The Annals of Statistics
2013, Vol. 41, No. 4, 2029–2074
DOI: 10.1214/13-AOS1134
© Institute of Mathematical Statistics, 2013

CENTRAL LIMIT THEOREMS FOR CLASSICAL LIKELIHOOD RATIO TESTS FOR HIGH-DIMENSIONAL NORMAL DISTRIBUTIONS

BY TIEFENG JIANG[1] AND FAN YANG

University of Minnesota

For random samples of size n obtained from p-variate normal distributions, we consider the classical likelihood ratio tests (LRT) for their means and covariance matrices in the high-dimensional setting. These test statistics have been extensively studied in multivariate analysis, and their limiting distributions under the null hypothesis were proved to be chi-square distributions as n goes to infinity and p remains fixed. In this paper, we consider the high-dimensional case where both p and n go to infinity with $p/n \to y \in (0, 1]$. We prove that the likelihood ratio test statistics under this assumption will converge in distribution to normal distributions with explicit means and variances. We perform the simulation study to show that the likelihood ratio tests using our central limit theorems outperform those using the traditional chi-square approximations for analyzing high-dimensional data.

1. Introduction. Traditional statistical theory, particularly in multivariate analysis, does not contemplate the demands of high dimensionality in data analysis due to technological limitations and/or motivations. Consequently, tests of hypotheses and many other modeling procedures in many classical textbooks of multivariate analysis such as Anderson (1958), Muirhead (1982) and Eaton (1983) are well developed under the assumption that the dimension of the dataset, denoted by p, is considered a fixed small constant or at least negligible compared with the sample size n. However, this assumption is no longer true for many modern datasets because their dimensions can be proportionally large compared with the sample size. For example, the financial data, the consumer data, the modern manufacturing data and the multimedia data all have this feature. More examples of high-dimensional data can be found in Donoho (2000) and Johnstone (2001).

Recently, Bai et al. (2009) develop corrections to the traditional likelihood ratio test (LRT) to make it suitable for testing a high-dimensional normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $H_0 : \boldsymbol{\Sigma} = \mathbf{I}_p$ vs $H_a : \boldsymbol{\Sigma} \neq \mathbf{I}_p$. The test statistic is chosen to be $L_n := \operatorname{tr}(\mathbf{S}) - \log|\mathbf{S}| - p$, where $\mathbf{S}$ is the sample covariance matrix from the data. In their

Received September 2012; revised April 2013.
[1]Supported in part by NSF FRG Grants DMS-04-49365 and DMS-12-08982.
MSC2010 subject classifications. Primary 62H15; secondary 62H10.
Key words and phrases. Likelihood ratio test, central limit theorem, high-dimensional data, multivariate normal distribution, hypothesis test, covariance matrix, mean vector, multivariate Gamma function.

derivation, the dimension p is no longer considered a fixed constant, but rather a variable that goes to infinity along with the sample size n, and the ratio between $p = p_n$ and n converges to a constant y, that is,

$$\lim_{n\to\infty} \frac{p_n}{n} = y \in (0, 1). \tag{1.1}$$

Jiang, Jiang and Yang (2012) further extend Bai's result to cover the case of $y = 1$.

In this paper, we study several other classical likelihood ratio tests for means and covariance matrices of high-dimensional normal distributions. Most of these tests have the asymptotic results for their test statistics derived decades ago under the assumption of a large n but a fixed p. Our results supplement these traditional results in providing alternatives to analyze high-dimensional datasets including the critical case $p/n \to 1$. We will briefly introduce these likelihood ratio tests next. In Section 2, for each LRT described below, we first review the existing literature, then give our central limit theorem (CLT) results when the dimension and the sample size are comparable. We also make graphs and tables on the sizes and powers of these CLTs based on our simulation study to show that, as both p and n are large, the traditional chi-square approximation behaves poorly and our CLTs improve the approximation very much.

- In Section 2.1, for the normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we study the sphericity test $H_0 : \Sigma = \lambda \mathbf{I}_p$ vs $H_a : \boldsymbol{\Sigma} \neq \lambda \mathbf{I}_p$ with λ unspecified. We derive the central limit theorem for the LRT statistic when $p/n \to y \in (0, 1]$. Its proof is given at Section 5.2.
- In Section 2.2, we derive the CLT for the LRT statistic in testing that several components of a vector with distribution $N_p(\mu, \boldsymbol{\Sigma})$ are independent. The proof is presented at Section 5.3.
- In Section 2.3, we consider the LRT with $H_0 : N_p(\mu_1, \boldsymbol{\Sigma}_1) = \cdots = N_p(\mu_k, \boldsymbol{\Sigma}_k)$, that is, several normal distributions are identical. We prove a CLT for the LRT statistic with the assumption $p/n_i \to y_i \in (0, 1]$ where n_i is the sample size of a data set from $N_p(\mu_i, \boldsymbol{\Sigma}_i)$ for $i = 1, 2, \ldots, k$. The proof of the theorem is arranged at Section 5.4.
- In Section 2.4, the test of the equality of the covariance matrices from several normal distributions are studied, that is, $H_0 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$. The LRT statistic is evaluated under the assumption $p/n_i \to y_i \in (0, 1]$ for $i = 1, \ldots, k$. This generalizes the work of Bai et al. (2009) and Jiang, Jiang and Yang (2012) from $k = 2$ to any $k \geq 2$. The proof of our result is given at Section 5.5.
- In Section 2.5, we investigate LRT with $H_0 : \boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \mathbf{I}_p$ for the population distribution $N_p(\mu, \boldsymbol{\Sigma})$. With the dimension p and the sample size n satisfying $p/n \to y \in (0, 1]$, we derive the CLT for the LRT statistic. The corresponding theorem is proved at Section 5.6.
- In Section 2.6, we study the test that the population correlation matrix of a normal distribution is equal to an identity matrix; that is, all of the components of a

normal vector are independent (but not necessarily identically distributed). This is different from the test in Section 2.2 that *several components* of a normal vector are independent. The proof is presented at Section 5.7.

- In Sections 3 and 4, we show some simulation results, state our method of the proofs and conclude by offering some open problems.

One can see the value of $y = \lim(p/n)$ or $y_i = \lim(p/n_i)$ introduced above is restricted to the range that $y \leq 1$. In fact, when $y > 1$, some matrices involved in the LRT statistics do not have a full rank, and consequently their determinants are equal to zero. As a result, the LRT statistics are not defined, or do not exist.

To our knowledge the central limit theorem of the LRT statistics mentioned above in the context of $p/n \to y \in (0, 1]$ are new in the literature. Similar research are Bai et al. (2009) and Jiang, Jiang and Yang (2012). The methods of the proofs in the three papers are different: the random matrix theory is used in Bai et al. (2009); the Selberg integral is used in Jiang, Jiang and Yang (2012). Here we obtain the central limit theorems by analyzing the moments of the LRT statistics.

The organization of the rest of the paper is stated as follows. In Section 2, we give the details for each of the six tests described above. A simulation study on the sizes and powers of these tests is presented in Section 3. A discussion is given in Section 4. The theorems appearing in each section are proved in Section 5. An auxiliary result on complex analysis is proved in the Appendix.

2. Main results. In this section we present the central limit theorems of six classical LRT statistics mentioned in the Introduction. The six central limit theorems are stated in the following six subsections.

2.1. *Testing covariance matrices of normal distributions proportional to identity matrix.* For distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we consider the spherical test

$$H_0 : \boldsymbol{\Sigma} = \lambda \mathbf{I}_p \quad \text{vs} \quad H_a : \Sigma \neq \lambda \mathbf{I}_p \tag{2.1}$$

with λ unspecified. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. $\mathbb{R}^p$-valued random variables with normal distribution $N_p(\mu, \boldsymbol{\Sigma})$. Recall

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \quad \text{and} \quad \mathbf{S} = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'. \tag{2.2}$$

The likelihood ratio test statistic of (2.1) is first derived by Mauchly (1940) as

$$V_n = |\mathbf{S}| \cdot \left(\frac{\operatorname{tr}(\mathbf{S})}{p}\right)^{-p}. \tag{2.3}$$

By Theorem 3.1.2 and Corollary 3.2.19 from Muirhead (1982), under H_0 in (2.1),

$$\frac{n}{\lambda} \cdot \mathbf{S} \quad \text{and} \quad \mathbf{Z}'\mathbf{Z} \text{ have the same distribution,} \tag{2.4}$$

where $\mathbf{Z} := (z_{ij})_{(n-1)\times p}$ and z_{ij}'s are i.i.d. with distribution $N(0,1)$. This says that, with probability one, $\mathbf{S}$ is not of full rank when $p \geq n$, and consequently $|\mathbf{S}| = 0$. This indicates that the likelihood ratio test of (2.1) only exists when $p \leq n-1$. The statistic V_n is commonly known as the ellipticity statistic. Gleser (1966) shows that the likelihood ratio test with the rejection region $\{V_n \leq c_\alpha\}$ (where c_α is chosen so that the test has a significance level of α) is unbiased. A classical asymptotic result shows that

$$-(n-1)\rho \log V_n \text{ converges to } \chi^2(f) \tag{2.5}$$

in distribution as $n \to \infty$ with p fixed, where

$$\rho = 1 - \frac{2p^2 + p + 2}{6(n-1)p} \quad \text{and} \quad f = \frac{1}{2}(p-1)(p+2). \tag{2.6}$$

This can be seen from, for example, Theorem 8.3.7 from Muirhead (1982), the Slutsky lemma and the fact that $\rho = \rho_n \to 1$ as $n \to \infty$ and p is fixed. The quantity ρ is a correction term to improve the convergence rate.

Now we consider the case when both n and p are large. For clarity of taking limit, let $p = p_n$; that is, p depends on n.

