(1 - \frac{p}{n-1}\right) - \frac{n-2}{n-1} p ;$$
$$\sigma^2_{n,0} = -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right].$$

According to Theorem 1 at $\mathbf{R} = \mathbf{I}_p$, the asymptotic size-$\alpha$ test is given by $R = \{\log |\hat{\mathbf{R}}_n| \leq c_\alpha\}$ with $c_\alpha = \mu_{n,0} + \sigma_{n,0}\Phi^{-1}(\alpha)$, where $\Phi(x) = (2\pi)^{-1/2} \times \int_{-\infty}^{x} e^{-t^2/2}\, dt$. By Theorem 1 again, the power function for the test is

$$\beta(\mathbf{R}) = P\big(\log |\hat{\mathbf{R}}_n| \leq c_\alpha | \mathbf{R}\big) \sim \Phi\left(\frac{c_\alpha - \mu_n}{\sigma_n}\right)$$

for all correlation matrix $\mathbf{R}$ satisfying the conditions in Theorem 1.

2. Proofs. The proof of Theorem 1 will follow the scheme described in Section 1.2. Since the proof is relatively lengthy, we now elaborate each step. In Section 2.1, we will prove the major tools: Propositions 1.1 and 1.2. In Section 2.2, auxiliary results on matrices and Gamma functions are given. In Section 2.3, we will derive a formula for the moments of logarithms of the determinants of sample correlation matrices; the inequality (1.9) will be proved. In Section 2.4, we study the convergence of random determinants, and the assertion (1.11) will be verified. In Section 2.5, the uniform integrability of random determinants is proved, and the statement (1.12) will be justified. We will finally prove Theorem 1 and Corollary 2 in Section 2.6.

2.1. *Proofs of major tools*: *Propositions 1.1 and 1.2.* Suppose $\xi_1, \ldots, \xi_m$ are i.i.d. $\mathbb{R}^p$-valued random vectors with distribution $N_p(\mathbf{0}, \Sigma)$, where $\mathbf{0} \in \mathbb{R}^p$ and Σ is a $p \times p$ positive definite matrix. We then say the $p \times p$ matrix $W = (\xi_1, \ldots, \xi_m)(\xi_1, \ldots, \xi_m)'$ follows the Wishart distribution and denote it by $W \sim W_p(m, \Sigma)$. For two $r \times s$ random matrices $\mathbf{M}_1$ and $\mathbf{M}_2$, we use notation $\mathbf{M}_1 \overset{d}{=} \mathbf{M}_2$ to represent that the two sets of rs random variables in order have the same joint distribution.

LEMMA 2.1. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. with distribution $N_p(\mu, \Sigma)$ and $n = m + 1 > p$. Let $\mathbf{W} = (W_{ij})$ follow the Wishart distribution $W_p(m, \Sigma)$. Review $\hat{\mathbf{R}}_n$ in (1.3). Then*

$$(2.1) \qquad \hat{\mathbf{R}}_n \overset{d}{=} \left(\frac{W_{ij}}{\sqrt{W_{ii}} \cdot \sqrt{W_{jj}}}\right)_{p \times p}.$$

In particular, $\hat{\mathbf{R}}_n$ is a positive definite matrix with $0 < |\hat{\mathbf{R}}_n| \leq 1$.

PROOF. Write $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)' = (x_{ij})_{n \times p} = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$. Define $\mathbf{e} = (1, \ldots, 1)' \in \mathbb{R}^n$. Review the definition of the Pearson correlation coefficient in (1.2); it is easily seen that the correlation between $\mathbf{y}_i = (x_{1i}, \ldots, x_{ni})'$ and $\mathbf{y}_j = (x_{1j}, \ldots, x_{nj})'$ is the same as that between $\mathbf{y}_i + a\mathbf{e}$ and $\mathbf{y}_j + b\mathbf{e}$ for any $a \in \mathbb{R}$ and $b \in \mathbb{R}$. Thus, without loss of generality, we can assume $\mu = \mathbf{0}$.

Set $\mathbf{M} = \mathbf{I}_n - \frac{1}{n}\mathbf{e}\mathbf{e}'$. Recall $\hat{\mathbf{R}}_n = (\hat{r}_{ij})_{p \times p}$. From (1.2) and (1.3), it is trivial to check that

$$(2.2) \qquad \hat{r}_{ij} = \frac{(\mathbf{M}\mathbf{y}_i)'(\mathbf{M}\mathbf{y}_j)}{\|\mathbf{M}\mathbf{y}_i\| \cdot \|\mathbf{M}\mathbf{y}_j\|}$$

for all $1 \le i, j \le p$. Let $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n$ be i.i.d. random vectors with distribution $N_p(\mathbf{0}, \mathbf{I}_p)$. Then

$$
\begin{aligned}
(\mathbf{M}\mathbf{y}_1, \ldots, \mathbf{M}\mathbf{y}_p) &= \mathbf{M}(\mathbf{x}_1, \ldots, \mathbf{x}_n)' \\
&\stackrel{d}{=} \mathbf{M}(\Sigma^{1/2}\tilde{\mathbf{x}}_1, \ldots, \Sigma^{1/2}\tilde{\mathbf{x}}_n)' \\
&= \mathbf{M}(\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n)'\Sigma^{1/2},
\end{aligned}
$$

(2.3)

where $\Sigma^{1/2}$ is a positive definite matrix satisfying $\Sigma^{1/2} \cdot \Sigma^{1/2} = \Sigma$. Since $\mathbf{M}^2 = \mathbf{M}$ and $\mathrm{tr}(\mathbf{M}) = n - 1 = m$, there is an $n \times n$ orthogonal matrix Γ such that $\mathbf{M} = \Gamma \, \mathrm{diag}(\mathbf{I}_m, 0)\Gamma'$. Hence, by the invariance property of $N_p(\mathbf{0}, \mathbf{I}_p)$,

$$
\begin{aligned}
\mathbf{M}(\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n)' &\stackrel{d}{=} \Gamma \, \mathrm{diag}(\mathbf{I}_m, \mathbf{0})(\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_n)' \\
&= \Gamma \begin{pmatrix} \mathbf{U} \\ \mathbf{0} \end{pmatrix} = \Gamma_1 \mathbf{U},
\end{aligned}
$$

where $\mathbf{0} = (0, \ldots, 0)' \in \mathbb{R}^p$, all of the entries of $\mathbf{U} = \mathbf{U}_{m \times p}$ are i.i.d. standard normals and Γ_1 is the $n \times m$ matrix by deleting the last column of Γ. This together with (2.3) implies that

$$(\mathbf{M}\mathbf{y}_1, \ldots, \mathbf{M}\mathbf{y}_p) \stackrel{d}{=} \Gamma_1 \mathbf{U}\Sigma^{1/2}.$$

Use the fact $\Gamma_1'\Gamma_1 = \mathbf{I}_m$ to have

$$
\begin{aligned}
(W_{ij})_{p \times p} &:= (\mathbf{M}\mathbf{y}_1, \ldots, \mathbf{M}\mathbf{y}_p)'(\mathbf{M}\mathbf{y}_1, \ldots, \mathbf{M}\mathbf{y}_p) \\
&\stackrel{d}{=} (\Sigma^{1/2}\mathbf{U}')(\Sigma^{1/2}\mathbf{U}')'.
\end{aligned}
$$

In particular, $W_{ij} = (\mathbf{M}\mathbf{y}_i)'(\mathbf{M}\mathbf{y}_j)$ and $W_{ii} = \|\mathbf{M}\mathbf{y}_i\|^2$ for all i, j. Evidently, the m columns of the $p \times m$ matrix $\Sigma^{1/2}\mathbf{U}'$ are i.i.d. random variables with distribution $N_p(\mathbf{0}, \Sigma)$. The conclusion (2.1) then follows from (2.2) and the definition of a Wishart matrix.

Finally, write $\hat{\mathbf{R}}_n = \mathbf{Q}\mathbf{W}\mathbf{Q}$ where $\mathbf{Q} := \mathrm{diag}(W_{11}^{-1/2}, \ldots, W_{pp}^{-1/2})$. For $m \ge p$, the Wishart matrix $\mathbf{W}$ is positive definite. Thus, $\hat{\mathbf{R}}_n$ is positive definite. Since all of the diagonal entries of $\hat{\mathbf{R}}_n$ are equal to 1, we see $|\hat{\mathbf{R}}_n| \le \prod_{i=1}^{p}(\hat{\mathbf{R}}_n)_{ii} = 1$ by the Hadamard inequality. $\square$

We now use Lemma 2.1 to show Proposition 1.1.

PROOF OF PROPOSITION 1.1.   Let us continue to use the notation in the proof of Lemma 2.1. By Theorem 3.2.1 from Muirhead (1982), the density function of **W** is given by

(2.4) $$C_{m,p} \cdot e^{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}W)} \cdot |W|^{(m-p-1)/2}$$

for all positive definite matrix W, where $\Gamma_p(\frac{1}{2}m)$ is the multivariate Gamma function defined at (1.4) and

$$C_{m,p} = \frac{1}{2^{mp/2}\Gamma_p(\frac{1}{2}m)|\Sigma|^{m/2}}.$$

By Lemma 2.1, $|\hat{\mathbf{R}}_n|$ has the same distribution as that of $|\mathbf{W}| \cdot (\prod_{i=1}^{p} W_{ii})^{-1}$. Thus,

$$E[|\hat{\mathbf{R}}_n|^t] = C_{m,p} \int_{W>0} \frac{|W|^t}{(\prod_{i=1}^{p} W_{ii})^t} e^{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}W)} \cdot |W|^{(m-p-1)/2} dW$$

$$= C_{m,p} \int_{W>0} \frac{|W|^{t+(m-p-1)/2}}{(\prod_{i=1}^{p} W_{ii})^t} e^{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}W)} dW.$$

Write

$$W_{ii}^{-t} = \frac{1}{2^t\Gamma(t)} \int_0^\infty e^{-\frac{1}{2}W_{ii}y_i} y_i^{t-1} dy_i$$

for all i (*this is the step we have to assume $t > 0$*). It follows that

$E[|\hat{\mathbf{R}}_n|^t]$

$$= C'_{m,p} \int_{(\mathbb{R}_+)^p} \prod_{i=1}^{p} y_i^{t-1} dy_1 \cdots dy_p \int_{W>0} |W|^{t+(m-p-1)/2} e^{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}W)}$$

$$\cdot e^{-\frac{1}{2}\sum_{i=1}^{p} W_{ii}y_i} dW,$$

where $\mathbb{R}_+ = [0, \infty)$ and

$$C'_{m,p} = C_{m,p} \cdot \frac{1}{2^{pt}\Gamma(t)^p}.$$

Write $Y = \operatorname{diag}(y_1, \ldots, y_p)$ and $\sum_{i=1}^{p} W_{ii}y_i = \operatorname{tr}(YW)$. Then

(2.5)

$$E[|\hat{\mathbf{R}}_n|^t] = C'_{m,p} \int_{(\mathbb{R}_+)^p} \prod_{i=1}^{p} y_i^{t-1} dy_1 \cdots dy_p \int_{W>0} |W|^{t+(m-p-1)/2}$$

$$\times e^{-\frac{1}{2}\operatorname{tr}((\Sigma^{-1}+Y)W)} dW$$

$$= C'_{m,p} \left[\Gamma_p\left(t + \frac{m}{2}\right) 2^{pt+(mp/2)} \right]$$

$$\times \int_{(\mathbb{R}_+)^p} |\Sigma^{-1} + Y|^{-((m/2)+t)} \prod_{i=1}^{p} y_i^{t-1} dy_1 \cdots dy_p$$

by Theorem 2.1.11 from Muirhead (1982) (taking $a = t + \frac{m}{2}$) provided $t > -\frac{1}{2}(m + 1 - p)$. From (2.4), the constant in front of the integral from (2.5) becomes

$$\frac{1}{2^{mp/2}\Gamma_p(\frac{1}{2}m)|\Sigma|^{m/2}} \cdot \frac{1}{2^{pt}\Gamma(t)^p} \cdot \Gamma_p\left(t + \frac{m}{2}\right)2^{pt+(mp/2)}$$

$$= \frac{1}{|\Sigma|^{m/2}\Gamma(t)^p} \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{1}{2}m)}.$$

Write $|\Sigma^{-1} + \mathbf{Y}| = |\Sigma|^{-1}|\mathbf{I} + \Sigma\mathbf{Y}|$. Then

$$E\big[|\hat{\mathbf{R}}_n|^t\big]$$

(2.6)
$$= \frac{|\Sigma|^t}{\Gamma(t)^p} \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{m}{2})} \int_{[0,\infty)^p} |\mathbf{I} + \Sigma Y|^{-(m/2)-t} \prod_{i=1}^p y_i^{t-1} \, dy_1 \cdots dy_p$$

for any $t > 0$.