THEOREM 1. *Let $n > p+1$ for all $n \geq 3$ and V_n be as in (2.3). Assume $\lim_{n\to\infty} p/n = y \in (0,1]$. Then, under H_0 in (2.1), $(\log V_n - \mu_n)/\sigma_n$ converges in distribution to $N(0,1)$ as $n \to \infty$, where*

$$\mu_n = -p - \left(n - p - \frac{3}{2}\right)\log\left(1 - \frac{p}{n-1}\right) \quad \textit{and}$$

$$\sigma_n^2 = -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right].$$

As discussed below (2.4), the LRT exists as $n \geq p+1$; however, we need a slightly stronger condition than $n > p+1$ because of the definition of σ_n^2. Though λ in (2.1) is unspecified, the limiting distribution in Theorem 1 is pivotal; that is, it does not depend on λ. This is because λ is canceled in the expression of V_n in (2.3): $|\alpha\mathbf{S}| = \alpha^p|\mathbf{S}|$ and $(\mathrm{tr}(\alpha\mathbf{S}))^{-p} = \alpha^{-p} \cdot (\mathrm{tr}(\mathbf{S}))^{-p}$ for any $\alpha > 0$.

Simulation is run on the approximation in (2.5) and the CLT in Theorem 1. The summary is given in Figure 1. It is seen from Figure 1 that the approximation in (2.5) becomes poorer as p becomes larger relative to n, and at the same time the CLT in Theorem 1 becomes more precise. In fact, the chi-square approximation in (2.5) is far from reasonable when p is large: the χ^2 curve and the histogram, which are supposed to be matched, separate from each other with the increase of the value of p. See the caption in Figure 1 for more details.

The sizes and powers of the tests by using (2.5) and by Theorem 1 are estimated from our simulation and summarized in Table 1 at Section 3. A further analysis on this results is presented in the same section.

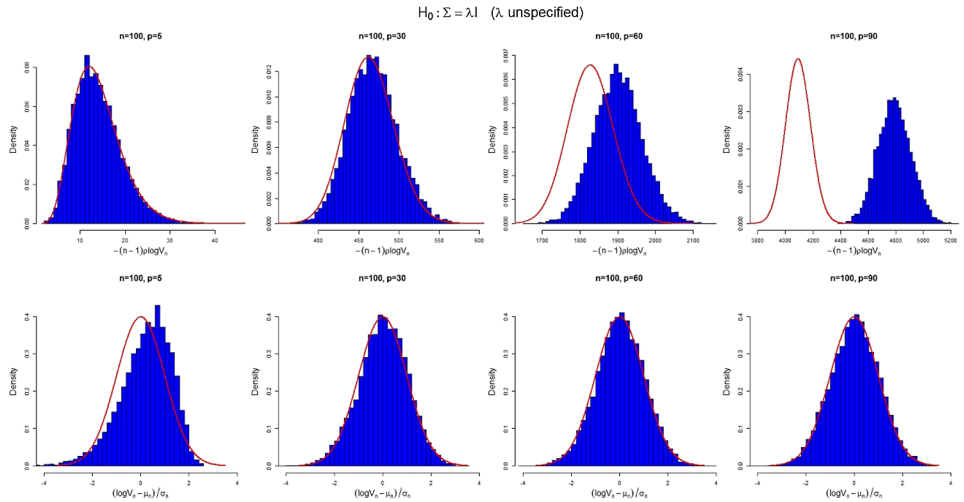

FIG. 1. *Comparison between Theorem 1 and (2.5). We choose* $n = 100$ *with* $p = 5, 30, 60, 90$. *The pictures in the top row show that the* χ^2 *curves stay farther away from the histogram of* $-(n-1)\rho \log V_n$ *as* p *grows. The bottom row shows that the* $N(0,1)$*-curve fits the histogram of* $(\log V_n - \mu_n)/\sigma_n$ *better as* p *becomes larger.*

Finally, when $p \geq n$, we know the LRT does not exist as mentioned above. There are some recent works on choosing other statistics to study the spherical test of (2.1); see, for example, Ledoit and Wolf (2002) and Chen, Zhang and Zhong (2010).

2.2. *Testing independence of components of normal distributions.* Let $k \geq 2$, $p_1, \ldots, p_k$ be positive integers. Denote $p = p_1 + \cdots + p_k$ and

$$\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})_{p\times p} \tag{2.7}$$

as a positive definite matrix, where $\boldsymbol{\Sigma}_{ij}$ is a $p_i \times p_j$ sub-matrix for all $1 \leq i, j \leq k$. Let $N_p(\boldsymbol{\mu}, \Sigma)$ be a p-dimensional normal distribution. We are testing

$$H_0 : \boldsymbol{\Sigma}_{ij} = \mathbf{0} \quad \text{for all } 1 \leq i < j \leq k \quad \text{vs} \quad H_a : H_0 \text{ is not true.} \tag{2.8}$$

In other words, H_0 is equivalent to that $\boldsymbol{\xi}_1, \ldots, \xi_k$ are independent, where $(\boldsymbol{\xi}_1', \ldots, \boldsymbol{\xi}_k')'$ has the distribution $N_p(\mu, \boldsymbol{\Sigma})$ and $\boldsymbol{\xi}_i \in \mathbb{R}^{p_i}$ for $1 \leq i \leq k$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be i.i.d. with distribution $N_p(\mu, \boldsymbol{\Sigma})$. Set $n = N - 1$. Let $\mathbf{S}$ be the covariance matrix as in (2.2). Now we partition $\mathbf{A} := n\mathbf{S}$ in the following way:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1k} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2k} \\ \vdots & \cdots & \cdots & \vdots \\ \mathbf{A}_{k1} & \mathbf{A}_{k2} & \cdots & \mathbf{A}_{kk} \end{pmatrix},$$

where $\mathbf{A}_{ij}$ is a $p_i \times p_j$ matrix. Wilks (1935) shows that the likelihood ratio statistic for testing (2.8) is given by

$$\Lambda_n = \frac{|\mathbf{A}|^{(n+1)/2}}{\prod_{i=1}^k |\mathbf{A}_{ii}|^{(n+1)/2}} := (W_n)^{(n+1)/2}; \tag{2.9}$$

see also Theorem 11.2.1 from Muirhead (1982). Notice that $W_n = 0$ if $p > N = n + 1$, since the matrix $\mathbf{A}$ is not of full rank. From (2.9), we know that the LRT of level α for testing H_0 in (2.8) is $\{\Lambda_n \le c_\alpha\} = \{W_n \le c'_\alpha\}$. Set

$$f = \frac{1}{2}\left(p^2 - \sum_{i=1}^k p_i^2\right) \quad \text{and} \quad \rho = 1 - \frac{2(p^3 - \sum_{i=1}^k p_i^3) + 9(p^2 - \sum_{i=1}^k p_i^2)}{6(n+1)(p^2 - \sum_{i=1}^k p_i^2)}.$$

When n goes to infinity while all p_i's remain fixed, the traditional chi-square approximation to the distribution of Λ_n is referenced from Theorem 11.2.5 in Muirhead (1982),

$$-2\rho \log \Lambda_n \text{ converges to } \chi_f^2 \text{ in distribution} \tag{2.10}$$

as $n \to \infty$. Now we study the case when p_i's are proportional to n. For convenience of taking limit, we assume that p_i depends on n for each $1 \le i \le k$.

THEOREM 2. *Assume $n > p$ for all n and $p_i/n \to y_i \in (0,1)$ as $n \to \infty$ for each $1 \le i \le k$. Recall W_n as defined in (2.9). Then, under H_0 in (2.8), $(\log W_n - \mu_n)/\sigma_n$ converges in distribution to $N(0,1)$ as $n \to \infty$, where*

$$\mu_n = -r_n^2\left(p - n + \frac{1}{2}\right) + \sum_{i=1}^k r_{n,i}^2\left(p_i - n + \frac{1}{2}\right) \quad \textit{and}$$

$$\sigma_n^2 = 2r_n^2 - 2\sum_{i=1}^k r_{n,i}^2$$

and $r_x = (-\log(1 - \frac{p}{x}))^{1/2}$ for $x > p$ and $r_{x,i} = (-\log(1 - \frac{p_i}{x}))^{1/2}$ for $x > p_i$ and $1 \le i \le k$.

Though H_0 in (2.8) involves with unknown $\mathbf{\Sigma}_{ii}$'s, the limiting distribution in Theorem 2 is pivotal. This actually can be quickly seen by transforming $\mathbf{y}_i = \Sigma^{-1/2}(\mathbf{x}_i - \mu)$ for $1 \le i \le N$. Then $\mathbf{y}_1, \ldots, \mathbf{y}_N$ are i.i.d. with distribution $N_p(\mathbf{0}, \mathbf{I}_p)$. Put this into (2.9), and the $\mathbf{\Sigma}_{ii}$'s are then canceled in the fraction under the null hypothesis. See also the interpretation in terms of group transformations on page 532 from Muirhead (1982).

We simulate the two cases in Figure 2: (i) the classical chi-square approximation (2.10); (ii) the central limit theorem based on Theorem 2. The results show that when p becomes large, the classical approximation in (2.10) is poor; however, $(\log W_n - \mu_n)/\sigma_n$ in Theorem 2 fits the standard normal curve very well.

In Table 2 from Section 3, we compare the sizes and powers of the two tests under the chosen H_a explained in the caption. See the detailed explanations in the same section.

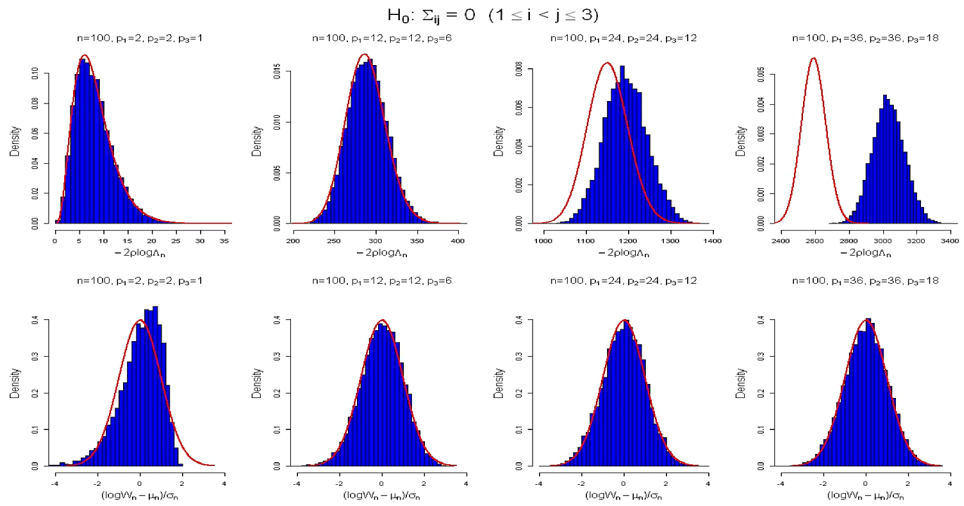

FIG. 2. *Comparison between Theorem 2 and (2.10). We choose* $k = 3$, $n = 100$ *and* $p = 5, 30, 60, 90$ *with* $p_1 : p_2 : p_3 = 2 : 2 : 1$. *The pictures in the top row show that the histogram of* $-2\rho \log \Lambda_n$ *move away gradually from* χ^2 *curve as* p *grows. The pictures in the bottom row indicate that* $(\log W_n - \mu_n)/\sigma_n$ *and* $N(0, 1)$*-curve match better as* p *becomes larger.*

2.3. *Testing that multiple normal distributions are identical.* Given normal distributions $N_p(\mu_i, \mathbf{\Sigma}_i)$, $i = 1, 2, \ldots, k$, we are testing that they are all identical; that is,

$$H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k, \qquad \mathbf{\Sigma}_1 = \cdots = \Sigma_k \quad \text{vs} \quad H_a : H_0 \text{ is not true.} \tag{2.11}$$

Let $\{\mathbf{y}_{ij}; 1 \le i \le k, 1 \le j \le n_i\}$ be independent p-dimensional random vectors, and $\{\mathbf{y}_{ij}; 1 \le j \le n_i\}$ be i.i.d. from $N(\mu_i, \mathbf{\Sigma}_i)$ for each $i = 1, 2, \ldots, k$. Set

$$\mathbf{A} = \sum_{i=1}^{k} n_i (\bar{\mathbf{y}}_i - \bar{\mathbf{y}})(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})', \qquad \mathbf{B}_i = \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)' \quad \text{and}$$

$$\mathbf{B} = \sum_{i=1}^{k} \mathbf{B}_i = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)',$$

where

$$\bar{\mathbf{y}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{y}_{ij}, \qquad \bar{\mathbf{y}} = \frac{1}{n} \sum_{i=1}^{k} n_i \bar{\mathbf{y}}_i, \qquad n = \sum_{i=1}^{k} n_i.$$

The following likelihood ratio test statistic for (2.11) is first derived by Wilks (1932):

$$\Lambda_n = \frac{\prod_{i=1}^{k} |\mathbf{B}_i|^{n_i/2}}{|\mathbf{A} + \mathbf{B}|^{n/2}} \cdot \frac{n^{pn/2}}{\prod_{i=1}^{k} n_i^{pn_i/2}}. \tag{2.12}$$

See also Theorem 10.8.1 from Muirhead (1982). The likelihood ratio test will reject the null hypothesis if $\Lambda_n \le c_\alpha$, where the critical value c_α is determined so that the significance level of the test is equal to α. Note that when $p > n_i$, the matrix $\mathbf{B}_i$ is not of full rank for $i = 1, \dots, k$, and consequently their determinants are equal to zero, so is the likelihood ratio statistic Λ_n. Therefore, to consider the test of (2.11), one needs $p \le \min\{n_i; 1 \le i \le k\}$. Perlman (1980) shows that the LRT is unbiased for testing H_0. Let

$$f = \frac{1}{2}p(k-1)(p+3) \quad \text{and} \quad \rho = 1 - \frac{2p^2+9p+11}{6(k-1)(p+3)n}\left(\sum_{i=1}^{k}\frac{n}{n_i} - 1\right). \tag{2.13}$$

When the dimension p is considered fixed, the following asymptotic distribution of $\log \Lambda_n$ under the null hypothesis (2.11) is a corollary from Theorem 10.8.4 in Muirhead (1982):

$$-2\rho \log \Lambda_n \text{ converges to } \chi_f^2 \tag{2.14}$$

in distribution as $\min_{1\le i\le k} n_i \to \infty$. When p grows with the same rate of n_i, we have the following theorem.

THEOREM 3. *Let $n_i = n_i(p) > p+1$ for all $p \ge 1$ and $\lim_{p\to\infty} p/n_i = y_i \in (0, 1]$ for all $1 \le i \le k$. Let Λ_n be as in (2.12). Then, under H_0 in (2.11),*

$$\frac{\log \Lambda_n - \mu_n}{n\sigma_n} \text{ converges in distribution to } N(0, 1)$$

as $p \to \infty$, where

$$\mu_n = \frac{1}{4}\left[-2kp - \sum_{i=1}^{k} y_i + nr_n^2(2p - 2n + 3) - \sum_{i=1}^{k} n_i r_{n_i'}^2(2p - 2n_i + 3)\right],$$

$$\sigma_n^2 = \frac{1}{2}\left(\sum_{i=1}^{k}\frac{n_i^2}{n^2}r_{n_i'}^2 - r_n^2\right) > 0,$$

$n_i' = n_i - 1$ *and* $r_x = (-\log(1 - \frac{p}{x}))^{1/2}$ *for* $x > p$.

The limiting distribution in Theorem 3 is independent of μ_i's and $\mathbf{\Sigma}_i$'s. This can be seen by defining $\mathbf{z}_{ij} = \mathbf{\Sigma}_1^{-1/2}(\mathbf{y}_{ij} - \mu_1)$, we then know $\mathbf{z}_{ij}$'s are i.i.d. with distribution $N_p(\mathbf{0}, \mathbf{I}_p)$ under the null. It can be easily verified that the μ_i's and Σ_i's are canceled from the numerator and the denominator of Λ_n in (2.12), and hence the right-hand side only depends on $\mathbf{z}_{ij}$'s.

From the simulation shown in Figure 3, we see that when p gets larger, the chi-square curve and the histogram are moving farther apart as p becomes large; however, the normal approximation in Theorem 3 becomes better. The sizes and powers are estimated and summarized in Table 3 at Section 3. See more detailed explanations in the same section.

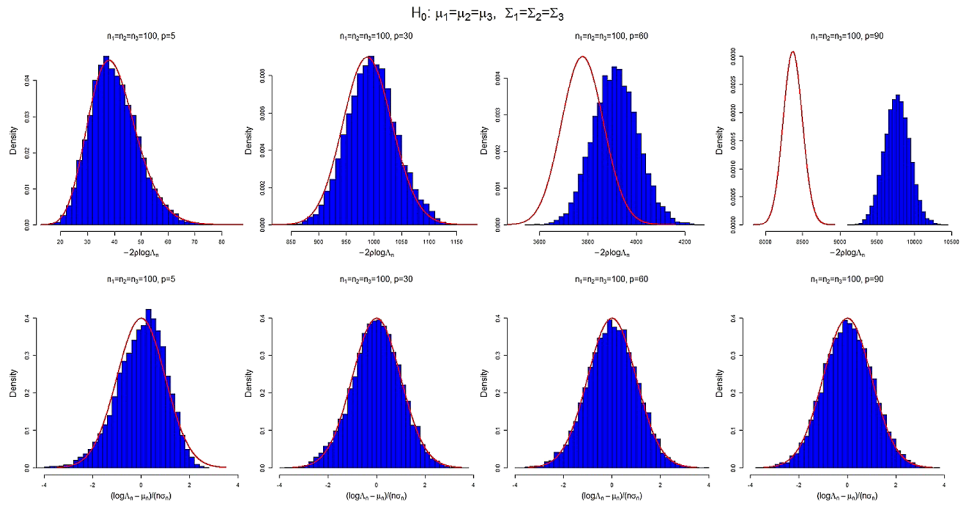

FIG. 3. *Comparison between Theorem 3 and (2.14). We choose* $n_1 = n_2 = n_3 = 100$ *with* $p = 5, 30, 60, 90$. *The pictures in the top row show that the* χ^2 *curves stay farther away gradually from the histogram of* $-2\rho \log \Lambda_n$ *as* p *grows. The pictures in the bottom row show that the* $N(0,1)$*-curve fits the histogram of* $(\log \Lambda_n - \mu_n)/(n\sigma_n)$ *very well as* p *becomes large.*

2.4. *Testing equality of several covariance matrices.* Let $k \geq 2$ be an integer. For $1 \leq i \leq k$, let $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}$ be i.i.d. $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$-distributed random vectors. We are considering

$$H_0 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k \quad \text{vs} \quad H_a : H_0 \text{ is not true.} \tag{2.15}$$

Denote

$$\overline{\mathbf{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{ij} \quad \text{and} \quad \mathbf{A}_i = \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \overline{\mathbf{x}}_i)(\mathbf{x}_{ij} - \overline{\mathbf{x}}_i)', \qquad 1 \leq i \leq k$$

and

$$\mathbf{A} = \mathbf{A}_1 + \cdots + \mathbf{A}_k \quad \text{and} \quad n = n_1 + \cdots + n_k.$$

Wilks (1932) gives the likelihood ratio test of (2.15) with a test statistic

$$\Lambda_n = \frac{\prod_{i=1}^k (\det \mathbf{A}_i)^{n_i/2}}{(\det \mathbf{A})^{n/2}} \cdot \frac{n^{np/2}}{\prod_{i=1}^k n_i^{n_i p/2}} \tag{2.16}$$

and the test rejects the null hypothesis H_0 at $\Lambda_n \leq c_\alpha$, where the critical value c_α is determined so that the test has the significance level of α. Note that $\mathbf{A}_i$ does not have a full rank when $p > n_i$ for any $i = 1, \ldots, k$, and hence their determinants are equal to zero. So the test statistic Λ_n is not defined. Therefore, we assume $p \leq n_i$ for all $i = 1, \ldots, k$ when we study the likelihood ratio test of (2.15). The drawback of the likelihood ratio test is on its bias; see Section 8.2.2 of Muirhead (1982).