We next will transfer (2.6) to the right-hand side of (1.5). Recall (1.1) and write $\Sigma = (\sigma_{ij})$. Then $\Sigma = \mathbf{L}\mathbf{R}_n\mathbf{L}$ where $\mathbf{L} := \text{diag}(\sigma_{11}^{1/2}, \ldots, \sigma_{pp}^{1/2})$. Noticing $\mathbf{L}$ and $\mathbf{Y}$ are both diagonal matrices. Plugging this into (2.6), we see that

$$E\big[|\hat{\mathbf{R}}_n|^t\big]$$

$$= \frac{|\mathbf{R}_n|^t(\prod_{i=1}^p \sigma_{ii})^t}{\Gamma(t)^p} \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{m}{2})} \int_{[0,\infty)^p} |\mathbf{I} + \mathbf{R}_n\mathbf{L}^2 Y|^{-(m/2)-t}$$

$$\cdot \prod_{i=1}^p y_i^{t-1} \, dy_1 \cdots dy_p.$$

Obviously, $\mathbf{L}^2 Y = \text{diag}(\sigma_{11}y_1, \ldots, \sigma_{pp}y_p)$. Set $s_i = \sigma_{ii}y_i$ for all i and $\mathbf{S} = \text{diag}(s_1, \ldots, s_p)$. We obtain

$$E\big[|\hat{\mathbf{R}}_n|^t\big]$$

$$= \frac{|\mathbf{R}_n|^t}{\Gamma(t)^p} \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{m}{2})} \int_{[0,\infty)^p} |\mathbf{I} + \mathbf{R}_n\mathbf{S}|^{-(m/2)-t} \prod_{i=1}^p s_i^{t-1} \, ds_1 \cdots ds_p$$

$$= \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2} + t)}\right)^p \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{m}{2})}$$

$$\cdot |\mathbf{R}_n|^t \cdot \left(\frac{\Gamma(\frac{m}{2} + t)}{\Gamma(t)\Gamma(\frac{m}{2})}\right)^p \int_{[0,\infty)^p} |\mathbf{I} + \mathbf{R}_n\mathbf{S}|^{-(m/2)-t} \prod_{i=1}^p s_i^{t-1} \, ds_1 \cdots ds_p.$$

Write $\mathbf{\Delta}_n = \mathbf{R}_n - \mathbf{I}$. Then $|\mathbf{I} + \mathbf{R}_n \mathbf{S}| = |\mathbf{I} + \mathbf{S} + \mathbf{\Delta}_n \mathbf{S}| = |\mathbf{I} + \mathbf{S}| \cdot |\mathbf{I} + \mathbf{\Delta}_n \mathbf{S}(\mathbf{I} + \mathbf{S})^{-1}|$. Thus, the last integral equals

$$\int_{[0,\infty)^p} |\mathbf{I} + \mathbf{\Delta}_n \mathbf{S}(\mathbf{I} + \mathbf{S})^{-1}|^{-(m/2)-t} \prod_{i=1}^{p} (1 + s_i)^{-(m/2)-t} s_i^{t-1} \, ds_1 \cdots ds_p.$$

Set $x_i = s_i/(1 + s_i)$ for $1 \le i \le p$. Then $1 + s_i = \frac{1}{1-x_i}$ and $s_i = \frac{x_i}{1-x_i}$. The integral above is equal to

$$\int_{[0,1]^p} |\mathbf{I} + \mathbf{\Delta}_n \cdot \mathrm{diag}(x_1, \ldots, x_p)|^{-(m/2)-t} \prod_{i=1}^{p} x_i^{t-1} (1 - x_i)^{\frac{m}{2}-1} \, dx_1 \cdots dx_p.$$

This says that

$$E\big[|\hat{\mathbf{R}}_n|^t\big] = \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2} + t)} \right)^p \cdot \frac{\Gamma_p(\frac{m}{2} + t)}{\Gamma_p(\frac{m}{2})}$$

$$\cdot |\mathbf{R}_n|^t \cdot E|\mathbf{I} + \mathbf{\Delta}_n \cdot \mathrm{diag}(V_1, \ldots, V_p)|^{-(m/2)-t},$$

where $V_1, \ldots, V_p$ are i.i.d. with Beta$(t, \frac{m}{2})$-distribution. $\square$

PROOF OF PROPOSITION 1.2.   We first claim that

$$(2.7) \qquad\qquad\qquad \sup_{n \ge 0} E\big\{|X_n|^p e^{tX_n}\big\} < \infty$$

for every integer $p \ge 1$ and every $t \in [0, \delta)$, where δ is as in the statement of the proposition. Obviously, $|X_n|^p e^{tX_n} \le |X_n|^{2p} e^{tX_n} + e^{tX_n}$. So, by the given condition $\sup_{n \ge 0} E(|X_n|^p) < \infty$ for each $p \ge 1$, we only need to prove (2.7) for even integer $p \ge 1$ and $t \in (0, \delta)$. In fact, for $\delta' \in (0, \delta)$, set $\delta'' = (\delta - \delta')/2$. By the Taylor expansion, $e^{\delta'' X_n} \ge \frac{(\delta'')^p}{p!} (X_n)^p$ for all even integers $p \ge 1$ if $X_n \ge 0$. Consequently, for all $t \in (0, \delta']$,

$$\begin{aligned}
(X_n)^p e^{tX_n} &= (X_n)^p e^{tX_n} I(X_n \ge 0) + (X_n)^p e^{tX_n} I(X_n < 0) \\
&\le \frac{p!}{\delta''^p} \cdot e^{(t+\delta'')X_n} I(X_n \ge 0) + \frac{1}{t^p} \cdot \sup_{x \le 0} \{x^p e^x\} \\
&\le \frac{p!}{\delta''^p} \cdot e^{\delta X_n} + \frac{1}{t^p} \cdot \sup_{x \le 0} \{x^p e^x\}
\end{aligned}$$

for all even integer $p \ge 1$. We then get (2.7) by taking expectations.

The condition $\sup_{n \ge 1} E(X_n^2) < \infty$ implies that $\{X_n; n \ge 1\}$ is tight. Thus, for any subsequence $\{X_{n_k}; k \ge 1\}$ such that X_{n_k} converges weakly to Y as $k \to \infty$, to prove the lemma, it is enough to show that Y and X_0 have the same distribution.

Now we assume that X_{n_k} converging weakly to Y as $k \to \infty$. Since $e^{tX_{n_k}}$ converges weakly to $e^{tY} \ge 0$, by assumption and the Fatou lemma, $Ee^{tY} \le Ee^{tX_0}$ for

all $t \in [0, \delta]$. This implies that $\{e^{tY}, e^{tX_{n_k}}; \; k \geq 1\}$ is uniformly integrable for each $t \in [0, \delta)$ by Hölder's inequality. Thus,

$$(2.8) \qquad Ee^{tY} = \lim_{k \to \infty} Ee^{tX_{n_k}} = Ee^{tX_0}$$

for all $t \in [0, \delta)$. Furthermore, $|X_{n_k}|^p e^{tX_{n_k}} \to |Y|^p e^{tY}$ weakly for every $t \in \mathbb{R}$. By (2.7) and the Fatou lemma again, we know that $E\{|Y|^p e^{tY}\} < \infty$ for every $p \geq 1$ and every $t \in [0, \delta)$. This and (2.7) imply

$$(2.9) \qquad E\{|Y|^p e^{tY}\} < \infty \quad \text{and} \quad E\{|X_0|^p e^{tX_0}\} < \infty$$

for every $p \geq 1$ and every $t \in [0, \delta)$. Fix $\delta_1 \in (0, \delta)$. Observe $|Y|^p e^{tY} \leq |Y|^p e^{\delta_1 Y} + |Y|^p$ for all $t \in [0, \delta_1]$. Then (2.9) can be strengthened to

$$(2.10) \qquad E\left(\sup_{t \in [0, \delta_1]} |Y|^p e^{tY} \right) < \infty \quad \text{and} \quad E\left(\sup_{t \in [0, \delta_1]} |X_0|^p e^{tX_0} \right) < \infty$$

for each $p \geq 1$. Now, by the mean-value theorem from calculus, for any $t \neq t_0 \in [0, \delta_1]$ and each integer $j \geq 0$, there exists ξ between t and t_0 such that $\frac{Y^j e^{tY} - Y^j e^{t_0 Y}}{t - t_0} = Y^{j+1} e^{\xi Y}$. Therefore,

$$\lim_{t \to t_0} \frac{Y^j e^{tY} - Y^j e^{t_0 Y}}{t - t_0} = Y^{j+1} e^{t_0 Y}$$

and

$$\sup_{t \neq t_0 \in [0, \delta_1]} \left| \frac{Y^j e^{tY} - Y^j e^{t_0 Y}}{t - t_0} \right| \leq \sup_{t \in [0, \delta_1]} \left\{ |Y|^{j+1} e^{tY} \right\}$$

for every integer $j \geq 0$. By (2.8), (2.10), the dominated convergence theorem and induction,

$$E(Y^p e^{tY}) = \frac{d^p (Ee^{tY})}{dt^p} = \frac{d^p (Ee^{tX_0})}{dt^p} = E(X_0^p e^{tX_0})$$

for each integer $p \geq 1$ and all $t \in [0, \delta)$. At $t = 0$, the derivatives above are understood as right derivatives. Therefore, $E(Y^p) = E(X_0^p)$ for all integer $p \geq 1$ by taking $t = 0$. By assumption, the moments of X_0 uniquely determine the distribution of X_0, we conclude that Y and X_0 have the same distribution. $\quad \square$

2.2. *Auxiliary results on matrices and Gamma functions.* In this section, we prove some facts on matrices and Gamma functions. We will also verify (1.7) in Lemma 2.4.

We say $\mathbf{R} = (r_{ij})_{p \times p}$ is a *correlation matrix* if $\mathbf{R}$ is a nonnegative definite matrix with $r_{ii} = 1$ for all $1 \leq i \leq p$. The following auxiliary results on matrices seem interesting in their own way.

LEMMA 2.2. *Let* $\mathbf{R} = (r_{ij})_{p \times p}$ *be a correlation matrix. Then we have* $\max_{1 \leq i \leq p} \sum_{j \neq i} r_{ij}^2 \leq \|\mathbf{R}\|$.

PROOF OF LEMMA 2.2. For a square matrix $\mathbf{M}$, let $\lambda_{\max}(\mathbf{M})$ and $\lambda_{\min}(\mathbf{M})$ be the largest and smallest eigenvalues of $\mathbf{M}$, respectively. By definition, $\|\mathbf{M}\| = \lambda_{\max}(\mathbf{M})$ if $\mathbf{M}$ is nonnegative definite.

First, assume $\mathbf{R}$ is invertible. Let $\mathbf{R}_1$ be the $(p-1) \times (p-1)$ upper-left corner of $\mathbf{R}$ and $\mathbf{y} = (r_{p1}, \dots, r_{p\,p-1})$. Then

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{y}' \\ \mathbf{y} & 1 \end{pmatrix}.$$

By the formula for the determinant of a partitioned matrix [see, e.g., page 470 from Bai and Silverstein (2010)], $|\mathbf{R}| = |\mathbf{R}_1| \cdot (1 - \mathbf{y}\mathbf{R}_1^{-1}\mathbf{y}')$. Since $\mathbf{R}$ is positive definite, $|\mathbf{R}| > 0$ and $|\mathbf{R}_1| > 0$. It follows that

$$1 \geq \mathbf{y}\mathbf{R}_1^{-1}\mathbf{y}'$$
$$\geq \lambda_{\min}(\mathbf{R}_1^{-1}) \cdot |\mathbf{y}|^2$$
$$= \frac{|\mathbf{y}|^2}{\lambda_{\max}(\mathbf{R}_1)} \geq \frac{|\mathbf{y}|^2}{\lambda_{\max}(\mathbf{R})},$$

where in the last step we apply the interlacing theorem to the situation that $\mathbf{R}_1$ is a submatrix of $\mathbf{R}$; see, for example, page 185 from Horn and Johnson (1985). It leads to

$$(2.11) \qquad \sum_{j \neq p} r_{pj}^2 \leq \|\mathbf{R}\|.$$

For any $1 \leq i \leq p - 1$, we simply exchange the ith row and pth row and then exchange the ith column and pth column. Since the new and old matrices are conjugate to each other, they have the same eigenvalues. We then apply (2.11) to get the desired conclusion.

If $\mathbf{R}$ is not invertible, consider $\mathbf{R}(x) := \frac{1}{1+x}(\mathbf{R} + x\mathbf{I})$ for $x > 0$. Then $\mathbf{R}(x)$ is a correlation matrix and is positive definite for any $x > 0$. By the proved conclusion,

$$\frac{1}{(1+x)^2} \max_{1 \leq i \leq p} \sum_{j \neq i} r_{ij}^2 \leq \frac{1}{1+x} \cdot \|\mathbf{R} + x\mathbf{I}\|.$$

Recall $\|\mathbf{M}\|$ is continuous in the entries of $\mathbf{M}$. The proof is complete by letting $x \downarrow 0$. □

LEMMA 2.3. *Let $\mathbf{R} = (r_{ij})_{p \times p}$ be a correlation matrix. Then $\lambda_{\min}(\mathbf{R}) \in [0, 1]$ and $\lambda_{\min}(\mathbf{R}) = 1$ if and only if $\mathbf{R} = \mathbf{I}$.*

PROOF. Let $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$ be the eigenvalues of $\mathbf{R}$. Take $x_0 = (1, 0, \dots, 0)' \in \mathbb{R}^p$. By the Rayleigh–Ritz formula,

$$\lambda_{\min}(\mathbf{R}) = \min_{\|x\|=1} x'\mathbf{R}x \leq x_0'\mathbf{R}x_0 = 1.$$

If $\lambda_{\min}(\mathbf{R}) = 1$, then $\lambda_1 \geq \cdots \geq \lambda_p \geq 1$. At the same time, the Hadamard inequality says $0 \leq \lambda_1 \cdots \lambda_p = |\mathbf{R}| \leq r_{11} \cdots r_{pp} = 1$. Hence, $\lambda_1 = \cdots = \lambda_p = 1$. This implies $\mathbf{R} = \mathbf{I}$. $\quad\square$

We now verify (1.7). The notation here is slightly different from (1.7) for a general purpose.

LEMMA 2.4. *Let* $\mathbf{A} = (a_{ij})_{k \times k}$ *with* $a_{ij} = 1$ *for all* $1 \leq i, j \leq k$. *Let* $b_1, \ldots, b_k \in \mathbb{R}$ *and* $\mathbf{B} = \operatorname{diag}(b_1, \ldots, b_k)$. *Then*

$$|\mathbf{I} + x(\mathbf{A} - \mathbf{I})\mathbf{B}| = \left[\prod_{i=1}^{k}(1 - b_i x)\right] \cdot \left(1 + \sum_{i=1}^{k} \frac{b_i x}{1 - b_i x}\right)$$

for all $x \in \mathbb{R}$ *satisfying* $b_i x \neq 1$ *for any* $1 \leq i \leq k$.

Let $f(x) = \prod_{i=1}^{k}(1 - b_i x)$. Then the right-hand side above is equal to $f(x) - xf'(x)$.