Bartlett (1937) suggests using a modified likelihood ratio test statistic Λ_n^* by substituting every sample size n_i with its degree of freedom $n_i - 1$ and substituting the total sample size n with $n - k$,

$$\Lambda_n^* = \frac{\prod_{i=1}^k (\det \mathbf{A}_i)^{(n_i-1)/2}}{(\det \mathbf{A})^{(n-k)/2}} \cdot \frac{(n-k)^{(n-k)p/2}}{\prod_{i=1}^k (n_i-1)^{(n_i-1)p/2}}. \tag{2.17}$$

The unbiased property of this modified likelihood ratio test is proved by Sugiura and Nagao (1968) for $k = 2$ and by Perlman (1980) for a general k. Let

$$f = \frac{1}{2}p(p+1)(k-1) \quad \text{and}$$

$$\rho = 1 - \frac{2p^2+3p-1}{6(p+1)(k-1)(n-k)}\left(\sum_{i=1}^k \frac{n-k}{n_i-1} - 1\right).$$

Box (1949) shows that when p remains fixed, under the null hypothesis (2.15),

$$-2\rho \log \Lambda_n^* \text{ converges to } \chi_f^2 \tag{2.18}$$

in distribution as $\min_{1\le i\le k} n_i \to \infty$; see also Theorem 8.2.7 from Muirhead (1982). Now, suppose p changes with the sample sizes n_i's. We have the following CLT.

THEOREM 4. *Assume $n_i = n_i(p)$ for all $1 \le i \le k$ such that $\min_{1\le i\le k} n_i > p + 1$ and $\lim_{p\to\infty} p/n_i = y_i \in (0, 1]$. Let Λ_n^* be as in (2.17). Then, under H_0 in (2.15), $(\log \Lambda_n^* - \mu_n)/((n-k)\sigma_n)$ converges in distribution to $N(0, 1)$ as $p \to \infty$, where*

$$\mu_n = \frac{1}{4}\Bigg[(n-k)(2n-2p-2k-1)\log\left(1-\frac{p}{n-k}\right) - \sum_{i=1}^k (n_i-1)(2n_i-2p-3)\log\left(1-\frac{p}{n_i-1}\right)\Bigg],$$

$$\sigma_n^2 = \frac{1}{2}\left[\log\left(1-\frac{p}{n-k}\right) - \sum_{i=1}^k \left(\frac{n_i-1}{n-k}\right)^2 \log\left(1-\frac{p}{n_i-1}\right)\right] > 0.$$

The limiting distribution in Theorem 4 is independent of μ_i's and $\mathbf{\Sigma}_i$'s. This is obvious: let $\mathbf{y}_{ij} = \Sigma_i^{-1/2}(\mathbf{x}_{ij} - \mu_i)$, then $\mathbf{y}_i$'s are i.i.d. with distribution $N_p(\mathbf{0}, \mathbf{I}_p)$ under the null. From the cancellation of $\mathbf{\Sigma}_i$'s in Λ_n^* from (2.17), we see that the distribution of Λ_n^* is free of μ_i's, and $\mathbf{\Sigma}_i$'s under H_0.

Bai et al. (2009) and Jiang, Jiang and Yang (2012) study Theorem 4 for the case $k = 2$. Theorem 4 generalizes their results for any $k \ge 2$. Further, the first four authors impose the condition $\max\{y_1, y_2\} < 1$ which excludes the critical case $\max\{y_1, y_2\} = 1$. There is no such a restriction in Theorem 4.

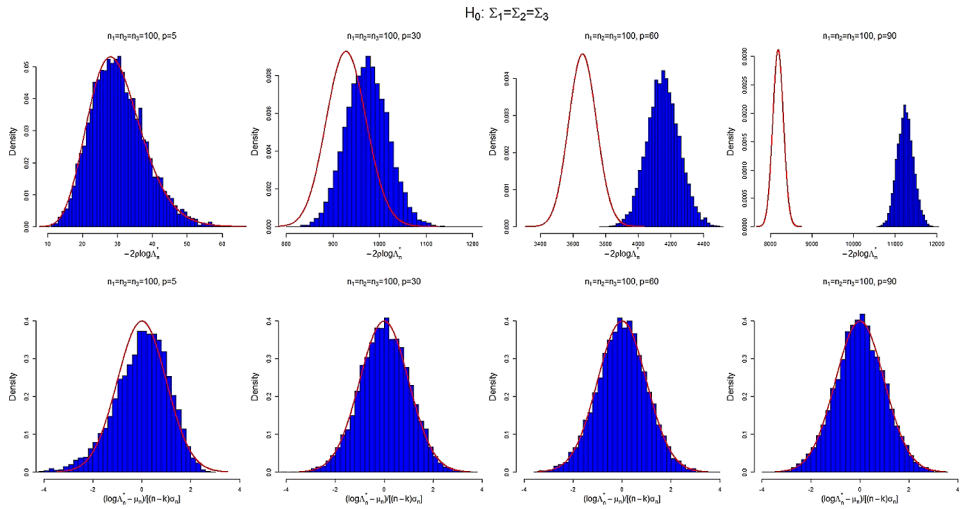

FIG. 4. *Comparison between Theorem 4 and (2.18). We chose* $n_1 = n_2 = n_3 = 100$ *with* $p = 5, 30, 60, 90$. *The pictures in the top row show that the* χ^2 *curves goes away quickly from the histogram of* $-2\rho \log \Lambda_n^*$ *as* p *grows. The pictures in the second row show that the* $N(0, 1)$*-curve fits the histogram of* $(\log \Lambda_n^* - \mu_n)/[(n-k)\sigma_n]$ *better as* p *becomes larger.*

Figure 4 presents our simulation with $k = 3$. It is interesting to see that the chi-square curve and the histogram almost separate from each other when p is large, and at the same time the normal approximation in Theorem 4 becomes very good. In Table 4 from Section 3, we estimate the sizes and powers of the two tests. The analysis is presented in the same section.

2.5. *Testing specified values for mean vector and covariance matrix.* Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. $\mathbb{R}^p$-valued random vectors from a normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathbb{R}^p$ is the mean vector and $\boldsymbol{\Sigma}$ is the $p \times p$ covariance matrix. Consider the following hypothesis test:

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad \text{and} \quad \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0 \quad \text{vs} \quad H_a : H_0 \text{ is not true},$$

where $\boldsymbol{\mu}_0$ is a specified vector in $\mathbb{R}^p$, and $\boldsymbol{\Sigma}_0$ is a specified $p \times p$ nonsingular matrix. By applying the transformation $\tilde{\mathbf{x}}_i = \boldsymbol{\Sigma}^{-1/2}(\mathbf{x}_i - \boldsymbol{\mu}_0)$, this hypothesis test is equivalent to the test of

$$H_0 : \boldsymbol{\mu} = \mathbf{0} \quad \text{and} \quad \boldsymbol{\Sigma} = \mathbf{I}_p \quad \text{vs} \quad H_a : H_0 \text{ is not true}. \tag{2.19}$$

Recall the notation

$$\bar{\boldsymbol{x}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \quad \text{and} \quad \mathbf{A} = \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\boldsymbol{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'. \tag{2.20}$$

The likelihood ratio test of size α of (2.19) rejects H_0 if $\Lambda_n \le c_\alpha$, where

$$\Lambda_n = \left(\frac{e}{n}\right)^{np/2} |\mathbf{A}|^{n/2} e^{-\operatorname{tr}(\mathbf{A})/2} e^{-n\bar{\mathbf{x}}'\bar{\mathbf{x}}/2}. \tag{2.21}$$

See, for example, Theorem 8.5.1 from Muirhead (1982). Note that the matrix $\mathbf{A}$ does not have a full rank when $p \ge n$ as discussed below (2.4); therefore its determinant is equal to zero. This indicates that the likelihood ratio test of (2.19) only exists when $p < n$. Sugiura and Nagao (1968) and Das Gupta (1969) show that this test with a rejection region $\{\Lambda_n \le c_\alpha\}$ is unbiased, where the critical value c_α is chosen so that the test has the significance level of α.

Theorem 8.5.5 from Muirhead (1982) implies that when the null hypothesis $H_0 : \boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \mathbf{I}_p$ is true, the statistic

$$-2\rho \log \Lambda_n \text{ converges to } \chi_f^2 \tag{2.22}$$

as $n \to \infty$ with p being fixed, where

$$\rho = 1 - \frac{2p^2 + 9p + 11}{6n(p+3)} \quad \text{and} \quad f = \frac{1}{2}p(p+3).$$

Obviously, $\rho = \rho_n \to 1$ in this case. Davis (1971) improves the above result with a second order approximation. Nagarsenker and Pillai (1973) study the exact null distribution of $-2\log\Lambda_n$ by using its moments. Now we state our CLT result when p grows with n.

THEOREM 5. *Assume that $p := p_n$ such that $n > 1 + p$ for all $n \ge 3$ and $\lim_{n\to\infty} p/n = y \in (0, 1]$. Let Λ_n be defined as in (2.21). Then under $H_0 : \boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I}_p$, $(\log\Lambda_n - \mu_n)/(n\sigma_n)$ converges in distribution to $N(0, 1)$ as $n \to \infty$, where*

$$\mu_n = -\frac{1}{4}\left[n(2n-2p-3)\log\left(1 - \frac{p}{n-1}\right) + 2(n+1)p\right] \quad \text{and}$$

$$\sigma_n^2 = -\frac{1}{2}\left(\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right) > 0.$$

The simulations shown in Figure 5 confirm that it is good to use Theorem 5 when p is large and proportional to n rather than the traditional chi-square approximation in (2.22). In Table 5 from Section 3, we study the sizes and powers for the two tests based on the χ^2 approximation and our CLT. The understanding of the table is elaborated in the same section.