PROOF OF LEMMA 2.4.   Since $\mathbf{A}$ is of rank one, the only nonzero eigenvalue is equal to $\operatorname{tr}(\mathbf{A}) = k$. It is easy to see that the corresponding eigenvector is $\mathbf{h} = (1, \ldots, 1)'/\sqrt{k} \in \mathbb{R}^k$. Then, by the spectral theorem of symmetric matrices, there exists an orthogonal matrix $\mathbf{O}$ such that $\mathbf{A} = \mathbf{O} \operatorname{diag}(k, 0, \ldots, 0)\mathbf{O}'$, where the first column of $\mathbf{O}$ is $\mathbf{h}$. Observe

$$|\mathbf{I} + x(\mathbf{A} - \mathbf{I})\mathbf{B}|$$
$$= |\operatorname{diag}(1 - b_1 x, \ldots, 1 - b_k x) + x\mathbf{A}\mathbf{B}|$$

(2.12)
$$= \left[\prod_{i=1}^{k}(1 - b_i x)\right]$$
$$\cdot |\mathbf{I} + x\mathbf{A} \cdot \operatorname{diag}(b_1(1 - b_1 x)^{-1}, \ldots, b_k(1 - b_k x)^{-1})|$$

for all x with $b_i x \neq 1$, $i = 1, \ldots, k$. Now, for any $c_1, \ldots, c_k \in \mathbb{R}^k$,

(2.13)
$$|\mathbf{I} + x\mathbf{A} \cdot \operatorname{diag}(c_1, \ldots, c_k)|$$
$$= |\mathbf{I} + x\mathbf{O} \operatorname{diag}(k, 0, \ldots, 0)\mathbf{O}' \cdot \operatorname{diag}(c_1, \ldots, c_k)|$$
$$= |\mathbf{I} + \operatorname{diag}(kx, 0, \ldots, 0) \cdot \mathbf{O}' \cdot \operatorname{diag}(c_1, \ldots, c_k)\mathbf{O}|.$$

Write $\mathbf{C} = \mathbf{O}' \cdot \operatorname{diag}(c_1, \ldots, c_k)\mathbf{O} = (c_{ij})$. Then, recall the first column of $\mathbf{O}$ is $\mathbf{h}$. Therefore,

$$c_{11} = \sum_{i=1}^{k}(\mathbf{O}')_{1i}\, c_i\, \mathbf{O}_{i\,1} = \frac{1}{k}\sum_{i=1}^{k} c_i.$$

Notice

$$\text{diag}(kx, 0, \ldots, 0) \cdot \mathbf{O}' \cdot \text{diag}(c_1, \ldots, c_k)\mathbf{O} = \begin{pmatrix} kxc_{11} & kxc_{12} & \cdots & kxc_{1k} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

From (2.13), we get $|\mathbf{I} + x\mathbf{A} \cdot \text{diag}(c_1, \ldots, c_k)| = 1 + kxc_{11} = 1 + x \sum_{i=1}^{k} c_i$. This together with (2.12) yields the conclusion. $\square$

LEMMA 2.5. *Let $b = b(x)$ be a function defined on $[0, \infty)$ with $b(x) = o(x)$ as $x \to \infty$. Then $\log \frac{\Gamma(x+b)}{x^b \Gamma(x)} = O(\frac{(|b|+1)^2}{x})$ as $x \to \infty$.*

PROOF. By Lemma 5.1 from Jiang and Qi (2015), as $x \to +\infty$,

$$\log \frac{\Gamma(x + b)}{\Gamma(x)} = (x + b) \log(x + b) - x \log x - b - \frac{b}{2x} + O\left(\frac{b^2 + 1}{x^2}\right)$$

holds uniformly on $b \in [-\delta x, \delta x]$ for any given $\delta \in (0, 1)$. By assumption, $b(x) = o(x)$, then

$$\begin{aligned}
&\log \frac{\Gamma(x + b)}{\Gamma(x)} \\
&= (x + b) \log(x + b) - x \log x - b + O\left(\frac{b^2 + bx + 1}{x^2}\right) \\
&= x \log\left(1 + \frac{b}{x}\right) + b \log x + b \log\left(1 + \frac{b}{x}\right) - b + O\left(\frac{b^2 + bx + 1}{x^2}\right) \\
&= x \left(\frac{b}{x} + O\left(\frac{b^2}{x^2}\right)\right) + b \log x + O\left(\frac{b^2}{x}\right) - b + O\left(\frac{b^2 + bx + 1}{x^2}\right) \\
&= b \log x + O\left(\frac{b^2 x + bx + b^2 + 1}{x^2}\right),
\end{aligned}$$

where the formula $\log(1 + t) = t + O(t^2)$ as $t \to 0$ is used in the third identity. Now

$$|b^2 x + bx + b^2 + 1| \le 2b^2 x + |b|x + 1 \le 2(|b| + 1)^2 x$$

as $x \ge 1$. The conclusion then follows. $\square$

We next provide upper and lower bounds for the standard deviation σ_n from Theorem 1.

LEMMA 2.6.   *Let $p := p_n$ satisfy that $m := n - 1 > p + 3$, $\mathbf{R}_n$ and σ_n be as in Theorem* 1. *Then*

$$(2.14) \qquad \left(\frac{p}{m}\right)^2 + \frac{2}{m} \operatorname{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right] \leq \sigma_n^2 \leq 2 \log m + \frac{2p^2}{m}$$

for all $n \geq 6$. Furthermore, $\lim_{n \to \infty} \frac{p}{\sigma_n} = \infty$.

PROOF.   By the Taylor expansion, $\log(1 - x) = -\sum_{i=1}^{\infty} \frac{x^i}{i}$ for all $|x| < 1$. Hence, $-[x + \log(1 - x)] \geq \frac{x^2}{2}$ for all $x \in [0, 1)$ and

$$\sigma_n^2 \geq \left(\frac{p}{m}\right)^2 + \frac{2}{m} \operatorname{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right].$$

Second, $\log(1 - \frac{p}{n-1}) = \log(\frac{n-1-p}{n-1}) \geq \log \frac{1}{n-1}$. Also, observe that $\operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2] = \sum_{1 \leq i \neq j \leq p} r_{ij}^2 \leq p^2$ since $|r_{ij}| \leq 1$. By the definition of σ_n^2, the second inequality follows.

Now we prove the remaining conclusion. It is enough to show $\frac{1}{p^2} \sigma_n^2 \to 0$. First,

$$\sigma_n^2 \leq -2 \log\left(1 - \frac{p}{m}\right) + \frac{2p^2}{m}.$$

Therefore,

$$\frac{1}{p^2} \sigma_n^2 \leq -\frac{2}{p^2} \log\left(1 - \frac{p}{m}\right) + \frac{2}{m}.$$

If $p \leq \frac{m}{2}$, then $1 - \frac{p}{m} \geq \frac{1}{2}$, and hence $-\log(1 - \frac{p}{m}) \leq \log 2$. This implies $\frac{1}{p^2} \sigma_n^2 \leq \frac{1}{p^2} \log 4 + \frac{2}{m}$. Moreover, if $p > \frac{m}{2}$, then

$$\frac{1}{p^2} \sigma_n^2 \leq -\frac{8}{m^2} \log \frac{m - p}{m} + \frac{2}{m}$$

$$\leq -\frac{8}{m^2} \log \frac{4}{m} + \frac{2}{m}.$$

Overall, $\frac{1}{p^2} \sigma_n^2 \leq \frac{1}{p^2} \log 4 + \frac{2}{m} + \frac{8}{m^2} \log m \to 0$. $\square$

2.3. *Moments on logarithms of determinants of sample correlation matrices.* In this section, we will prove (1.9) that is one of the crucial steps in proving Theorem 1. In the following, we write $\boldsymbol{\Sigma}$ for $\boldsymbol{\Sigma}_n$ for short notation. Review $W_p(m, \boldsymbol{\Sigma})$ as defined at the beginning of Section 2.1.

LEMMA 2.7.   *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. from $N_p(\mu, \boldsymbol{\Sigma})$ and $\hat{\mathbf{R}}_n$ be as in (1.3). Assume $n > p$ and $\boldsymbol{\Sigma}^{-1}$ exists. Let $\mathbf{R}_n$ be the correlation matrix generated by*

$\mathbf{\Sigma} = (\sigma_{ij})$ as in (1.1). Suppose $\bar{\mathbf{W}} = (\bar{W}_{ij})$ follows $W_p(m, \mathbf{I}_p)$ with $m = n - 1$. Set $\mathbf{W} = (W_{ij}) = \mathbf{\Sigma}^{1/2}\bar{\mathbf{W}}\mathbf{\Sigma}^{1/2}$. Then

$$|\hat{\mathbf{R}}_n| \stackrel{d}{=} |\hat{\mathbf{R}}_{n,0}| \cdot |\mathbf{R}_n| \cdot \left(\prod_{i=1}^{p} \bar{W}_{ii}\right) \cdot \prod_{i=1}^{p} \frac{\sigma_{ii}}{W_{ii}},$$

where $\hat{\mathbf{R}}_{n,0}$ is the $\hat{\mathbf{R}}_n$ corresponding to $\mathbf{\Sigma} = \mathbf{I}$.

PROOF. $\mathbf{W}$ follows the distribution of $W_p(m, \mathbf{\Sigma})$. By Lemma 2.1,

$$|\hat{\mathbf{R}}_n| \stackrel{d}{=} |\mathbf{W}| \cdot \prod_{i=1}^{p} \frac{1}{W_{ii}} = |\mathbf{\Sigma}| \cdot |\bar{\mathbf{W}}| \cdot \prod_{i=1}^{p} \frac{1}{W_{ii}}.$$

Define

$$|\hat{\mathbf{R}}_{n,0}| = |\bar{\mathbf{W}}| \cdot \prod_{i=1}^{p} \frac{1}{\bar{W}_{ii}},$$

which has the same distribution as that of $\hat{\mathbf{R}}_n$ as $\mathbf{\Sigma} = \mathbf{I}$. Then

(2.15) $$|\hat{\mathbf{R}}_n| \stackrel{d}{=} |\hat{\mathbf{R}}_{n,0}| \cdot \left(\prod_{i=1}^{p} \frac{\bar{W}_{ii}}{W_{ii}}\right) \cdot |\mathbf{\Sigma}|.$$

By the definition of $\mathbf{R}_n$,

$$|\mathbf{R}_n| = |\mathbf{\Sigma}| \cdot \prod_{i=1}^{p} \frac{1}{\sigma_{ii}}.$$

Replacing $|\mathbf{\Sigma}|$ in (2.15) with the one from the above leads to the desired conclusion. $\square$

LEMMA 2.8. *Assume the notation and conditions in Lemma 2.7 hold. Define* $T_n = (\prod_{i=1}^{p} \bar{W}_{ii}) \cdot \prod_{i=1}^{p} \frac{\sigma_{ii}}{W_{ii}}$. *Then, for any integer* $k \geq 1$,

$$\sup_{n \geq 3}\left\{\left[\frac{p^2}{m^2} + \frac{2}{m}\operatorname{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right]\right]^{-k} \cdot E\left[(\log T_n)^{2k}\right]\right\} < \infty.$$

PROOF. The proof is divided into several steps.
Step 1: *a workable form of* T_n. Write

$$\log T_n = \sum_{i=1}^{p} \log \frac{\bar{W}_{ii}}{m} - \sum_{i=1}^{p} \log \frac{W_{ii}}{m\sigma_{ii}}.$$

Define ε_i and $\varepsilon_{i,0}$ through

(2.16)
$$\log \frac{W_{ii}}{m\sigma_{ii}} = \frac{W_{ii}}{m\sigma_{ii}} - 1 + \left(\frac{W_{ii}}{m\sigma_{ii}} - 1\right)^2 \varepsilon_i;$$

$$\log \frac{\bar{W}_{ii}}{m} = \frac{\bar{W}_{ii}}{m} - 1 + \left(\frac{\bar{W}_{ii}}{m} - 1\right)^2 \varepsilon_{i,0}.$$

The estimates of ε_i and $\varepsilon_{i,0}$ will be given after (2.21). It follows that

$$\sum_{i=1}^{p} \log \frac{W_{ii}}{m\sigma_{ii}} = \sum_{i=1}^{p} \left(\frac{W_{ii}}{m\sigma_{ii}} - 1\right) + \sum_{i=1}^{p} \left(\frac{W_{ii}}{m\sigma_{ii}} - 1\right)^2 \varepsilon_i$$

and

$$\sum_{i=1}^{p} \log \frac{\bar{W}_{ii}}{m} = \sum_{i=1}^{p} \left(\frac{\bar{W}_{ii}}{m} - 1\right) + \sum_{i=1}^{p} \left(\frac{\bar{W}_{ii}}{m} - 1\right)^2 \varepsilon_{i,0}.$$

Consequently,

$$\sum_{i=1}^{p} \log \frac{\bar{W}_{ii}}{m} - \sum_{i=1}^{p} \log \frac{W_{ii}}{m\sigma_{ii}}$$

$$= \frac{1}{m}\big[\mathrm{tr}(\bar{\mathbf{W}}) - \mathrm{tr}(\mathbf{T}\mathbf{W})\big] - \sum_{i=1}^{p} \left(\frac{W_{ii}}{m\sigma_{ii}} - 1\right)^2 \varepsilon_i + \sum_{i=1}^{p} \left(\frac{\bar{W}_{ii}}{m} - 1\right)^2 \varepsilon_{i,0},$$

where $\mathbf{T} = \mathrm{diag}(1/\sigma_{ii}, \ 1 \le i \le p)$. From Lemma 2.7, it is readily seen that $\mathrm{tr}(\mathbf{T}\mathbf{W}) = \mathrm{tr}(\Sigma^{1/2}\mathbf{T}\Sigma^{1/2}\bar{\mathbf{W}})$. It gives us

$$\mathrm{tr}(\bar{\mathbf{W}}) - \mathrm{tr}(\mathbf{T}\mathbf{W}) = \mathrm{tr}(\mathbf{M}\bar{\mathbf{W}})$$

with $\mathbf{M} = \mathbf{I} - \Sigma^{1/2}\mathbf{T}\Sigma^{1/2}$. By the spectral theorem, there is an orthogonal matrix $\mathbf{O} = \mathbf{O}_{p \times p}$ such that $\mathbf{M} = \mathbf{O}' \mathrm{diag}(\lambda_1, \ldots, \lambda_p)\mathbf{O}$, where $\{\lambda_i; \ 1 \le i \le p\}$ are the eigenvalues of $\mathbf{M}$. Since the Wishart matrix $\bar{\mathbf{W}} = (\bar{W}_{ij})$ is orthogonally invariant,

$$\mathrm{tr}(\mathbf{M}\bar{\mathbf{W}}) = \mathrm{tr}\big[\mathrm{diag}(\lambda_1, \ldots, \lambda_p)\mathbf{O}\bar{\mathbf{W}}\mathbf{O}'\big]$$

$$\stackrel{d}{=} \mathrm{tr}\big[\mathrm{diag}(\lambda_1, \ldots, \lambda_p)\bar{\mathbf{W}}\big]$$

$$= \sum_{i=1}^{p} \lambda_i \bar{W}_{ii},$$

where $\{\bar{W}_{ii}; \ 1 \le i \le p\}$ are i.i.d. random variables with distribution $\chi^2(m)$. Also, observe $\sum_{i=1}^{p} \lambda_i = \mathrm{tr}(\mathbf{M}) = p - \mathrm{tr}(\mathbf{R}_n) = 0$ since $\mathrm{tr}(\Sigma^{1/2}\mathbf{T}\Sigma^{1/2}) = \mathrm{tr}(\Sigma\mathbf{T}) =$

$tr(\mathbf{T}^{1/2} \Sigma \mathbf{T}^{1/2})$. We conclude

$$\log T_n \stackrel{d}{=} \frac{1}{m} \sum_{i=1}^{p} \lambda_i (\bar{W}_{ii} - m) - \sum_{i=1}^{p} \left(\frac{W_{ii}}{m\sigma_{ii}} - 1 \right)^2 \varepsilon_i$$

(2.17)
$$+ \sum_{i=1}^{p} \left(\frac{\bar{W}_{ii}}{m} - 1 \right)^2 \varepsilon_{i,0}$$

$$= Z_1 - Z_2 + Z_3.$$

By the definition of correlation matrix from (1.1), we have $\mathbf{T}^{1/2} \Sigma \mathbf{T}^{1/2} = \mathbf{R}_n$. Since $\mathbf{M}_1 \mathbf{M}_2$ and $\mathbf{M}_2 \mathbf{M}_1$ have the same eigenvalues for any $p \times p$ matrices $\mathbf{M}_1$ and $\mathbf{M}_2$, it is easy to check that $\lambda_1, \ldots, \lambda_p$ are also the eigenvalues of $\mathbf{I} - \mathbf{R}_n$.