2.6. *Testing complete independence.* In this section, we study the likelihood ratio test of the complete independence of the coordinates of a high-dimensional

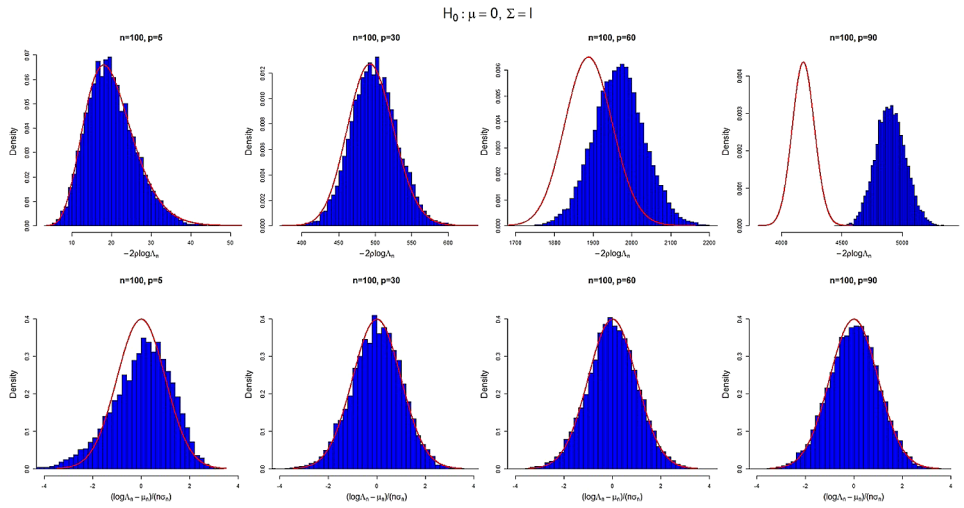

FIG. 5. *Comparison between Theorem 5 and (2.22). We chose* $n = 100$ *with* $p = 5, 30, 60, 90$. *The pictures in the top row show that the* χ^2*-curve stays away gradually from the histogram of* $-2\log\Lambda_n$ *as* p *grows, whereas the* $N(0,1)$*-curve fits statistic* $(\log\Lambda_n - \mu_n)/(n\sigma_n)$ *better as shown from the bottom row.*

normal random vector. Precisely, let $\mathbf{R} = (r_{ij})_{p\times p}$ be the correlation matrix generated from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{x} = (x_1, \ldots, x_p)' \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The test is

$$H_0 : \mathbf{R} = \mathbf{I} \quad \text{vs} \quad H_a : \mathbf{R} \neq \mathbf{I}. \tag{2.23}$$

The null hypothesis H_0 is equivalent to that $x_1, \ldots, x_p$ are independent, or $\boldsymbol{\Sigma}$ is diagonal. To study the LRT, we need to understand the determinant of a sample correlation matrix generated by normal random vectors. In fact we will have a conclusion on the class of spherical distributions, which is more general than that of the normal distributions. Let us first review two terminologies.

Let $\mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_n)' \in \mathbb{R}^n$. Recall the *Pearson correlation coefficient* r defined by

$$r = r_{\mathbf{x},\mathbf{y}} = \frac{\sum_{i=1}^n (x_i - \bar{\mathbf{x}})(y_i - \bar{\mathbf{y}})}{\sqrt{\sum_{i=1}^n (x_i - \bar{\mathbf{x}})^2 \cdot \sum_{i=1}^n (y_i - \bar{\mathbf{y}})^2}}, \tag{2.24}$$

where $\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^n x_i$ and $\bar{\mathbf{y}} = \frac{1}{n}\sum_{i=1}^n y_i$.

We say a random vector $\mathbf{x} \in \mathbb{R}^n$ has a *spherical distribution* if $\mathbf{Ox}$ and $\mathbf{x}$ have the same probability distribution for all $n \times n$ orthogonal matrices $\mathbf{O}$. Examples include the multivariate normal distribution $N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$, the "$\varepsilon$-contaminated" normal distribution $(1-\varepsilon)N_n(\mathbf{0}, \mathbf{I}_n) + \varepsilon N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$ with $\sigma > 0$ and $\varepsilon \in [0,1]$ and the multivariate t distributions. See page 33 from Muirhead (1982) for more discussions.

Let $\mathbf{X} = (x_{ij})_{n\times p} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)' = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$ be an $n \times p$ matrix such that $\mathbf{y}_1, \ldots, \mathbf{y}_p$ are independent random vectors with n-variate spherical distributions

and $P(\mathbf{y}_i = 0) = 0$ for all $1 \le i \le p$ (these distributions may be different). Let $r_{ij} = r_{\mathbf{y}_i,\mathbf{y}_j}$, that is, the Pearson correlation coefficient between $\mathbf{y}_i$ and $\mathbf{y}_j$ for $1 \le i \le j \le n$. Then

$$\mathbf{R}_n := (r_{ij})_{p\times p} \tag{2.25}$$

is the sample correlation matrix. It is known that $\mathbf{R}_n$ can be written as $\mathbf{R}_n = \mathbf{U}'\mathbf{U}$ where $\mathbf{U}$ is an $n \times p$ matrix; see, for example, Jiang (2004a). Thus, $\mathbf{R}_n$ does not have a full rank and hence $|\mathbf{R}_n| = 0$ if $p > n$. According to Theorem 5.1.3 from Muirhead (1982), the density function of $\mathbf{R}_n$ is given by

$$\text{Constant} \cdot |\mathbf{R}_n|^{(n-p-2)/2} d\mathbf{R}_n. \tag{2.26}$$

In the aspect of random matrix theory, the limiting behavior of the largest eigenvalues of $\mathbf{R}_n$ and the empirical distributions of the eigenvalues of $\mathbf{R}_n$ are investigated by Jiang (2004a). For considering the construction of compressed sensing matrices, the statistical testing problems, the covariance structures of normal distributions, high-dimensional regression in statistics and a wide range of applications including signal processing, medical imaging and seismology, the largest off-diagonal entries of $\mathbf{R}_n$ are studied by Jiang (2004b), Li and Rosalsky (2006), Zhou (2007), Liu, Lin and Shao (2008), Li, Liu and Rosalsky (2010), Li, Qi and Rosalsky (2012) and Cai and Jiang (2011, 2012).

Let us now focus on the LRT of (2.23). According to page 40 from Morrison (2004), the likelihood ratio test will reject the null hypothesis of (2.23) if

$$|\mathbf{R}_n|^{n/2} \le c_\alpha, \tag{2.27}$$

where c_α is determined so that the test has significant level of α. It is also known [see, e.g., Bartlett (1954) or page 40 from Morrison (2004)] that when the dimension p remains fixed and the sample size $n \to \infty$,

$$-\left(n - 1 - \frac{2p+5}{6}\right) \log|\mathbf{R}_n| \xrightarrow{d} \chi^2_{p(p-1)/2}. \tag{2.28}$$

This asymptotic result has been used for testing the complete independence of all the coordinates of a normal random vector in the traditional multivariate analysis when p is small relative to n.

Now we study the LRT statistic when p and n are large and at the same scale. First, we give a general CLT result on spherical distributions.

THEOREM 6. *Let $p = p_n$ satisfy $n \ge p+5$ and $\lim_{n\to\infty} p/n = y \in (0, 1]$. Let $\mathbf{X} = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$ be an $n \times p$ matrix such that $\mathbf{y}_1, \ldots, \mathbf{y}_p$ are independent random vectors with n-variate spherical distribution and $P(\mathbf{y}_i = \mathbf{0}) = 0$ for all $1 \le i \le p$ (these distributions may be different). Recall $\mathbf{R}_n$ in (2.25). Then $(\log|\mathbf{R}_n| - \mu_n)/\sigma_n$ converges in distribution to $N(0, 1)$ as $n \to \infty$, where*

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

$$\sigma_n^2 = -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right].$$

In the definition of σ_n^2 above, we need the condition $n \geq p+2$. However, the assumption "$n \geq p+5$" still looks a bit stronger than "$n \geq p+2$." In fact, we use the stronger one as a technical condition in the proof of Lemma 5.10 which involves the complex analysis.

Notice that when the random vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. from a p-variate normal distribution $N_p(\mu, \mathbf{\Sigma})$ with complete independence (i.e., $\mathbf{\Sigma}$ is a diagonal matrix or the correlation matrix $\mathbf{R} = \mathbf{I}_p$). Write $\mathbf{X} = (x_{ij})_{n\times p} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)' = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$. Then, $\mathbf{y}_1, \ldots, \mathbf{y}_p$ are independent random vectors from n-variate normal distributions (these normal distributions may differ by their covariance matrices). It is also obvious that in this case $P(\mathbf{y}_i = 0) = 0$ for all $1 \leq i \leq p$. Therefore, we have the following corollary.

COROLLARY 1. *Assume that $p := p_n$ satisfy that $n \geq p+5$ and $\lim_{n\to\infty} p/n = y \in (0, 1]$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be i.i.d. from $N_p(\mu, \mathbf{\Sigma})$ with the Pearson sample correlation matrix $\mathbf{R}_n$ as defined in (2.25). Then, under H_0 in (2.23), $(\log|\mathbf{R}_n| - \mu_n)/\sigma_n$ converges in distribution to $N(0, 1)$ as $n \to \infty$, where*

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

$$\sigma_n^2 = -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right] > 0.$$

According to Corollary 1, the set $\{(\log|\mathbf{R}_n| - \mu_n)/\sigma_n \leq -z_\alpha\}$ is the rejection region with an asymptotic $1-\alpha$ confidence level for the LRT of (2.23), where the critical value $z_\alpha > 0$ satisfies that $P(N(0, 1) > z_\alpha) = \alpha$ for all $\alpha \in (0, 1)$. Figure 6 shows that the chi-square approximation in (2.28) is good when p is small, but behaves poorly as p is large. At the same time, the normal approximation in Corollary 1 becomes better.

We simulate the sizes and powers of the two tests according to the chi-square approximation in (2.28) and the CLT in Corollary 1 in Table 6 at Section 3. See more analysis in the same section.

As mention earlier, when $p > n$, the LRT statistic $\log|\mathbf{R}_n|$ is not defined. So one has to choose other statistics rather than $\log|\mathbf{R}_n|$ to study (2.23). See, for example, Schott (2005) and Cai and Ma (2013) for recent progress.

3. Simulation study: Sizes and powers. In this part, for each of the six LRTs discussed earlier, we run simulation with 10,000 iterations to estimate the sizes and the powers of the LRTs using the CLT approximation and the classical χ^2 approximation. An analysis for each table is given. In the following discussion, the notation $\mathbf{J}_p$ stands for the $p \times p$ matrix whose entries are all equal to 1, and $[x]$ stands for the integer part of $x > 0$.

(1) *Table* 1. This table corresponds to the sphericity test that, for $N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, $H_0 : \Sigma = \lambda\mathbf{I}_p$ vs $H_a : \mathbf{\Sigma} \neq \lambda\mathbf{I}_p$ with λ unspecified. It is studied in Section 2.1. As

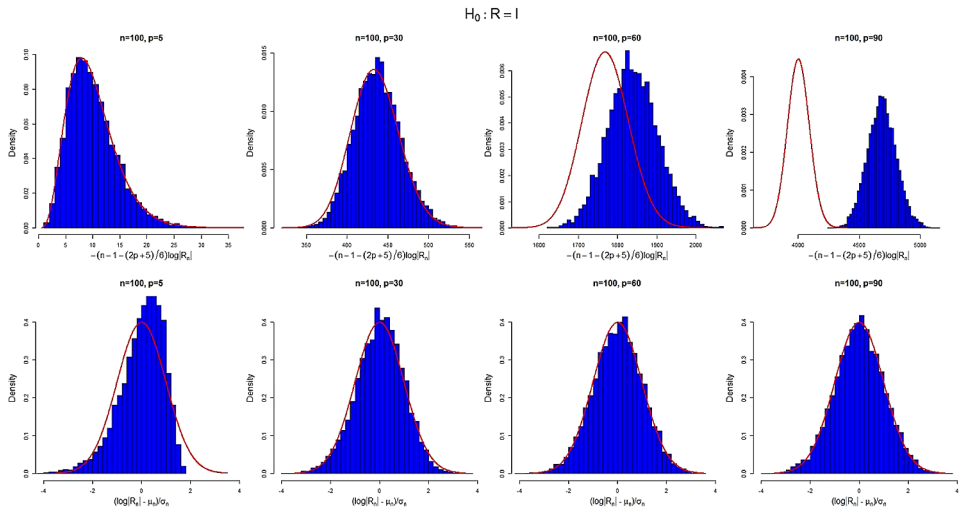

FIG. 6. *Comparison between Corollary 1 and (2.28). We choose* $n = 100$ *with* $p = 5, 30, 60, 90$. *The pictures in the first row show that, as* p *becomes large, the* χ^2*-curve fits the histogram of* $-(n - 1 - \frac{2p+5}{6}) \log |\mathbf{R}_n|$ *poorly. Those in the second row indicate that the* $N(0, 1)$*-curve fits the histogram of* $(\log |\mathbf{R}_n| - \mu_n)/\sigma_n$ *very well as* p *becomes large.*

expected, the χ^2 approximation is good when p is small relative to n, but not when p is large. For example, at $n = 100$ and $p = 60$, the size (type I error or alpha error) for our normal approximation is 0.0511 and power is 0.7914, but the size for χ^2 approximation is 0.3184, which is too large to be used in practice. It is very interesting to see that our normal approximation is also as good as the χ^2 approximation even when p is small. Moreover, for $n = 100$ and $p = 90$ where the ratio $y = 0.9$ is close to 1, the type I error in the CLT case is close to 5%, and the power is still decent at 0.5406. Further, the power for the case of CLT drops as

TABLE 1
Size and power of LRT for sphericity in Section 2.1

	Size under H_0		**Power under H_a**	
	CLT	**χ^2 approx.**	**CLT**	**χ^2 approx.**
$n = 100, p = 5$	0.0562	0.0491	0.7525	0.7317
$n = 100, p = 30$	0.0581	0.0686	0.8700	0.8867
$n = 100, p = 60$	0.0511	0.3184	0.7914	0.9759
$n = 100, p = 90$	0.0518	1.0000	0.5406	1.0000

Notes: The sizes (alpha errors) are estimated based on 10,000 simulations from $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers are estimated under the alternative hypothesis that $\boldsymbol{\Sigma} = \text{diag}(1.69, \ldots, 1.69, 1, \ldots, 1)$, where the number of 1.69 on the diagonal is equal to $[p/2]$.

TABLE 2
Size and power of LRT for independence of three components in Section 2.2

	Size under H_0		Power under H_a	
	CLT	**χ^2 approx.**	**CLT**	**χ^2 approx.**
$n = 100$, $p_1 = 2$, $p_2 = 2$, $p_3 = 1$	0.0663	0.0475	0.7607	0.7144
$n = 100$, $p_1 = 12$, $p_2 = 12$, $p_3 = 6$	0.0552	0.0586	0.9784	0.9793
$n = 100$, $p_1 = 24$, $p_2 = 24$, $p_3 = 12$	0.0507	0.2206	0.8638	0.9733
$n = 100$, $p_1 = 36$, $p_2 = 36$, $p_3 = 18$	0.0547	0.9997	0.4823	1.0000

Notes: The sizes (alpha errors) are estimated based on 10,000 simulations from $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers are estimated under the alternative hypothesis that $\boldsymbol{\Sigma} = 0.15\mathbf{J}_p + 0.85\mathbf{I}_p$.

the ratio p/n increases to 1. This makes sense because the convergence rate of the CLT becomes slow. This can be seen from Theorem 1 that $\sigma_n \to \infty$ as $p/n \to 1$.

(2) *Table* 2. In this table, we compare the sizes and powers of two tests under the chosen H_a explained in the caption. The first one is the classical χ^2-approximation in (2.10) and the second is the CLT in Theorem 2 for the hypothesis that some components of a normal distribution are independent. We observe from the chart that our CLT approximation and the classical χ^2 approximation are comparable for the small values of p_i's. However, when p_i's are large, noticing the last two rows in the table, our test is good whereas the χ^2 approximation is no longer applicable because of the large sizes (type I errors). The power for the CLT drops when the values of p_i's become large. This follows from Theorem 2 that $\sigma_n \to \infty$ as $\sum p_i/n \to 1$, and hence the CLT-approximation does not perform well.

(3) *Table* 3. We create this table for testing that several normal distributions are identical in Section 2.3. It is easily seen that our CLT is good in all cases (except at the case of $p = 5$ where the type I error in our test is 0.0621, slightly higher

TABLE 3
Size and power of LRT for equality of three distributions in Section 2.3

	Size under H_0		Power under H_a	
	CLT	**χ^2 approx.**	**CLT**	**χ^2 approx.**
$n_1 = n_2 = n_3 = 100$, $p = 5$	0.0621	0.0512	0.7420	0.7135
$n_1 = n_2 = n_3 = 100$, $p = 30$	0.0588	0.0743	0.8727	0.8936
$n_1 = n_2 = n_3 = 100$, $p = 60$	0.0531	0.4542	0.6864	0.9770
$n_1 = n_2 = n_3 = 100$, $p = 90$	0.0488	1.0000	0.3493	1.0000

Notes: The sizes (alpha errors) are estimated based on 10,000 simulations from three normal distributions of $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers were estimated under the alternative hypothesis that $\boldsymbol{\mu}_1 = (0, \ldots, 0)'$, $\boldsymbol{\Sigma}_1 = 0.5\mathbf{J}_p + 0.5\mathbf{I}_p$; $\boldsymbol{\mu}_2 = (0.1, \ldots, 0.1)'$, $\boldsymbol{\Sigma}_2 = 0.6\mathbf{J}_p + 0.4\mathbf{I}_p$; $\boldsymbol{\mu}_3 = (0.1, \ldots, 0.1)'$, $\boldsymbol{\Sigma}_3 = 0.5\mathbf{J}_p + 0.31\mathbf{I}_p$.

TABLE 4
Size and power of LRT for equality of three covariance matrices in Section 2.4

	Size under H_0		Power under H_a	
	CLT	**χ^2 approx.**	**CLT**	**χ^2 approx.**
$n_1 = n_2 = n_3 = 100$, $p = 5$	0.0805	0.0567	0.7157	0.6586
$n_1 = n_2 = n_3 = 100$, $p = 30$	0.0516	0.2607	0.6789	0.9218
$n_1 = n_2 = n_3 = 100$, $p = 60$	0.0525	0.9998	0.4493	1.0000
$n_1 = n_2 = n_3 = 100$, $p = 90$	0.0535	1.0000	0.2297	1.0000

Notes: The sizes (alpha errors) are estimated based on 10,000 simulations from $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers are estimated under the alternative hypothesis that $\mathbf{\Sigma}_1 = \mathbf{I}_p$, $\mathbf{\Sigma}_2 = 1.1\mathbf{I}_p$, and $\mathbf{\Sigma}_3 = 0.9\mathbf{I}_p$.

than 0.0512 in the classical case). But when $p = 60$ and $n_1 = n_2 = n_3 = 100$, the size in the classical case is 0.4542, too large to be used. It is worthwhile to notice that the power on the CLT becomes smaller as the value of p becomes larger. It is easily understood from Theorem 3 that the standard deviation diverges to infinity when $p/n \to 1$. Equivalently, the convergence rate is poorer when p gets closer to n.

(4) *Table* 4. This table relates to the test of the equality of the covariance matrices from k normal distributions studied in Section 2.4. We take $k = 3$ in our simulations. The sizes and powers of the chi-square approximation and the CLT in Theorem 4 are summarized in the table. When $p = 5$ and $n_1 = n_2 = n_3 = 100$, our CLT approximation gives a reasonable size of the test while the classical χ^2 approximation is a bit better. However, for the same values of n_i's, when $p = 30, 60, 90$, the size for the χ^2 approximation is 0.2607, 0.9998 and 1, respectively, which are not recommended to be used in practice. Similar to the previous tests, $\sigma_n \to \infty$ as $p/n \to 1$, where σ_n^2 is as in Theorem 4. This implies that the convergence of the CLT is slow in this case. So it is not surprising to see that the power of the test based on the CLT in the table reduces as $p/n \to 1$.