Step 2: *the moment of* Z_1. Review the definition of $W_p(m, \mathbf{I}_p)$ at the beginning of Section 2.1. We are able to write $\bar{W}_{ii} - m = \sum_{j=1}^{m} (\xi_{ij}^2 - 1)$ for all $1 \leq i \leq p$, where $\{\xi_{ij}; 1 \leq i \leq p, 1 \leq j \leq m\}$ are i.i.d. $N(0, 1)$-distributed random variables. By the Marcinkiewicz–Zygmund inequality [e.g., Theorem 2 on page 367 from Chow and Teicher (1988)], for any integer $k \geq 1$, there exists a constant $C_k > 0$ such that

$$E(Z_1^{2k}) \leq C_k m^{-2k} E\left[\left(\sum_{i=1}^{p} \sum_{j=1}^{m} \lambda_i^2 (\xi_{ij}^2 - 1)^2 \right)^k \right]$$

(2.18)
$$= C_k m^{-2k} (\lambda_1^2 + \cdots + \lambda_p^2)^k E\left[\left(E^N \sum_{j=1}^{m} (\xi_{Nj}^2 - 1)^2 \right)^k \right],$$

where N is a random variable independent of ξ_{ij}'s and

$$P(N = i) = \frac{\lambda_i^2}{\lambda_1^2 + \cdots + \lambda_p^2}, \qquad 1 \leq i \leq p.$$

In the sequel, the notation C_k, which denotes constants depending on k but not on n, may be different from line to line. By Hölder's inequality and the convex inequality on $g(x) := x^k$ on $[0, \infty)$,

$$\left(E^N \sum_{j=1}^{m} (\xi_{Nj}^2 - 1)^2 \right)^k \leq E^N \left[\left(\sum_{j=1}^{m} (\xi_{Nj}^2 - 1)^2 \right)^k \right] \leq m^{k-1} E^N \sum_{j=1}^{m} (\xi_{Nj}^2 - 1)^{2k}.$$

From Fubini's theorem,

$$E\left[\left(E^N \sum_{j=1}^{m} (\xi_{Nj}^2 - 1)^2 \right)^k \right] \leq m^{k-1} E^N \sum_{j=1}^{m} E[(\xi_{Nj}^2 - 1)^{2k}]$$

$$= m^k E[(\xi_{11}^2 - 1)^{2k}].$$

This and (2.18) imply

$$(2.19) \qquad E(Z_1^{2k}) \leq C_k m^{-k} (\lambda_1^2 + \cdots + \lambda_p^2)^k = C_k \cdot \left(\frac{1}{m} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2] \right)^k.$$

Step 3: *the moments of* Z_2 *and* Z_3. Recall Lemma 2.7. Let $\Sigma^{1/2} = (\tau_{ij}) = (\tau_{ij})_{p \times p}$. Then $(\sigma_{ij}) = \Sigma = (\tau_{ij})(\tau_{ij})'$. In particular, $\sigma_{ii} = \sum_{j=1}^p \tau_{ij}^2$. By definition $\bar{\mathbf{W}} = \mathbf{Y}'\mathbf{Y}$, where $\mathbf{Y}$ is a $m \times p$ matrix in which all entries are i.i.d. $N(0, 1)$-distributed random variables. Since $(W_{ij}) = \Sigma^{1/2} \bar{\mathbf{W}} \Sigma^{1/2} = \Sigma^{1/2} (\mathbf{Y}'\mathbf{Y}) \Sigma^{1/2}$, we see

$$\begin{aligned} W_{ii} &= (\tau_{i1}, \ldots, \tau_{ip}) \mathbf{Y}'\mathbf{Y} (\tau_{i1}, \ldots, \tau_{ip})' \\ &= \| \mathbf{Y}(\tau_{i1}, \ldots, \tau_{ip})' \|^2 \\ &\sim \sigma_{ii} \cdot \chi^2(m) \end{aligned}$$

by the invariance of i.i.d. normal random variables. We then get a simple formula

$$(2.20) \qquad \frac{W_{ii}}{\sigma_{ii}} \sim \chi^2(m)$$

for each $1 \leq i \leq p$. Let us calculate the moment of Z_2. Keep in mind that we will only use the marginal distribution as in (2.20). Then, by Hölder's inequality,

$$\begin{aligned} E(Z_2^{2k}) &\leq p^{2k-1} \sum_{i=1}^p E\left[\left(\frac{W_{ii}}{m\sigma_{ii}} - 1 \right)^{4k} \varepsilon_i^{2k} \right] \\ &\leq p^{2k-1} \sum_{i=1}^p \left\| \frac{W_{ii}}{m\sigma_{ii}} - 1 \right\|_{8k}^{4k} \cdot [E(\varepsilon_i^{4k})]^{1/2}, \end{aligned}$$

where $\|\xi\|_r := [E(|X|^r)]^{1/r}$ for any random variable ξ. From (2.20) and Corollary 2 on page 368 from Chow and Teicher (1988), which is an application of the Marcinkiewicz–Zygmund inequality used in *Step* 2,

$$E\left[\left(\frac{W_{ii}}{m\sigma_{ii}} - 1 \right)^{8k} \right] = \frac{1}{m^{8k}} E\left[\left(\sum_{i=1}^m (\eta_i^2 - 1) \right)^{8k} \right] \leq \frac{C_k}{m^{4k}},$$

where $\eta_1, \ldots, \eta_m$ are i.i.d. $N(0, 1)$-distributed random variables. Therefore,

$$(2.21) \qquad E(Z_2^{2k}) \leq C_k \cdot \frac{p^{2k-1}}{m^{2k}} \sum_{i=1}^p [E(\varepsilon_i^{4k})]^{1/2}.$$

Define $\psi(x) = \frac{\log x - x + 1}{(x-1)^2}$ for $x > 0$. It is easy to verify, there exists a constant $C > 0$ such that $|\psi(x)| \leq \frac{C}{x}$ for all $x > 0$. Take $x = \frac{W_{ii}}{m\sigma_{ii}}$ in (2.16) to see $|\varepsilon_i| \leq C (\frac{W_{ii}}{m\sigma_{ii}})^{-1}$. By (2.20),

$$E(\varepsilon_i^{4k}) \leq (Cm)^{4k} E[\chi^2(m)^{-4k}].$$

The density of $\chi^2(m)$ is $f(x) := \frac{1}{\Gamma(\frac{m}{2})2^{m/2}}x^{(m/2)-1}e^{-x/2}$ for $x > 0$. Consequently,

$$
\begin{aligned}
E[\chi^2(m)^{-4k}] &= \frac{1}{\Gamma(\frac{m}{2})2^{m/2}}\int_0^\infty x^{(m/2)-4k-1}e^{-x/2}\,dx \\
&= \frac{2^{(m/2)-4k}}{\Gamma(\frac{m}{2})2^{m/2}}\int_0^\infty v^{(m/2)-4k-1}e^{-v}\,dv \\
&= \frac{\Gamma(\frac{m}{2}-4k)}{\Gamma(\frac{m}{2})2^{4k}} \sim m^{-4k}
\end{aligned}
$$

(2.22)

by the transform $v = x/2$ and then Lemma 2.4 from Dong, Jiang and Li (2012) that $\lim_{x\to\infty}\frac{\Gamma(x+a)}{\Gamma(x)x^a} = 1$ for any number $a \in \mathbb{R}$, which can also be obtained from (2.5). The assertion (2.22) is consistent with the law of large numbers $\chi^2(m) \sim m$. We then have $E(\varepsilon_i^{4k}) \leq C_k$. By (2.21), we eventually arrive at

$$
E(Z_2^{2k}) \leq C_k\left(\frac{p}{m}\right)^{2k}. \tag{2.23}
$$

Inspecting the above proof, we get (2.23) from (2.20) without using any information on σ_{ii}'s. Applying (2.23) to the case $\Sigma = \mathbf{I}$, we get $E(Z_3^{2k}) \leq C_k(\frac{p}{m})^{2k}$.

Finally, the two inequalities on $E(Z_2^{2k})$ and $E(Z_3^{2k})$ together with (2.17) and (2.19) entail

$$
\begin{aligned}
E[(\log T_n)^{2k}] &\leq C_k\left[E(|Z_1|^{2k}) + E(|Z_2|^{2k}) + E(|Z_3|^{2k})\right] \\
&\leq C_k\left[\frac{p^2}{m^2} + \frac{2}{m}\operatorname{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right]\right]^k.
\end{aligned}
$$

The proof is complete. $\square$

We will need the following notation later:

$$
\begin{aligned}
\mu_{n,0} &= \left(p - n + \frac{3}{2}\right)\log\left(1 - \frac{p}{n-1}\right) - \frac{n-2}{n-1}p; \\
\sigma_{n,0}^2 &= -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right].
\end{aligned}
$$

(2.24)

LEMMA 2.9. *Assume $p := p_n$ satisfy that $n > p + 4$ and $p \to \infty$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. from $N_p(\mu, \mathbf{I})$ and $\hat{\mathbf{R}}_{n,0}$ be as in Lemma 2.7. Then there exists $s_0 > 0$ such that the following holds*:

(i) *For any subsequence $\{n_j; j \geq 1\}$ of positive integers with $\lim_{j\to\infty}\frac{p_{n_j}}{n_j} = y \in [0, 1]$,*

$$
\lim_{j\to\infty} E\exp\left(\frac{\log|\hat{\mathbf{R}}_{n_j,0}| - \mu_{n_j,0}}{\sigma_{n_j,0}}s\right) = e^{s^2/2}, \qquad |s| \leq s_0.
$$

(ii) *For all $|s| \le s_0$,*

$$\sup_{n \ge 6} E \exp\left(\frac{\log|\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}} s\right) < \infty.$$

PROOF.   Recall $\hat{\mathbf{R}}_{n,0}$ is the $\hat{\mathbf{R}}_n$ corresponding to $\Sigma = \mathbf{I}$.

By (1.6), $E \exp(\frac{\log|\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}} s) < \infty$ for all $n \ge 6$ and $|s| \le s_0$;

(i) The assertion (5.68) from Jiang and Yang (2013) confirms the case for $y \in (0, 1]$; the limit right above (5.6) from Jiang and Qi (2015) confirms the case for $y = 0$. So (ii) is true for any $y \in [0, 1]$.

(ii) If not, there exists $s' \ne 0$ and a subsequence $\{n_j; j \ge 1\}$ of positive integers such that $|s'| \le s_0$ and

$$\lim_{j \to \infty} E \exp\left(\frac{\log|\hat{\mathbf{R}}_{n_j,0}| - \mu_{n_j,0}}{\sigma_{n_j,0}} s'\right) = \infty.$$

Since $0 < \frac{p_n}{n} \le 1$ for all $n \ge 6$, there exists $y \in [0, 1]$ and a further subsequence $\{n_{j_k}; k \ge 1\}$ such that $\lim_{k \to \infty} \frac{p_{n_{j_k}}}{n_{j_k}} = y$ and

$$\lim_{k \to \infty} E \exp\left(\frac{\log|\hat{\mathbf{R}}_{n_{j_k},0}| - \mu_{n_{j_k},0}}{\sigma_{n_{j_k},0}} s'\right) = \infty,$$

which contradicts (i). The proof is completed. $\square$

We now are ready to prove the main conclusion in this section.

LEMMA 2.10.   *Assume $p := p_n$ satisfy that $n > p + 4$ and $p \to \infty$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. from $N_p(\mu, \Sigma)$ and $\hat{\mathbf{R}}_n$ be as in (1.3). Let μ_n and σ_n be as in Theorem 1. Let $\mathbf{R}_n$ be as in (1.1). Assume $\mathbf{R}_n^{-1}$ exists. Then, for any $k \ge 1$,*

$$\sup_{n \ge 6} E\left[\left(\frac{\log|\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n}\right)^{2k}\right] < \infty.$$

PROOF.   Review (2.24) for $\mu_{n,0}$ and $\sigma_{n,0}^2$. Define $m = n - 1$. Then

$$\mu_n = \mu_{n,0} + \log|\mathbf{R}_n| \quad \text{and} \quad \sigma_n^2 = \sigma_{n,0}^2 + \frac{2}{m} \text{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right].$$

By Lemma 2.7,

$$\frac{\log|\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n} = \frac{\log|\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}} \cdot \frac{\sigma_{n,0}}{\sigma_n} + \frac{1}{\sigma_n} \log T_n,$$

where $T_n = [(\prod_{i=1}^{p} \bar{W}_{ii}) \cdot \prod_{i=1}^{p} \frac{\sigma_{ii}}{W_{ii}}]$ is as in Lemma 2.8. Therefore,

(2.25)
$$2^{1-2k} E\left[\left(\frac{\log |\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n}\right)^{2k}\right]$$
$$\leq E\left[\left(\frac{\log |\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}}\right)^{2k}\right] + \frac{1}{\sigma_n^{2k}} \cdot E[(\log T_n)^{2k}]$$

for any $n \geq 6$. By (iii) of Lemma 2.9,

$$\sup_{n \geq 6} E \exp\left(\frac{\log |\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}} s\right) < \infty$$

for all $|s| \leq s_0$. Use the inequality $\frac{x^{2k}}{(2k)!} \leq e^{|x|} \leq e^x + e^{-x}$ for any $x \in \mathbb{R}$ to get

(2.26)
$$\sup_{n \geq 6} E\left[\left(\frac{\log |\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}}\right)^{2k}\right] < \infty$$

for any integer $k \geq 1$.

On the other hand, by Lemma 2.6,

$$\sigma_n^2 \geq \left(\frac{p}{m}\right)^2 + \frac{2}{m} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2].$$

Thus, from Lemma 2.8 we see

$$\sup_{n \geq 6}\left\{\frac{1}{\sigma_n^{2k}} \cdot E[(\log T_n)^{2k}]\right\} < \infty$$

for any integer $k \geq 1$. This, (2.25) and (2.26) entail the desired result. $\square$

2.4. *Convergence of random determinants.* We now prove (1.11), the convergence in probability, a key step to compute I_n in (1.10). The major conclusion is Lemma 2.12.

Throughout the rest of the paper, we always assume the following:

(2.27)
Let σ_n be as in Theorem 1. Given $s > 0$, set $t = t_n = \dfrac{s}{\sigma_n}$ and $m = n - 1$. Let $V_1, \ldots, V_p$ be i.i.d. random variables with the distribution $\operatorname{Beta}\left(t, \dfrac{m}{2}\right)$.

Recall $\mathbf{R}_n$ and $\mathbf{\Delta}_n$ from Proposition 1.1. Let $\mathbf{D}_n = \operatorname{diag}(\sqrt{V_1}, \ldots, \sqrt{V_p})$ and $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of $\mathbf{D}_n \mathbf{\Delta}_n \mathbf{D}_n$. Observing $(\mathbf{\Delta}_n)_{ii} = 0$, then

$(\mathbf{D}_n \boldsymbol{\Delta}_n \mathbf{D}_n)_{ii} = 0$ and $(\mathbf{D}_n \boldsymbol{\Delta}_n \mathbf{D}_n)_{ij} = r_{ij} \sqrt{V_i V_j}$ for all $i \neq j$. These say

$$
\sum_{i=1}^{p} \lambda_i = \mathrm{tr}(\mathbf{D}_n \boldsymbol{\Delta}_n \mathbf{D}_n) = 0 \quad \text{and}
$$

(2.28)

$$
\sum_{i=1}^{p} \lambda_i^2 = \mathrm{tr}\big[(\mathbf{D}_n \boldsymbol{\Delta}_n \mathbf{D}_n)^2\big] = \sum_{i \neq j} r_{ij}^2 V_i V_j.
$$

LEMMA 2.11. *Let $p := p_n$ satisfy that $n > p$ and $p \to \infty$ and σ_n be as in Theorem* 1. *Given $s > 0$, set $t = t_n = \frac{s}{\sigma_n}$ and $m = n - 1$. Then $m \sum_{i=1}^{p} \lambda_i^2 = \frac{4t^2}{m} \cdot \mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2] + \varepsilon_n$ with $\varepsilon_n \to 0$ in probability as $n \to \infty$.*

PROOF.   First, if $\xi \sim \mathrm{Beta}(\alpha, \beta)$, then

$$
E\xi = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \mathrm{Var}(\xi) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)};
$$

(2.29)

$$
E(\xi^k) = \prod_{i=0}^{k-1} \frac{\alpha + i}{\alpha + \beta + i}, \qquad k = 1, 2, \dots.
$$

In our case, $V_1 \sim \mathrm{Beta}(t, \frac{m}{2})$. From (2.14), $\frac{t}{m} \to 0$, hence

$$
(2.30) \quad EV_1 = \frac{t}{\frac{m}{2} + t} \sim \frac{2t}{m}; \qquad E(V_1^2) = \frac{t(t+1)}{(\frac{m}{2} + t)(\frac{m}{2} + t + 1)} \sim \frac{4t(t+1)}{m^2};
$$

$$
(2.31) \ \mathrm{Var}(V_1) = \frac{mt/2}{(\frac{m}{2} + t)^2(\frac{m}{2} + t + 1)} \sim \frac{4t}{m^2}
$$

as $n \to \infty$. Let $U_n = \sum_{i \neq j} r_{ij}^2 V_i V_j$. Then

$$
EU_n = (EV_1)^2 \sum_{i \neq j} r_{ij}^2
$$

(2.32)

$$
= (EV_1)^2 \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]
$$

$$
= \big(1 + o(1)\big) \cdot \frac{4t^2}{m^2} \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big].
$$

Write $\bar{V}_i = V_i - EV_i$ for $i = 1, \dots, p$. Then

$$
U_n - EU_n = \sum_{i \neq j} r_{ij}^2 \bar{V}_i \bar{V}_j + 2 \sum_{i \neq j} r_{ij}^2 \bar{V}_i \cdot EV_j
$$

(2.33)

$$
= \sum_{i \neq j} r_{ij}^2 \bar{V}_i \bar{V}_j + 2(EV_1) \sum_{i=1}^{p} \Big(\sum_{j \neq i} r_{ij}^2\Big) \bar{V}_i.
$$

By independence, the two terms in (2.33) are uncorrelated and the summands from each sum are also uncorrelated, respectively. Therefore,

$$\mathrm{Var}\,(U_n)$$

$$= E\big[(U_n - EU_n)^2\big]$$

(2.34)
$$= \mathrm{Var}\,(V_1)^2 \cdot \sum_{i \neq j} r_{ij}^4 + 4(EV_1)^2 \cdot \mathrm{Var}\,(V_1) \cdot \sum_{i=1}^{p}\bigg(\sum_{j \neq i} r_{ij}^2\bigg)^2$$

$$\leq \mathrm{Var}\,(V_1)^2 \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big] + 4\|\mathbf{R}_n\| \cdot (EV_1)^2 \cdot \mathrm{Var}\,(V_1) \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]$$

since $\sum_{i \neq j} r_{ij}^4 \leq \sum_{i \neq j} r_{ij}^2 = \mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2]$ and

$$\sum_{i=1}^{p}\bigg(\sum_{j \neq i} r_{ij}^2\bigg)^2 \leq \|\mathbf{R}_n\| \cdot \sum_{j \neq i} r_{ij}^2$$

by Lemma 2.2. Then, from (2.14), (2.31) and the notation $t = t_n = \frac{s}{\sigma_n}$,

$$m^2\,\mathrm{Var}\,(V_1)^2 \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]$$

(2.35)
$$\sim \frac{16s^2}{m^2\sigma_n^2} \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]$$

$$\leq \big(16s^2\big) \cdot \frac{\mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2]}{p^2 + m\,\mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2]} \leq \frac{16s^2}{m} \to 0$$

as $n \to \infty$. Now,

$$m^2 \cdot \|\mathbf{R}_n\| \cdot (EV_1)^2 \cdot \mathrm{Var}\,(V_1) \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]$$

$$\leq 2 \cdot m^2 \cdot \frac{4t^2}{m^2} \cdot \frac{4t}{m^2} \cdot \|\mathbf{R}_n\| \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]$$

$$= 32 \cdot \frac{\|\mathbf{R}_n\|t}{m} \cdot \bigg(t^2 \cdot \frac{1}{m}\,\mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big]\bigg)$$

as n is sufficiently large. By (2.14), $t^2 \cdot \frac{1}{m}\,\mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2] \leq s^2$. Recall the notation $\|\mathbf{R}_n\| = \lambda_{\max}(\mathbf{R}_n)$. Trivially, $\mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2] \geq (\|\mathbf{R}_n\| - 1)^2$. Then, from (2.14) again,

$$\frac{\|\mathbf{R}_n\|t}{m} \leq s \cdot \frac{|\|\mathbf{R}_n\| - 1| + 1}{m} \cdot \bigg(\frac{p^2}{m^2} + \frac{1}{m}\big(\|\mathbf{R}_n\| - 1\big)^2\bigg)^{-1/2}$$

$$= s \cdot \frac{|\|\mathbf{R}_n\| - 1| + 1}{\sqrt{p^2 + m \cdot (\|\mathbf{R}_n\| - 1)^2}}.$$

Evidently, the fraction is controlled by $\frac{1}{\sqrt{m}} + \frac{1}{p}$. Therefore,

$$m^2 \cdot \|\mathbf{R}_n\| \cdot (EV_1)^2 \cdot \mathrm{Var}\,(V_1) \cdot \mathrm{tr}\big[(\mathbf{R}_n - \mathbf{I})^2\big] \to 0$$

as $n \to \infty$. Connecting this to (2.34) and (2.35), we obtain $\text{Var}(mU_n) \to 0$. Therefore, $mU_n - mEU_n \to 0$ in probability. The inequality in (2.14) implies $\frac{4t^2}{m} \cdot \text{tr}[(\mathbf{R}_n - \mathbf{I})^2]$ is bounded. Hence, we see from (2.28) and (2.32) that $m\sum_{i=1}^{p} \lambda_i^2 = mU_n = \frac{4t^2}{m} \cdot \text{tr}[(\mathbf{R}_n - \mathbf{I})^2] + \varepsilon_n$ with $\varepsilon_n \to 0$ in probability as $n \to \infty$. $\square$

Recall the notation in (2.27). Here comes our main result for convergence in probability.

LEMMA 2.12. *Let $p := p_n$ satisfy that $n > p$ and $p \to \infty$ and σ_n be as in Theorem 1. Review $\mathbf{R}_n$ defined in (1.1). Set $\mathbf{\Delta}_n = \mathbf{R}_n - \mathbf{I}$. Then*

$$-\left(\frac{m}{2} + t\right) \log|\mathbf{I} + \mathbf{\Delta}_n \cdot \text{diag}(V_1, \ldots, V_p)| - \frac{t^2}{m} \cdot \text{tr}(\Delta_n^2) \to 0$$

in probability as $n \to \infty$.

PROOF.   Let $\mathbf{D}_n = \text{diag}(\sqrt{V_1}, \ldots, \sqrt{V_p})$ and $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of $\mathbf{D}_n \mathbf{\Delta}_n \mathbf{D}_n$. Then

$$\log|\mathbf{I} + \mathbf{\Delta}_n \cdot \text{diag}(V_1, \ldots, V_p)| = \log|\mathbf{I} + \mathbf{D}_n \mathbf{\Delta}_n \mathbf{D}_n|$$

(2.36)
$$= \sum_{i=1}^{p} \log(1 + \lambda_i).$$

We first need a few of estimates. It is easy to check from (2.14) and notation $t = \frac{s}{\sigma_n}$ that

(2.37)
$$\frac{|t|^3}{m^2} \cdot \text{tr}(\mathbf{\Delta}_n^2) = \frac{|s|^3}{m\sigma_n} \cdot \frac{1}{m\sigma_n^2} \text{tr}(\mathbf{\Delta}_n^2) \leq \frac{|s|^3}{p} \to 0$$

by using $m\sigma_n \geq p$ and $\sigma_n^2 \geq \frac{2}{m} \text{tr}(\mathbf{\Delta}_n^2)$. Similarly,

(2.38)
$$\frac{4t^2}{m^3} \cdot \text{tr}(\mathbf{\Delta}_n^2) = \frac{4s^2}{m^2} \cdot \frac{1}{m\sigma_n^2} \text{tr}(\mathbf{\Delta}_n^2) \leq \frac{4s^2}{m^2}.$$

From Lemma 2.11 and (2.37), for any sequence $\{\delta_n; n \geq 4\}$ with $\delta_n = o(1)$,

$$-\left(\frac{m}{2} + t\right) \cdot \left(-\frac{1}{2} + \delta_n\right) \cdot \sum_{i=1}^{p} \lambda_i^2$$

(2.39)
$$= (1 - 2\delta_n) \cdot \frac{1}{4} \cdot \left[\frac{4t^2}{m} \cdot \text{tr}(\mathbf{\Delta}_n^2) + \varepsilon_n\right]$$

$$+ \frac{1 - 2\delta_n}{2} \cdot \frac{t}{m} \cdot \left[\frac{4t^2}{m} \cdot \text{tr}(\mathbf{\Delta}_n^2) + \varepsilon_n\right]$$

$$= \frac{t^2}{m} \cdot \text{tr}(\mathbf{\Delta}_n^2) + \varepsilon_n',$$

where $\varepsilon_n \to 0$ and $\varepsilon'_n \to 0$ in probability as $n \to \infty$, and the assertion $\frac{t}{m}\varepsilon_n \to 0$ in probability is due to the fact $\frac{t}{m} \to 0$; the fact $\frac{4t^2}{m} \cdot \text{tr}[(\mathbf{R}_n - \mathbf{I})^2]$ is bounded comes from (2.14).

By Lemma 2.11 again, $m\sum_{i=1}^{p}\lambda_i^2 = \frac{4t^2}{m} \cdot \text{tr}(\mathbf{\Delta}_n^2) + \varepsilon_n$. Thus, we see from (2.38) that $P(\sum_{i=1}^{p}\lambda_i^2 \leq \frac{5s^2}{m}) \to 1$ as $n \to \infty$. Over the set $\{\sum_{i=1}^{p}\lambda_i^2 \leq \frac{5s^2}{m}\}$, we know $\max_{1 \leq i \leq p}|\lambda_i| \leq \frac{3s}{\sqrt{m}}$. In particular, $\max_{1 \leq i \leq p}|\lambda_i| \leq \frac{1}{2}$ as n is large enough.

Write $\log(1 + x) = x - \frac{1}{2}x^2 + x^3 \cdot \varepsilon(x)$ for $|x| < 1$. By Taylor's expansion, $\varepsilon(x) = \sum_{i=0}^{\infty}(-1)^i\frac{1}{i+3}x^i$. Thus, $|\varepsilon(x)| \leq \frac{1}{3}\sum_{i=0}^{\infty}|x|^i = \frac{1}{3(1-|x|)}$ for all $|x| < 1$. So $\sup_{|x| \leq 1/2}|\varepsilon(x)| \leq 1$. Define $\Omega_n = \{\sum_{i=1}^{p}\lambda_i^2 \leq \frac{5s^2}{m}\}$, which is a subset of the whole sample space Ω. From now on, we assume n is sufficiently large. On Ω_n, it is seen from the first identity in (2.28) that

$$\sum_{i=1}^{p}\log(1 + \lambda_i) = \sum_{i=1}^{p}\lambda_i - \frac{1}{2}\sum_{i=1}^{p}\lambda_i^2 + \sum_{i=1}^{p}\lambda_i^3 \cdot \varepsilon_i(\lambda_i)$$

$$= -\frac{1}{2}\sum_{i=1}^{p}\lambda_i^2 + \sum_{i=1}^{p}\lambda_i^3 \cdot \varepsilon_i(\lambda_i)$$

by (2.28) where $\varepsilon_i(x)$, $1 \leq i \leq p$, are functions satisfying $\sup_{1 \leq i \leq p}|\varepsilon(\lambda_i)| \leq 1$. Trivially,

$$\left|\sum_{i=1}^{p}\lambda_i^3 \cdot \varepsilon_i(\lambda_i)\right| \leq \max_{1 \leq i \leq p}\left|\lambda_i\right| \cdot \sum_{i=1}^{p}\lambda_i^2 \leq \frac{3s}{\sqrt{m}} \cdot \sum_{i=1}^{p}\lambda_i^2$$

on Ω_n. Hence,

$$-\left(\frac{m}{2} + t\right)\sum_{i=1}^{p}\log(1 + \lambda_i) = -\left(\frac{m}{2} + t\right) \cdot \left[-\frac{1}{2} + O\left(\frac{1}{\sqrt{m}}\right)\right] \cdot \sum_{i=1}^{p}\lambda_i^2$$

on Ω_n. The lemma then follows from (2.36), (2.39) and the fact $P(\Omega_n) \to 1$ as $n \to \infty$. $\square$

2.5. *Uniform integrability of random determinants*. In this section, we will show (1.12) en route to prove Theorem 1. Recall the notation in (2.27). First, we need some concentration inequalities.