(5) *Table* 5. We generate this table by considering the LRT with $H_0 : \boldsymbol{\mu} = \mathbf{0}$, $\Sigma = \mathbf{I}_p$ for the population distribution $N_p(\boldsymbol{\mu}, \Sigma)$. The CLT is developed in Theorem 5. In this table we study the sizes and powers for the two cases based on the χ^2 approximation and the CLT. At $n = 100$, $p = 5$ (p is small), the χ^2 test outperforms ours. The two cases are equally good at $n = 100$, $p = 30$. When the values of p are large at 60 and 90, our CLT is still good, but the χ^2 approximation is no longer useful. At the same time, it is easy to spot from the fourth column of the table that the power for the CLT-test drops as the ratio p/n becomes large. It is obvious from Theorem 5 that the standard deviation σ_n goes to infinity as the ratio approaches one. This causes drop in precision when the sample size is not large.

(6) *Table* 6. This chart is created on the test that all of the components of a normal vector are independent (but not necessarily identically distributed). It is studied in Corollary 1. The sizes and powers of the two tests are estimated from

TABLE 5
Size and power of LRT for specified normal distribution in Section 2.5

	Size under H_0		**Power under H_a**	
	CLT	**χ^2 approx.**	**CLT**	**χ^2 approx.**
$n = 100$, $p = 5$	0.0986	0.0471	0.5106	0.3818
$n = 100$, $p = 30$	0.0611	0.0657	0.7839	0.7898
$n = 100$, $p = 60$	0.0584	0.3423	0.7150	0.9583
$n = 100$, $p = 90$	0.0571	1.0000	0.4752	1.0000

Notes: Sizes (alpha errors) are estimated based on 10,000 simulations from $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers are estimated under the alternative hypothesis that $\boldsymbol{\mu} = (0.1, \ldots, 0.1, 0, \ldots, 0)'$ where the number of 0.1 is equal to $[p/2]$ and $\boldsymbol{\Sigma} = \{\sigma_{ij}\}$ where $\sigma_{ij} = 1$ for $i = j$, $\sigma_{ij} = 0.1$ for $0 < |i - j| \le 3$, and $\sigma_{ij} = 0$ for $|i - j| > 3$.

simulation using the chi-square approximation in (2.28) and the CLT in Corollary 1 from Section 3 (the H_a is explained in the caption). At all of the four cases of $n = 100$ with $p = 5, 30, 60$ and 90, the performance of our CLT-test is good, and it is even comparable with the classical χ^2-test at the small value of $p = 5$. When $p = 60$ and 90, the sizes of the χ^2-test are too big, while those of the CLT-test keep around 0.05. For the CLT-test itself, looking at the third and fourth rows of the table, though the performance corresponding to $y = p/n = 0.6$ is better than that corresponding to the high value of $y = p/n = 0.9$ as expected, they are quite close. The only difference is the declining of the power as the rate p/n increases. Again, this is easily seen from Corollary 1 that the standard deviation σ_n is divergent as p is close to n.

4. Conclusions and discussions. In this paper, we consider the likelihood ratio tests for the mean vectors and covariance matrices of high-dimensional normal

TABLE 6
Size and power of LRT for complete independence in Section 2.6

	Size under H_0		**Power under H_a**	
	CLT	**χ^2 approx.**	**CLT**	**χ^2 approx.**
$n = 100$, $p = 5$	0.0548	0.0520	0.4311	0.4236
$n = 100$, $p = 30$	0.0526	0.0606	0.6658	0.6945
$n = 100$, $p = 60$	0.0522	0.3148	0.5828	0.9130
$n = 100$, $p = 90$	0.0560	1.0000	0.3811	1.0000

Notes: Sizes (alpha errors) are estimated based on 10,000 simulations from $N_p(\mathbf{0}, \mathbf{I}_p)$. The powers are estimated under the alternative hypothesis that the correlation matrix $\mathbf{R} = (r_{ij})$ where $r_{ij} = 1$ for $i = j$, $r_{ij} = 0.1$ for $0 < |i - j| \le 3$, and $r_{ij} = 0$ for $|i - j| > 3$.

distributions. Traditionally, these tests were performed by using the chi-square approximation. However, this approximation relies on a theoretical assumption that the sample size n goes to infinity, while the dimension p remains fixed. As many modern datasets discussed in the Introduction feature high dimensions, these traditional likelihood ratio tests were shown to be less accurate in analyzing those datasets.

Motivated by the pioneer work of Bai et al. (2009) and Jiang, Jiang and Yang (2012), who prove two central limit theorems of the likelihood ratio test statistics for testing the high-dimensional covariance matrices of normal distributions, we examine in this paper other LRTs that are widely used in the multivariate analysis and prove the central limit theorems for their test statistics. By using the method developed in Jiang, Jiang and Yang (2012), that is, the asymptotic expansion of the multivariate Gamma function with high dimension p, we are able to derive the central limit theorems without relying on concrete random matrix models as demonstrated in Bai et al. (2009). Our method also has an advantage that the central limit theorems for the critical cases $\lim(p/n) = y = 1$ or $\lim(p/n_i) = y_i = 1$ are all derived, which is not the case in Bai et al. (2009) because of the restriction of their tools from the random matrix theory. In real data analysis, as long as $n > p + 1$ or $n_i > p + 1$ in Theorems 1–5, or $n \geq p + 5$ in Theorem 6, we simply take $y = p/n$ or $y_i = p/n_i$ to use the theorems. As Figures 1–6 and Tables 1–6 show, our CLT-approximations are all good even though p is relatively small.

The proofs in this paper are based on the analysis of the moments of the LRT statistics (five of six such moments are from literature, and the last one is derived by us as in Lemma 5.10). The moment method we use here is different from that of the random matrix theory employed in Bai et al. (2009) and the Selberg integral used in Jiang, Jiang and Yang (2012).

Our research also brings out the following four interesting open problems:

(1) All our central limit theorems in this paper are proved under the null hypothesis. As people want to assess the power of the test in many cases, it is also interesting to study the distribution of the test statistic under an alternative hypothesis. In the traditional case where p is considered to be fixed while n goes to infinity, the asymptotic distributions of many likelihood ratio statistics under the alternative hypotheses are derived by using the zonal polynomials; see, for example, Section 8.2.6, Section 8.3.4, Section 8.4.5 from Muirhead (1982). It can be conjectured that in the high-dimensional case, there could be some new results regarding the limiting distributions of the test statistics under the alternative hypotheses. However, this is nontrivial and may require more investigation of the high-dimensional zonal polynomials. Some new understanding about the connection between the random matrix theory and the Jack polynomials (the zonal polynomials, the Schur polynomials and the zonal spherical functions are special cases) is given by Jiang and Matsumoto (2011). In a recent work, Bai et al. (2009) study the high-dimensional LRTs through the random matrix theory. So the connection among the random matrix theory, LRTs and the Jack polynomials is obvious. We

are almost sure that the understanding by Jiang and Matsumoto (2011) will be useful in exploring the LRT statistics under the alternative hypotheses.

(2) Except Theorem 6 where the condition $n \geq p + 5$ is imposed due to a technical constraint, all of the central limit theorems in this paper are proved under the condition $n > p + 1$ or $n_i > p + 1$. This is because when this is not the case in the five theorems, the likelihood ratio statistics will become undefined in these five cases. This indicates that tests other than the likelihood ratio ones shall be developed for analyzing a dataset with p greater than n. For recent progress, see, for example, Ledoit and Wolf (2002) and Chen, Zhang and Zhong (2010) for the sphericity test, and Schott (2001, 2007) for testing the equality of multiple covariance matrices and Srivastava (2005) for testing the covariance matrix of a normal distribution. A power study for sphericity test is tried by Onatski, Moreira and Hallin (2013). Despite these enlightening works mentioned above, other hypothesis tests for $p > n$ or $p > n_i$ are still an open area with many interesting problems to be solved.

(3) In this paper we consider the cases when p and n or n_i are proportional to each other, that is, $\lim p/n = y \in (0, 1]$ or $\lim p/n_i = y_i \in (0, 1]$. In practice, p may be large, but may not be large enough to be at the same scale of n or n_i. So it is useful to derive the central limit theorems appearing in this paper under the assumption that $p \to \infty$ such that $p/n \to 0$ or $p/n_i \to 0$.

(4) To understand the robustness of the six likelihood tests in this paper, one has to study the limiting behaviors of the LRT statistics without the normality assumptions. This is feasible. For example, in Section 2.2 we test the independence of several components of a normal distribution. The LRT statistic W_n in (2.9) can be written as the product of some independent random variables, say, V_i's with beta distributions; see, for example, Theorem 11.2.4 from Muirhead (1982). Therefore, it is possible that we can derive the CLT of W_n for general V_i's with the same means and variances as those of the beta distributions.

Finally, it is worthwhile to mention that some recent works consider similar problems under the nonparametric setting; see, for example, Cai, Liu and Xia (2013), Cai and Ma (2013), Chen, Zhang and Zhong (2010), Li and Chen (2012), Qiu and Chen (2012) and Xiao and Wu (2013).

5. Proofs. This section is divided into some subsections. In each of them we prove a theorem introduced in the Introduction. We first develop some tools. The following is some standard notation.

For two sequence of numbers $\{a_n; n \geq 1\}$ and $\{b_n; n \geq 1\}$, the notation $a_n = O(b_n)$ as $n \to \infty$ means that $\limsup_{n\to\infty} |a_n/b_n| < \infty$. The notation $a_n = o(b_n)$ as $n \to \infty$ means that $\lim_{n\to\infty} a_n/b_n = 0$. For two functions $f(x)$ and $g(x)$, the notation $f(x) = O(g(x))$ and $f(x) = o(g(x))$ as $x \to x_0 \in [-\infty, \infty]$ are similarly interpreted.

Throughout the paper $\Gamma(z)$ is the Gamma function defined on the complex plane $\mathbb{C}$.