LEMMA 2.13. *Let $p := p_n$ satisfy that $m := n - 1 > p \to \infty$ and σ_n be as in Theorem 1. Then, for any $\rho \in (0, \frac{1}{2})$, there exist constants $M = M(\rho) > 0$ and $n_0 = n_0(\rho, s) \geq 1$ such that*

$$\sup_{y \geq M\frac{pt}{m}}\left\{e^{\rho my} \cdot P\left(\sum_{i=1}^{p}V_i > y\right)\right\} \leq 1$$

as $n \geq n_0$.

Lemma 2.13 is not a standard Chernoff bound since the mean of V_1 goes to zero. An extra effort has to be paid to get the subtle rate. In fact, the rate $e^{-C_1 m}$ for some $C_1 > 0$ in the lemma is different from $e^{-C_2 p}$ for some constants $C_2 > 0$, which is usually seen in a Chernoff bound.

PROOF OF LEMMA 2.13.   Notice $\sum_{i=1}^{p} V_i \leq p$. So without loss of generality, we assume $y \leq p$. Set $r = \frac{m}{2}$. By the Chernoff bound [see, e.g., Dembo and Zeitouni (1998)],

$$(2.40) \qquad P\left(\frac{1}{p}\sum_{i=1}^{p} V_i \geq x\right) \leq e^{-pI(x)}$$

for all $x > EV_1 \sim \frac{2t}{m}$ from (2.30), where

$$I(x) = \sup_{\theta > 0}\{\theta x - \log E e^{\theta V_1}\}.$$

Step 1. *Estimate of the moment generating function* $E e^{\theta V_1}$. Notice

$$E e^{\theta V_1} = \frac{1}{B(t, r)}\int_0^1 x^{t-1}(1-x)^{r-1}e^{\theta x}\,dx.$$

Define $w = r - 1$. Then $(1-x)^{r-1} \leq e^{-wx}$ since $r - 1 = \frac{m}{2} - 1 \geq 0$ as $n \geq 3$. Hence,

$$\begin{aligned} E e^{\theta V_1} &\leq \frac{1}{B(t, r)}\int_0^1 x^{t-1}e^{-(w-\theta)x}\,dx \\ &\leq \frac{(w-\theta)^{-t}}{B(t, r)}\int_0^\infty y^{t-1}e^{-y}\,dy \\ &= \frac{\Gamma(r+t)}{\Gamma(r)}(w-\theta)^{-t} \end{aligned}$$

for $\theta < w$, where the transform $y = (w - \theta)x$ is used in the second step and the formula $B(t, r) = \frac{\Gamma(r)\Gamma(t)}{\Gamma(r+t)}$ is applied in the last identity. By using $t = \frac{s}{\sigma_n} = O(\frac{m}{p})$ and Lemma 2.5, we see $\frac{\Gamma(r+t)}{\Gamma(r)} = r^t u$ where $u = u_n = \exp[O((t^2 + 1)m^{-1})]$ not depending on θ. Then

$$(2.41) \qquad E e^{\theta V_1} \leq r^t u \cdot (w-\theta)^{-t}$$

uniformly for all $\theta < w$ as n is sufficiently large.

Step 2. *Evaluation of the rate function* $I(x)$. From (2.41),

$$I(x) \geq -\log(r^t u) + \sup_{0 < \theta < w}\{\theta x + t\log(w - \theta)\}.$$

Set $\varphi(\theta) = \theta x + t \log(w - \theta)$ for $0 < \theta < w$. Then $\varphi'(\theta) = x - \frac{t}{w-\theta}$ and $\varphi''(\theta) = -t(w-\theta)^{-2} < 0$. So the maximizer θ_0 satisfies $w - \theta_0 = \frac{t}{x}$ or $\theta_0 = w - \frac{t}{x} \in (0, w)$ if $x > \frac{t}{w}$. Thus, under the restriction

$$(2.42) \qquad x > \frac{t}{\frac{m}{2} - 1} \sim \frac{2t}{m},$$

we obtain $I(x) \geq -\log(r^t u) + wx - t + t \log \frac{t}{x}$. By (2.40),

$$P\left(\frac{1}{p}\sum_{i=1}^{p} V_i \geq x\right) \leq (r^t u)^p e^{-pJ(x)}$$

for all x satisfying (2.42), where $J(x) = wx - t + t \log \frac{t}{x}$. Therefore,

$$P\left(\sum_{i=1}^{p} V_i > y\right) = P\left(\frac{1}{p}\sum_{i=1}^{p} V_i > x\right) \leq (r^t u)^p e^{-pJ(x)}$$

with $x = \frac{y}{p}$ if (2.42) holds, which is ensured if $y \geq \frac{3pt}{m}$. Now

$$J(x) = \frac{wy}{p} - t + t \log \frac{pt}{y}.$$

Consequently,

$$P\left(\sum_{i=1}^{p} V_i > y\right) \leq (r^t u)^p \cdot \exp\left\{-wy + pt - pt \log \frac{pt}{y}\right\}$$

$$(2.43) \qquad = \exp\left\{-wy + pt - pt \log \frac{pt}{y} + pt \log r + p \log u\right\}$$

$$= \exp\left\{-wy + pt \log \frac{ery}{pt} + p \log u\right\}$$

for all $y \geq \frac{3pt}{m}$.

Step 3. In this part, we will show that the second and third terms in $\{\cdot\}$ from (2.43) are small relative to the first term wy. Review $u = \exp[O((t^2 + 1)m^{-1})]$. It is easy to see, as n is sufficiently large,

$$\frac{p \log u}{wy} = O\left(\frac{p(t^2 + 1)}{m^2 y}\right) = O\left(\frac{t^2 + 1}{mt}\right)$$

uniformly for all $y \geq 3\frac{pt}{m}$. From Lemma 2.6,

$$(2.44) \qquad \frac{t^2 + 1}{mt} = \frac{1}{m}\left(t + \frac{1}{t}\right) = \frac{s}{m\sigma_n} + \frac{\sigma_n}{ms}$$

$$\leq \frac{s}{p} + \frac{1}{s}\left(\frac{2 \log m}{m^2} + \frac{2p^2}{m^3}\right)^{1/2} \to 0$$

as $n \to \infty$ since $p < m$ by assumption. This implies

$$(2.45) \qquad \tau_n := \sup_{y \geq 3pt/m} \frac{p \log u}{wy} \to 0$$

as $n \to \infty$. Moreover, noting $\frac{r}{r-1} = \frac{m}{m-2} \leq 2$ for all $m = n - 1 \geq 4$, we have

$$\left(pt \log \frac{ery}{pt} \right) \cdot \frac{1}{wy} = \frac{pt}{wy} \log \frac{ery}{pt}$$

$$= \frac{er}{r-1} \frac{\log A}{A} \leq \frac{2e \log A}{A}$$

for all $n \geq 5$, where $A := \frac{ery}{pt}$. Evidently, given $M \geq 3$, we see $A \geq \frac{m}{pt} y \geq M$ if $y \geq M \frac{pt}{m}$. Realizing $h(x) = \frac{\log x}{x}$ is decreasing on $[e, \infty)$, we see

$$\sup_{y > Mpt/m} \left\{ \left(pt \log \frac{ery}{pt} \right) \cdot \frac{1}{wy} \right\} \leq \frac{6 \log M}{M}$$

for all $M \geq 3$ and $n \geq 5$. This, (2.43) and (2.45) imply

$$(2.46) \qquad P\left(\sum_{i=1}^{p} V_i > y \right) \leq \exp\left\{ -wy \left(1 - \frac{6 \log M}{M} - \tau_n \right) \right\}$$

uniformly for all $y > M \frac{pt}{m}$, $M \geq 3$ and large n not depending on M. Trivially, $wy \sim \frac{my}{2}$ as $n \to \infty$. From the fact $\lim_{M \to \infty} \frac{\log M}{M} = 0$ and $\lim_{n \to \infty} \tau_n = 0$, we get the desired conclusion by picking $M = M(\rho)$ through (2.46) and $n_0 = n_0(\rho, s) \geq 1$ through (2.44) and (2.45). $\quad \square$

Lemma 2.13 is a type of large deviations. We also need to estimate a similar probability when the "y" in the lemma is small. This belongs to the zone of moderate deviations. We will apply the following inequality by Yurinskiĭ (1976) to achieve this purpose.

LEMMA 2.14. *Let $\xi_1, \ldots, \xi_p$ be independent random variables taking values in a separable Banach space $(\mathbb{B}, \|\cdot\|)$ satisfying*

$$E\left(\|\xi_j\|^m \right) \leq \frac{m!}{2} b_j^2 H^{m-2}, \qquad m = 2, 3, \ldots.$$

Let $\beta_p \geq E\|\xi_1 + \cdots + \xi_p\|$ and $B_p^2 = b_1^2 + \cdots + b_p^2$. For any $x > \frac{\beta_p}{B_p}$, set $\bar{x} = x - \frac{\beta_p}{B_p}$. Then

$$P\left(\|\xi_1 + \cdots + \xi_p\| \geq x B_p \right) \leq \exp\left(-\frac{\bar{x}^2}{8} \frac{1}{1 + (\bar{x} H)/(2 B_p)} \right).$$

Specializing the above lemma to our set-up, we get the following result.

LEMMA 2.15. *Let $p := p_n$ satisfy that $m := n - 1 > p + 3$ and $p \to \infty$ and σ_n be as in Theorem 1. Assume either $\inf_{n \geq 6} \frac{p_n}{n} > 0$ or $\inf_{n \geq 6} \frac{1}{n} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2] > 0$. Then, for some $\delta > 0$,*

$$P\left(\sum_{i \neq j} r_{ij}^2 V_i V_j \geq y\right) \leq \exp\left(-\frac{1}{256} \cdot \frac{m^2 y}{pt + m\sqrt{y}}\right)$$

for all $y > \frac{1}{m}$, $s \in (0, \delta]$ and $n \geq 6$.

The Hanson–Wright inequality [Hanson and Wright (1971)] can also provide an upper bound for the probability in Lemma 2.15. A most recent exposition of the inequality is given by Rudelson and Vershynin (2013). However, since all V_i's take small values, the Hanson–Wright inequality seems not able to catch the precision we want. The Yurinskii inequality supplies an ideal upper bound for us.

PROOF OF LEMMA 2.15. By Lemma 2.6, there exists a constant $s_0 > 0$ such that

$$\sigma_n^2 \geq \left(\frac{p}{m}\right)^2 + \frac{2}{m} \operatorname{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right] \geq s_0^2 > 0$$

for all $n \geq 6$ provided $\inf_{n \geq 6} \frac{p_n}{n} > 0$ or $\inf_{n \geq 6} \frac{1}{n} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2] > 0$. This says

(2.47) $$t = t_n = \frac{s}{\sigma_n} \leq 1$$

for all $n \geq 6$ and $0 < s \leq s_0$.

In the application of Lemma 2.14, next we take $\mathbb{B} = \mathbb{R}^p$ and $\|\cdot\|$ is the Euclidean norm.

Now, for $\mathbf{R}_n = (r_{ij})$ is nonnegative definite, the Hadamard product $\mathbf{R}_n \circ \mathbf{R}_n = (r_{ij}^2)$ is also nonnegative definite by the Schur product theorem; see, for example, page 458 from Horn and Johnson (1985). So we can write $(r_{ij}^2) = \mathbf{A}^2 = \mathbf{A}'\mathbf{A}$, where $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_p)$ is a nonnegative definite matrix. In particular, $r_{ij}^2 = \mathbf{a}_i' \mathbf{a}_j$ for $i \neq j$ and $\|\mathbf{a}_i\| = 1$. Set $W_n = \sum_{i \neq j} r_{ij}^2 V_i V_j$. Then

(2.48) $$W_n \leq \sum_{1 \leq i, j \leq p} r_{ij}^2 V_i V_j = \sum_{1 \leq i, j \leq p} \mathbf{a}_i' \mathbf{a}_j V_i V_j = \|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\|^2.$$

Define $r = \frac{m}{2}$ as before. Note $E\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\| \leq (E\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\|^2)^{1/2}$ and

$$E\left(\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\|^2\right) = E \sum_{1 \leq i, j \leq p} r_{ij}^2 V_i V_j$$

$$= \operatorname{tr}\left[(\mathbf{R}_n - \mathbf{I})^2\right] \cdot \frac{t^2}{(r+t)^2} + p \frac{t(t+1)}{(r+t)(r+t+1)}$$

$$\leq \frac{2s^2}{m} + \frac{8pt}{m^2},$$

where (2.30) is used in the second identity and the fact $\frac{2}{m}\,\mathrm{tr}[(\mathbf{R}_n - \mathbf{I})^2] \leq \sigma_n^2$ and (2.47) are employed in the last step. From Lemma 2.6, $\sigma_n \geq \frac{p}{m}$, it follows

$$E\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\|$$

(2.49)
$$\leq \left(\frac{2s^2}{m} + \frac{8s}{m}\right)^{1/2}$$

$$\leq \frac{1}{2\sqrt{m}} := \beta_p$$

for all $0 < s \leq 1/40$. Now let us bound the probability $P(\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\| \geq x)$. First, by (2.29),

$$E\big[\|\mathbf{a}_1 V_1\|^m\big] = E(V_1^m)$$

$$= \frac{t(t+1)\cdots(t+m-1)}{(r+t)(r+t+1)\cdots(r+t+m-1)}$$

$$\leq \frac{t(t+1)}{r^2} \cdot \frac{(t+2)(t+3)\cdots(t+m-1)}{r^{m-2}}$$

$$\leq \frac{2t}{r^2} \cdot \frac{\frac{1}{2}m!}{r^{m-2}}.$$

According to the notation from Lemma 2.14,

(2.50)
$$b_j^2 = \frac{2t}{r^2}, \qquad B_p^2 = \frac{2pt}{r^2}, \qquad H = \frac{1}{r}.$$

If $x \geq 2\frac{\beta_p}{B_p}$, then $\frac{x}{2} \leq \bar{x} := x - \frac{\beta_p}{B_p} \leq x$. Therefore, we have from Lemma 2.14 and (2.49) that

$$P\big(\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\| \geq x B_p\big) \leq \exp\left(-\frac{\bar{x}^2}{8} \frac{1}{1 + (\bar{x}H)/(2B_p)}\right)$$

$$\leq \exp\left(-\frac{x^2}{32} \frac{1}{1 + (xH)/(2B_p)}\right)$$

for all $x \geq 2\frac{\beta_p}{B_p}$ and $0 < s \leq \min\{s_0, \frac{1}{40}\}$. Now, from (2.48),

$$P(W_n \geq y) \leq P\big(\|\mathbf{a}_1 V_1 + \cdots + \mathbf{a}_p V_p\| \geq x B_p\big)$$

with $x = \frac{\sqrt{y}}{B_p}$. Hence, if $x = \frac{\sqrt{y}}{B_p} \geq 2\frac{\beta_p}{B_p}$, that is, $y \geq 4\beta_p^2 = \frac{1}{m}$, we have

$$P(W_n \geq y) \leq \exp\left(-\frac{y}{32B_p^2} \frac{1}{1 + (H\sqrt{y})/(2B_p^2)}\right).$$

Finally, from (2.50),

$$\frac{y}{32B_p^2} \frac{1}{1 + (H\sqrt{y})/(2B_p^2)} = \frac{1}{32} \cdot \frac{r^2 y}{2pt + r\sqrt{y}/2} \geq \frac{1}{256} \cdot \frac{m^2 y}{pt + m\sqrt{y}}.$$

In summary,

$$P(W_n \geq y) \leq \exp\left(-\frac{1}{256} \cdot \frac{m^2 y}{pt + m\sqrt{y}}\right)$$

for all $y \geq \frac{1}{m}$, $0 < s \leq \min\{s_0, \frac{1}{40}\}$ and $n \geq 6$. $\quad\square$

Let $\{X_t; t \in T\}$ be a collection of random variables. We say they are *uniformly integrable* if $\sup_{t \in T} E[|X_t| I(|X_t| \geq \beta)] \to 0$ as $\beta \to \infty$. Now we prove the key result in this section. Review $\boldsymbol{\Delta}_n = \mathbf{R}_n - \mathbf{I}$ and (2.27).

LEMMA 2.16. *Let $p := p_n$ satisfy that $m := n - 1 > p$ and $p \to \infty$. Set $Q_n = |\mathbf{I} + \boldsymbol{\Delta}_n \cdot \operatorname{diag}(V_1, \ldots, V_p)|^{-\frac{m}{2} - t}$. Assume $\inf_{n \geq 6} \lambda_{\min}(\mathbf{R}_n) > \frac{1}{2}$. Then there exists $s_0 > 0$ such that $\{Q_n; n \geq 6\}$ is uniformly integrable for any $s \in (0, s_0]$ provided one of the following holds:*

(i) $\inf_{n \geq 6} \frac{p_n}{n} > 0$;
(ii) $\inf_{n \geq 6} \frac{1}{n} \operatorname{tr}(\boldsymbol{\Delta}_n^2) > 0$;
(iii) $\sup_{n \geq 6} \frac{p_n \|\boldsymbol{\Delta}_n\|_\infty}{\|\boldsymbol{\Delta}_n\|_2} < \infty$.

PROOF. Write $a_n = -\lambda_{\min}(\boldsymbol{\Delta}_n)$ for short notation. Lemma 2.3 and the condition $\inf_{n \geq 6} \lambda_{\min}(\mathbf{R}_n) > \frac{1}{2}$ imply that $a_n > 0$ for each $n \geq 6$ and

$$(2.51) \qquad \sup_{n \geq 6} a_n < \frac{1}{2}.$$

Let $\mathbf{D}_n = \operatorname{diag}(\sqrt{V_1}, \ldots, \sqrt{V_p})$. Easily, $\mathbf{M}_n := \mathbf{D}_n(\boldsymbol{\Delta}_n + a_n \mathbf{I}_p)\mathbf{D}_n$ is nonnegative definite. This and the fact that $0 \leq V_i \leq 1$ for each i imply that $\mathbf{I} + a_n \mathbf{D}_n^2$ is positive definite. Hence,

$$|\mathbf{I} + \boldsymbol{\Delta}_n \cdot \operatorname{diag}(V_1, \ldots, V_p)| = |(\mathbf{I} - a_n \mathbf{D}_n^2) + \mathbf{M}_n|$$

$$\geq |\mathbf{I} - a_n \mathbf{D}_n^2|.$$

Set $v = \frac{m}{2} + t$. Notice

$$(2.52) \qquad Q_n \leq R_n := |\mathbf{I} - a_n \mathbf{D}_n^2|^{-v} \leq \frac{1}{(1 - a_n)^{vp}}.$$

By the definition of uniform integrability, it suffices to show that there exists $n_0 \geq 6$ such that

$$\sup_{n \geq n_0} E[Q_n I(Q_n \geq \beta)] \to 0$$

as $\beta \to \infty$. Set $U_n = \sum_{i=1}^p V_i$. Then, for any number $\beta \geq 1$ and $b > 0$,

$$E[Q_n I(Q_n \geq \beta)] \leq E[R_n I(U_n \geq b)] + E[Q_n I(Q_n \geq \beta, U_n \leq b)]$$

$$:= F_n + G_n.$$

In the following two steps, we will show $F_n \to 0$ and $G_n \to 0$, respectively.

Step 1: *the proof that* $F_n \to 0$. It is known $\log(1 - x) \geq -\frac{x}{1-x}$ for $x \in [0, 1)$. Thus, $\log[\prod_{i=1}^{p}(1 - a_n V_i)] \geq -\frac{a_n}{1-a_n} \sum_{i=1}^{p} V_i$. Hence, from (2.52),

$$(2.53) \qquad\qquad R_n \leq \exp\left(\frac{a_n v}{1 - a_n} U_n\right).$$

Set $h = \frac{a_n}{1-a_n} v$. Then $0 \leq h \leq v$ by (2.51). It is trivial to check

$$e^{h U_n} I(b \leq U_n \leq c) \leq e^{hb} I(U_n \geq b) + h \int_b^c e^{hy} I(U_n \geq y)\, dy$$

for any $b < c$ and any realization of U_n by considering if $b \leq U_n \leq c$ is true or not. Take expectations on both sides to have

$$(2.54) \qquad E\left[e^{h U_n} I(b \leq U_n \leq c)\right] \leq e^{hb} P(U_n \geq b) + h \int_b^c e^{hy} P(U_n \geq y)\, dy.$$

Since $\sup_{n \geq 6} a_n < \frac{1}{2}$, it is apparent that $\tau := \sup_{n \geq 6} \frac{a_n}{1-a_n} < 1$ from (2.51). Take $\rho \in (\frac{\tau}{2}, \frac{1}{2})$. By Lemma 2.13, there exist constants $n_1 = n_1 \geq 6$ and $M = M(\rho) > 0$ such that

$$P(U_n \geq y) \leq e^{-\rho m y}$$

for all $y \geq b := b_n = M \frac{pt}{m}$ and $n \geq n_1$. In what follows, n_i denotes constants not depending on n. Thus, we have from (2.54) with $c = \infty$ that

$$E\left[e^{h U_n} I(U_n \geq b)\right]$$
$$\leq \exp\left[\frac{a_n}{1-a_n} vb - \rho m b\right] + h \int_b^\infty \exp\left[\frac{a_n}{1-a_n} vy - \rho m y\right] dy$$
$$= e^{-\beta_n m b} + h \int_b^\infty e^{-\beta_n m y}\, dy,$$

where

$$\beta_n = \rho - \frac{a_n}{1-a_n} \frac{v}{m}.$$

For $\frac{v}{m} \to \frac{1}{2}$ and $0 \leq h \leq m$ for large n, from the choice of ρ, it is readily seen $\beta_n \geq \frac{1}{2}(\rho - \frac{1}{2}\tau) = \frac{1}{4}(2\rho - \tau) > 0$ as $n \geq n_2 \geq n_1$. This gives

$$E\left[e^{h U_n} I(U_n \geq b)\right] \leq \left(1 + \frac{4h}{(2\rho - \tau)m}\right) \exp\left(-\frac{1}{4}(2\rho - \tau)mb\right)$$
$$\leq \left(1 + \frac{4}{2\rho - \tau}\right) \cdot \exp\left\{-\frac{1}{4}(2\rho - \tau)Mpt\right\}$$

for all $n \geq n_3 \geq n_2$. Review (2.53) and the notation $h = \frac{a_n}{1-a_n} v$, we get

$$E\left[R_n I(U_n \geq b)\right] \leq \left(1 + \frac{4}{2\rho - \tau}\right) \cdot \exp\left\{-\frac{1}{4}(2\rho - \tau)Mpt\right\}$$

for all $s > 0$ and $n \geq n_3$. From Lemma 2.6, $pt \to \infty$. Therefore,

$$\lim_{n \to \infty} E[R_n I(U_n \geq b)] = 0.$$

Step 2: *the proof that* $G_n \to 0$. It suffices to show there exist $n_0' \geq 6$ and $s_1 > 0$ such that

$$(2.55) \qquad \lim_{\beta \to \infty} \sup_{n \geq n_0'} E[Q_n I(Q_n \geq \beta, U_n \leq b)] = 0$$

for any $s \in (0, s_1)$, where $b = M \frac{pt}{m}$ is defined as before.

Let $\lambda_1, \dots, \lambda_p$ be the eigenvalues of $\mathbf{D}_n \Delta_{\mathbf{n}} \mathbf{D}_n$. From (2.28),

$$(2.56) \qquad \sum_{i=1}^{p} \lambda_i^2 = \sum_{i \neq j} r_{ij}^2 V_i V_j \leq \sum_{i \neq j} V_i V_j \leq U_n^2.$$

Then, under condition $U_n \leq b = M \frac{pt}{m}$, we see $\sum_{i=1}^{p} \lambda_i^2 \leq \frac{M^2 p^2 t^2}{m^2}$. Recall M does not depend on s. Set $s_0 = \frac{1}{2M}$. Thus, from Lemma 2.6, we see $t^2 = \frac{s^2}{\sigma_n^2} \leq \frac{m^2 s^2}{p^2}$, and hence

$$\sum_{i=1}^{p} \lambda_i^2 \leq \frac{M^2 p^2 t^2}{m^2} \leq \frac{1}{4}$$

for any $s \in (0, s_0)$. In particular, $\max_{1 \leq i \leq p} |\lambda_i| < \frac{1}{2}$. From now on, we assume $s \in (0, s_0)$.

Write $\log(1 + x) = x - x^2 \cdot \varepsilon(x)$ for $|x| < 1$. Then, by Taylor's expansion, $\varepsilon(x) = \sum_{i=0}^{\infty} (-1)^i \frac{1}{i+2} x^i$. Thus, $|\varepsilon(x)| \leq \frac{1}{2} \sum_{i=0}^{\infty} |x|^i = \frac{1}{2(1-|x|)}$ for all $|x| < 1$. So $\sup_{|x| \leq 1/2} |\varepsilon(x)| \leq 1$. This indicates that, under $U_n \leq b$,

$$\log Q_n = -\left(\frac{m}{2} + t\right) \cdot |\mathbf{I} + \mathbf{D}_n \Delta_{\mathbf{n}} \mathbf{D}_n|$$

$$= -\left(\frac{m}{2} + t\right) \sum_{i=1}^{p} \log(1 + \lambda_i)$$

$$= v \sum_{i=1}^{p} \lambda_i^2 \varepsilon(\lambda_i)$$

by the first identity from (2.28), where $\max_{1 \leq i \leq p} |\varepsilon(\lambda_i)| \leq 1$. This and (2.28) say that

$$(2.57) \qquad Q_n I(U_n \leq b) \leq e^{v \Psi_n} \leq e^{m \Psi_n}$$

for all $n \geq n_4 \geq n_3$, where $\Psi_n := \sum_{i \neq j} r_{ij}^2 V_i V_j$. Now we show (2.55) by distinguishing three cases.

Cases (i) *and* (ii). Review Case (i) says $\inf_{n\geq 6}\frac{p_n}{n} > 0$ and Case (ii) is that $\inf_{n\geq 6}\frac{1}{n}\operatorname{tr}(\Delta_n^2) > 0$. The inequalities from (2.56) and (2.57) imply that

$$(2.58) \qquad Q_n I(Q_n \geq \beta, U_n \leq b) \leq e^{m\Psi_n} I\left(\frac{\log\beta}{m} \leq \Psi_n \leq \frac{M^2 p^2 t^2}{m^2}\right)$$

for all $\beta > 0$. By Lemma 2.15, there exists some $0 < \delta < s_0$ for which

$$P\left(\sum_{i\neq j} r_{ij}^2 V_i V_j \geq y\right) \leq \exp\left(-\frac{my}{256}\cdot\frac{m}{pt + m\sqrt{y}}\right)$$

for all $y > \frac{1}{m}$, $s \in (0, \delta]$ and $n \geq n_4 + 6$. Since $pt = \frac{ps}{\sigma_n} \leq ms$ by Lemma 2.6, we have

$$\frac{m}{pt + m\sqrt{y}} \geq \frac{m}{pt + Mpt} \geq \frac{1}{(M+1)s} \geq 512$$

for all $\frac{1}{m} < y \leq \frac{M^2 p^2 t^2}{m^2}$ and $s \in (0, s_1]$ where $s_1 := \min\{\delta, (512(M+1))^{-1}\}$. Thus,

$$P(\Psi_n \geq y) \leq e^{-2my}$$

for all $\frac{1}{m} < y \leq \frac{M^2 p^2 t^2}{m^2}$, $s \in (0, s_1]$ and $n \geq n_4 + 6$. In (2.54), taking $b = b_1 := \frac{\log\beta}{m}$ with $\beta \geq e^2$ and $c = \frac{M^2 p^2 t^2}{m^2}$, then replacing U_n by Ψ_n, we obtain

$$E\left[e^{m\Psi_n} I(b_1 \leq \Psi_n \leq c)\right] \leq e^{mb_1} P(\Psi_n \geq b_1) + m\int_{b_1}^{c} e^{my} P(\Psi_n \geq y)\,dy$$

$$\leq e^{-mb_1} + m\int_{b_1}^{\infty} e^{-my}\,dy$$

$$= 2e^{-mb_1} = \frac{2}{\beta}.$$

This and (2.58) indicate

$$\sup_{n\geq n_4} E\left[Q_n I(Q_n \geq \beta, U_n \leq b)\right] \leq \frac{2}{\beta}$$

for every $s \in (0, s_1]$, every $\beta \geq e^2$ and $n \geq n_4 + 6$. This concludes (2.55).

Case (iii): $\sup_{n\geq 6}\frac{p_n\|\Delta_n\|_\infty}{\|\Delta_n\|_2} < \infty$. Let us continue from (2.57). Under $U_n \leq b$,

$$\Psi_n \leq \|\Delta_n\|_\infty^2 \sum_{i\neq j} V_i V_j$$

$$\leq \|\Delta_n\|_\infty^2 \cdot U_n^2$$

$$\leq M^2\|\Delta_n\|_\infty^2 \cdot \frac{p^2 t^2}{m^2}.$$

Set $K = \sup_{n \geq 6} \frac{p_n \|\boldsymbol{\Delta}_n\|_\infty}{\|\Delta_n\|_2}$. Then $\|\boldsymbol{\Delta}_n\|_\infty \leq K \frac{1}{p_n} \|\boldsymbol{\Delta}_n\|_2$, and thus

$$\Psi_n \leq \frac{(MKs)^2}{m} \cdot \left(\frac{2}{m} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2] \right) \cdot \frac{1}{\sigma_n^2} \leq \frac{(MKs)^2}{m},$$

where the last assertion comes from Lemma 2.6. Therefore, by (2.57), we finally see

$$Q_n I(U_n \leq b) \leq e^{(MKs)^2}$$

for all $n \geq n_4$. Therefore,

$$Q_n I(Q_n \geq \beta, U_n \leq b) = 0$$

as long as $\beta > e^{(MKs)^2}$. This proves (2.55). □

2.6. *Proofs of Theorem 1 and Corollary 2.* With the understanding from Sections 2.1–2.5, we are now fully prepared to prove Theorem 1.

PROOF OF THEOREM 1. Since $N(0,1)$ is uniquely determined by its moments $E[N(0,1)^k]$ for $k = 1, 2, \ldots$, to prove the theorem by Proposition 1.2 and Lemma 2.10, it is enough to show

$$(2.59) \qquad \lim_{n \to \infty} E \exp\left(\frac{\log |\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n} s \right) = e^{s^2/2}$$

for all $s \in (0, s_0)$, where s_0 is a constant not depending on n. From the second conclusion in Lemma 2.1, we have $\log |\hat{\mathbf{R}}_n| \leq 0$, and hence the expectation above is finite for any $s \geq 0$ and $n \geq 6$. Define

$$\mu_{n,0} = \left(p - n + \frac{3}{2} \right) \log\left(1 - \frac{p}{n-1} \right) - \frac{n-2}{n-1} p;$$

$$\sigma_{n,0}^2 = -2 \left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1} \right) \right] \quad \text{and} \quad \sigma_{n,1}^2 = \frac{2}{n-1} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2].$$

We then have

$$(2.60) \qquad \mu_n = \mu_{n,0} + \log |\mathbf{R}_n| \quad \text{and} \quad \sigma_n^2 = \sigma_{n,0}^2 + \sigma_{n,1}^2.$$

Set $t = \frac{s}{\sigma_n}$ for $s > 0$ fixed. By Proposition 1.1,

$$E[|\hat{\mathbf{R}}_n|^t] = \left(\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2} + t)} \right)^p \cdot \frac{\Gamma_p(\frac{n-1}{2} + t)}{\Gamma_p(\frac{n-1}{2})}$$

$$\cdot |\mathbf{R}_n|^t \cdot E[|\mathbf{I} + \boldsymbol{\Delta}_n \cdot \operatorname{diag}(V_1, \ldots, V_p)|^{-\frac{n-1}{2} - t}]$$

$$= E[|\hat{\mathbf{R}}_{n,0}|^t] \cdot |\mathbf{R}_n|^t \cdot E[|\mathbf{I} + \boldsymbol{\Delta}_n \cdot \operatorname{diag}(V_1, \ldots, V_p)|^{-\frac{n-1}{2} - t}],$$

where $V_1, \ldots, V_p$ are i.i.d. with Beta$(t, \frac{n-1}{2})$-distribution and

$$E\big[|\hat{\mathbf{R}}_{n,0}|^t\big] = E\big[|\hat{\mathbf{R}}_n|^t\big]|_{\Delta_n=\mathbf{0}} = \left(\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2}+t)}\right)^p \cdot \frac{\Gamma_p(\frac{n-1}{2}+t)}{\Gamma_p(\frac{n-1}{2})}.$$

Rewrite $E\exp(\frac{\log|\hat{\mathbf{R}}_n|-\mu_n}{\sigma_n}s) = e^{-\mu_n t} \cdot E[|\hat{\mathbf{R}}_n|^t]$. Then, to show (2.59), it suffices to check

$$\big\{\log E\big[|\hat{\mathbf{R}}_{n,0}|^t\big] - \mu_{n,0}t\big\} + \log E\big[|\mathbf{I} + \boldsymbol{\Delta}_n \cdot \mathrm{diag}(V_1, \ldots, V_p)|^{-\frac{n-1}{2}-t}\big]$$

(2.61)

$$:= A_n + B_n \to \frac{s^2}{2}$$

for each $s \in (0, s_0)$. The number "s_0" will be specified later. We will analyze A_n and B_n separately next.

Step 1: *Analysis of A_n.* By Lemma 2.9, there exists $s_0 > 0$ such that, for any subsequence $\{n_j; \ j \geq 1\}$ of positive integers with $\lim_{j\to\infty}\frac{p_{n_j}}{n_j} = y \in [0, 1]$,

$$(2.62) \qquad \lim_{j\to\infty} E\exp\left(\frac{\log|\hat{\mathbf{R}}_{n_j,0}| - \mu_{n_j,0}}{\sigma_{n_j,0}}s\right) = e^{s^2/2}, \qquad |s| \leq s_0.$$

Based on the fact $\frac{p_n}{n} \in [0, 1]$ for all $n \geq 6$, by a subsequence argument, it is easy to see that

$$\lim_{n\to\infty} E\exp\left(\frac{\log|\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}}s\right) = e^{s^2/2}$$

for all $|s| \leq 2s_1 := s_0$. This immediately implies

$$(2.63) \qquad \frac{\log|\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_{n,0}} \to N(0, 1)$$

weakly as $n \to \infty$.

From (2.60), $\frac{\sigma_{n,i}^2}{\sigma_n^2} \in [0, 1]$ for $i = 0, 1$ and all $n \geq 6$. Then, for any subsequence $\{l_k\}$, there exists a further subsequence l_{k_j} such that

$$(2.64) \qquad \lim_{j\to\infty}\frac{\sigma_{n,i}^2}{\sigma_n^2} = \alpha_i \in [0, 1] \quad \text{and} \quad \alpha_0 + \alpha_1 = 1$$

with $n = l_{k_j}$ for $i = 0, 1$. Define $Z_n = \exp(\frac{\log|\hat{\mathbf{R}}_{n,0}|-\mu_{n,0}}{\sigma_n}s)$ for $n \geq 6$ and $s \geq 0$. Then $Z_n \to \exp(s\sqrt{\alpha_0}N(0, 1))$ weakly along the subsequence $\{l_{k_j}\}$ by (2.63). Furthermore, since $0 < |\hat{\mathbf{R}}_{n,0}| \leq 1$, then Z_n is a bounded random variable for all $n \geq 6$. Thus, $\sup_{n\geq 6} E(Z_n^2) < \infty$ for all $|s| \leq s_1$ by (2.62) and the fact $\sigma_n^2 \geq \sigma_{n,0}^2$. It follows that $\{Z_n; \ n \geq 6\}$ are uniformly integrable. So

$$\lim_{j\to\infty} E\exp\left(\frac{\log|\hat{\mathbf{R}}_{n,0}| - \mu_{n,0}}{\sigma_n}s\right) = E\exp(s\sqrt{\alpha_0}N(0, 1)) = e^{\alpha_0 s^2/2}$$

for all $|s| \leq s_1$ if "n" is replaced by "l_{k_j}." So

$$(2.65) \qquad A_n \to \frac{\alpha_0 s^2}{2}$$

along the subsequence $\{l_{k_j}\}$ for all $|s| \leq s_1$.

Step 2: *Analysis of* B_n. Recall $\sigma_{n,1}^2 = \frac{2}{n-1} \operatorname{tr}[(\mathbf{R}_n - \mathbf{I})^2]$. Then

$$\frac{t^2}{n-1} \operatorname{tr}(\mathbf{\Delta}_n^2) = \frac{s^2}{2} \cdot \frac{\sigma_{n,1}^2}{\sigma_n^2} \to \frac{\alpha_1 s^2}{2}$$

along the subsequence $\{l_{k_j}\}$ by (2.64). We then have from Lemma 2.12 that

$$\left| \mathbf{I} + \mathbf{\Delta}_n \cdot \operatorname{diag}(V_1, \ldots, V_p) \right|^{-\frac{n-1}{2}-t} \to e^{\alpha_1 s^2/2}$$

in probability along the subsequence $\{l_{k_j}\}$ for any $s \geq 0$. Lemma 2.16 confirms that the left-hand side above is uniformly integrable for each $s \in [0, s_2]$ where s_2 is a constant. We thus conclude

$$(2.66) \qquad B_n = \log E\left[\left| \mathbf{I} + \mathbf{\Delta}_n \cdot \operatorname{diag}(V_1, \ldots, V_p) \right|^{-\frac{n-1}{2}-t} \right] \to \frac{\alpha_1 s^2}{2}$$

along the subsequence $\{l_{k_j}\}$ for each $s \in [0, s_2]$.

We now make a summary. It is shown that, for any subsequence $\{l_k; k \geq 1\}$ of $\{n; n \geq 1\}$, we find a further subsequence $\{l_{k_j}; j \geq 1\}$ such that (2.65) and (2.66) hold along the subsequence $\{l_{k_j}; j \geq 1\}$ for all $s \in [0, s_3]$ where $s_3 := s_1 \wedge s_2$. This ensures (2.61) by the second equality in (2.64) and the subsequence argument again. $\square$

PROOF OF COROLLARY 2. By Lemma 2.3, $\lambda_{\min}(\mathbf{R}_n) \in [0, 1]$. Let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of $\mathbf{R}_n$. Then, from the Geršgorin disc theorem [see, e.g., page 344 from Horn and Johnson (1985)], we have

$$|1 - \lambda_k| \leq \max_{1 \leq i \leq p} \sum_{j \neq i} |r_{ij}|$$

for any $1 \leq k \leq p$. In particular, this entails $1 - \lambda_{\min}(\mathbf{R}_n) \leq \max_{1 \leq i \leq p} \sum_{j \neq i} |r_{ij}|$. Now we use this fact to prove the corollary:

(i) Notice $\max_{1 \leq i \leq p} \sum_{j \neq i} |r_{ij}| \leq 2(|\rho| + \cdots + |\rho|^{p-1}) \leq \frac{2|\rho|}{1-|\rho|}$. Hence,

$$\lambda_{\min}(\mathbf{R}_n) \geq 1 - \frac{2|\rho|}{1-|\rho|} > \frac{1}{2}$$

if $|\rho| < \frac{1}{5}$. Then the CLT in Theorem 1 holds assuming that $\inf_{n \geq 6} \frac{p_n}{n} > 0$.

(ii) First, fix $n \geq 6$. Then

$$
\begin{aligned}
1 - \lambda_{\min}(\mathbf{R}_n) &\leq \max_{1 \leq i \leq p} \sum_{j \neq i} |r_{ij}| \\
&\leq \max_{1 \leq i \leq p} \left\{ j \neq i; |i - j| \leq k \right\} \cdot \max_{i \neq j} |r_{ij}| \\
&\leq (2k) \cdot \sup_{n \geq 6} \max_{i \neq j} |r_{ij}|.
\end{aligned}
$$

If $\sup_{n \geq 6} \max_{i \neq j} |r_{ij}| < \frac{1}{4k}$, then $\inf_{n \geq 6} \lambda_{\min}(\mathbf{R}_n) > \frac{1}{2}$. Hence, the CLT in Theorem 1 is valid provided $\inf_{n \geq 6} \frac{p_n}{n} > 0$. □

REFERENCES

ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. Wiley, New York. MR0091588

BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. Springer, New York. MR2567175

BAO, Z., PAN, G. and ZHOU, W. (2012). Tracy–Widom law for the extreme eigenvalues of sample correlation matrices. *Electron. J. Probab.* **17** 1–32.

BARTLETT, M. S. (1954). A note on multiplying factors for various chi-squared approximations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **16** 296–298.

BILLINGSLEY, P. (1986). *Probability and Measure*, 2nd ed. Wiley, New York. MR0830424

BROCKWELL, P. J. and DAVIS, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer, New York.

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

CAI, T., LIANG, T. and ZHOU, H. (2015). Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional Gaussian distributions. *J. Multivariate Anal.* **137** 161–172.

CHOW, Y. S. and TEICHER, H. (1988). *Probability Theory: Independence, Interchangeability, Martingales*, 2nd ed. Springer, New York. MR0953964

DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics (New York)* **38**. Springer, New York. MR1619036

DONG, Z., JIANG, T. and LI, D. (2012). Circular law and arc law for truncation of random unitary matrix. *J. Math. Phys.* **53** Article ID 013301. MR2919538

EATON, M. L. (1983). *Multivariate Statistics: A Vector Space Approach*. Wiley, New York. MR0716321

HANSON, D. L. and WRIGHT, F. T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. MR0279864

HORN, R. A. and JOHNSON, C. R. (1985). *Matrix Analysis*. Cambridge Univ. Press, Cambridge. MR0832183

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

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

JIANG, T. and QI, Y. (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scand. J. Stat.* **42** 988–1009. MR3426306

JIANG, T. and YANG, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. Statist.* **41** 2029–2074. MR3127857

LI, D., LIU, W. 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

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

MORRISON, D. F. (2004). *Multivariate Statistical Methods*, 4th ed. Duxbury Press, Pacific Grove, CA.

MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

NGUYEN, H. H. and VU, V. (2014). Random matrices: Law of the determinant. *Ann. Probab.* **42** 146–167. MR3161483

RUDELSON, M. and VERSHYNIN, R. (2013). Hanson–Wright inequality and sub-Gaussian concentration. *Electron. Commun. Probab.* **18** Article ID 82. MR3125258

SMALE, S. (2000). Mathematical problems for the next century. In *Mathematics*: *Frontiers and Perspectives* (V. Arnold, M. Atiyah, P. Lax and B. Mazur, eds.) 271–294. Amer. Math. Soc., Providence, RI. MR1754783

TAO, T. and VU, V. (2012). A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.* **231** 74–101. MR2935384

WILKS, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* **24** 471–494.

YURINSKIĬ, V. V. (1976). Exponential inequalities for sums of random vectors. *J. Multivariate Anal.* **6** 473–499. MR0428401

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

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