5.1. *A preparation.*

LEMMA 5.1. *Let $b := b(x)$ be a real-valued function defined on $(0, \infty)$. Then*

$$\log \frac{\Gamma(x+b)}{\Gamma(x)} = b \log x + \frac{b^2 - b}{2x} + c(x)$$

as $x \to +\infty$, where

$$c(x) = \begin{cases} O\big(x^{-1/2}\big), & \text{if } b(x) = O(\sqrt{x}); \\ O\big(x^{-2}\big), & \text{if } b(x) = O(1). \end{cases}$$

Further, for any constants $d > c$, as $x \to +\infty$,

$$\sup_{c \le t \le d} \left| \log \frac{\Gamma(x+t)}{\Gamma(x)} - t \log x \right| \to 0.$$

PROOF. Recall the Stirling formula [see, e.g., page 368 from Gamelin (2001) or (37) on page 204 from Ahlfors (1979)],

$$\log \Gamma(x) = \left(x - \frac{1}{2}\right) \log x - x + \log \sqrt{2\pi} + \frac{1}{12x} + O\left(\frac{1}{x^3}\right) \tag{5.1}$$

as $x \to +\infty$. We have that

$$\begin{aligned} \log \frac{\Gamma(x+b)}{\Gamma(x)} &= (x+b) \log(x+b) - x \log x - b - \frac{1}{2}\big(\log(x+b) - \log x\big) \\ &\quad + \frac{1}{12}\left(\frac{1}{x+b} - \frac{1}{x}\right) + O\left(\frac{1}{x^3}\right) \end{aligned} \tag{5.2}$$

as $x \to +\infty$. First, use the fact that $\log(1+t) = t - (t^2/2) + O(t^3)$ as $t \to 0$ to get

$$\begin{aligned} (x+b) \log(x+b) - x \log x &= (x+b)\left(\log x + \log\left(1 + \frac{b}{x}\right)\right) - x \log x \\ &= (x+b)\left(\log x + \frac{b}{x} - \frac{b^2}{2x^2} + O\left(\frac{b^3}{x^3}\right)\right) - x \log x \\ &= b \log x + b + \frac{b^2}{2x} + O\left(\frac{b^3}{x^2}\right) - \frac{b^3}{2x^2} + O\left(\frac{b^4}{x^3}\right) \\ &= b \log x + b + \frac{b^2}{2x} + c_1(x) \end{aligned}$$

as $x \to +\infty$, where

$$c_1(x) = \begin{cases} O\big(x^{-1/2}\big), & \text{if } b(x) = O(\sqrt{x}); \\ O\big(x^{-2}\big), & \text{if } b(x) = O(1). \end{cases}$$

Similarly, as $x \to +\infty$,

$$\log(x+b) - \log x = \log\left(1 + \frac{b}{x}\right) = \begin{cases} \dfrac{b}{x} + O\big(x^{-1}\big), & \text{if } b(x) = O(\sqrt{x}); \\ \dfrac{b}{x} + O\big(x^{-2}\big), & \text{if } b(x) = O(1) \end{cases}$$

and

$$\frac{1}{x+b} - \frac{1}{x} = -\frac{b}{x(x+b)} = \begin{cases} O\big(x^{-3/2}\big), & \text{if } b(x) = O(\sqrt{x}); \\ O\big(x^{-2}\big), & \text{if } b(x) = O(1). \end{cases}$$

Substituting these two assertions in (5.2), we have

$$\log \frac{\Gamma(x+b)}{\Gamma(x)} = b \log x + \frac{b^2 - b}{2x} + c(x) \tag{5.3}$$

with

$$c(x) = \begin{cases} O\big(x^{-1/2}\big), & \text{if } b(x) = O(\sqrt{x}); \\ O\big(x^{-2}\big), & \text{if } b(x) = O(1) \end{cases}$$

as $x \to +\infty$.

For the last part, reviewing the whole proof above, we have from (5.3) that

$$\log \frac{\Gamma(x+t)}{\Gamma(x)} = t \log x + \frac{t^2 - t}{2x} + O\big(x^{-2}\big)$$

as $x \to +\infty$ uniformly for all $c \le t \le d$. This implies the conclusion. □

LEMMA 5.2. *Given $a > 0$, define*

$$\eta(t) = \sup_{x \ge a} \left| \log \frac{\Gamma(x+t)}{\Gamma(x)} - t \log x \right|$$

for all $t > -a$. Then $\lim_{t \to 0} \eta(t) = 0$.

PROOF. Let $d > c > 0$ be two constants. Since $\Gamma(x) > 0$ is continuous on $(0, \infty)$, then $g(x) := \log \Gamma(x)$ is uniformly continuous over the compact interval $[c/2, 2d]$. It then follows that

$$\sup_{c \le x \le d} \left| \log \frac{\Gamma(x+\varepsilon)}{\Gamma(x)} \right| = \sup_{c \le x \le d} \big| g(x+\varepsilon) - g(x) \big| \to 0 \tag{5.4}$$

as $\varepsilon \to 0$. On the other hand, by the second part of Lemma 5.1, for any $\varepsilon > 0$, there exists $x_0 > 2a$ such that

$$\sup_{|t| \le a} \left| \log \frac{\Gamma(x+t)}{\Gamma(x)} - t \log x \right| < \varepsilon$$

for all $x \geq x_0$. Therefore,

$$\sup_{x \geq x_0} \sup_{|t| \leq a} \left| \log \frac{\Gamma(x+t)}{\Gamma(x)} - t \log x \right| \leq \varepsilon.$$

Then,

$$\begin{aligned} \eta(t) &\leq \varepsilon + \sup_{a \leq x \leq x_0} \left| \log \frac{\Gamma(x+t)}{\Gamma(x)} - t \log x \right| \\ &\leq \varepsilon + \big(|\log a| + |\log x_0| \big) \cdot |t| + \sup_{a \leq x \leq x_0} \left| \log \frac{\Gamma(x+t)}{\Gamma(x)} \right| \end{aligned}$$

for all $|t| \leq a$. Consequently, we have from (5.4) that $\limsup_{t \to 0} \eta(t) \leq \varepsilon$ for all $\varepsilon > 0$, which completes the lemma. □

PROPOSITION 5.1 [Proposition 2.1 from Jiang, Jiang and Yang (2012)]. *Let $n > p = p_n$ and $r_n = (-\log(1 - \frac{p}{n}))^{1/2}$. Assume that $p/n \to y \in (0, 1]$ and $t = t_n = O(1/r_n)$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\log \prod_{i=n-p}^{n-1} \frac{\Gamma(i/2 - t)}{\Gamma(i/2)} = pt(1 + \log 2 - \log n) + r_n^2 \big(t^2 + (p - n + 1.5)t \big) + o(1).$$

LEMMA 5.3. *Let $n > p = p_n$ and $r_n = (-\log(1 - \frac{p}{n}))^{1/2}$. Assume $\frac{p}{n} \to y \in (0, 1]$ and $t = t_n = O(1/r_n)$ as $n \to \infty$. Then*

$$\log \left[\frac{\Gamma(n/2 + t)}{\Gamma(n/2)} \cdot \frac{\Gamma((n-p)/2)}{\Gamma((n-p)/2 + t)} \right] = r_n^2 t + o(1) \tag{5.5}$$

as $n \to \infty$.

PROOF. We prove the lemma by considering two cases.

Case (i): $y \in (0, 1)$. In this case, $n - p \to \infty$ and $\lim_{n \to \infty} r_n = (-\log(1 - y))^{1/2} \in (0, \infty)$, and hence $\{t_n\}$ is bounded. By Lemma 5.1,

$$\begin{aligned} \log \frac{\Gamma(n/2 + t)}{\Gamma(n/2)} &= t \log \frac{n}{2} + O\left(\frac{1}{n} \right), \\ \log \frac{\Gamma((n-p)/2)}{\Gamma((n-p)/2 + t)} &= -t \log (n-p)/2 + O\left(\frac{1}{n-p} \right) \end{aligned}$$

as $n \to \infty$. Add the two assertions up, and we get that the left-hand side of (5.5) is equal to

$$-t \log \left(1 - \frac{p}{n} \right) + o(1) = r_n^2 t + o(1) \tag{5.6}$$

as $n \to \infty$. So the lemma holds for $y \in (0, 1)$.

Case (ii): $y = 1$. In this case, $r_n \to +\infty$ and $t_n \to 0$ as $n \to \infty$. Recalling Lemma 5.2, we know that

$$\left|\log\frac{\Gamma((n-p)/2+t_n)}{\Gamma((n-p)/2)} - t_n \log\frac{n-p}{2}\right| \le \eta(t_n) \to 0$$

as $n \to \infty$ by taking $a = 1/2$ since $n - p \ge 1$. That is,

$$\log\frac{\Gamma((n-p)/2)}{\Gamma((n-p)/2+t_n)} = -t_n \log\frac{n-p}{2} + o(1) \tag{5.7}$$

as $n \to \infty$. By Lemma 5.1 and the fact that $\lim_{n\to\infty} t_n = 0$,

$$\log\frac{\Gamma(n/2+t_n)}{\Gamma(n/2)} = t_n \log\frac{n}{2} + o(1)$$

as $n \to \infty$. Adding up the above two terms, then using the same argument as in (5.6), we obtain (5.5). □

Define

$$\Gamma_p(z) := \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(z - \frac{1}{2}(i-1)\right) \tag{5.8}$$

for complex number z with $\mathrm{Re}(z) > \frac{1}{2}(p-1)$. See page 62 from Muirhead (1982).

LEMMA 5.4. *Let $\Gamma_p(z)$ be as in (5.8). Let $n > p = p_n$ and $r_n = (-\log(1 - \frac{p}{n}))^{1/2}$. Assume $\frac{p}{n} \to y \in (0, 1]$ and $s = s_n = O(1/r_n)$ and $t = t_n = O(1/r_n)$ as $n \to \infty$. Then*

$$\log\frac{\Gamma_p(n/2+t)}{\Gamma_p(n/2+s)} = p(t-s)(\log n - 1 - \log 2) + r_n^2\left[(t^2 - s^2) - \left(p - n + \frac{1}{2}\right)(t-s)\right] + o(1)$$

as $n \to \infty$.

PROOF. First,

$$\Gamma_p\left(\frac{n}{2}+t\right) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(\frac{n-i}{2} + t + \frac{1}{2}\right) = \pi^{p(p-1)/4} \prod_{j=n-p}^{n-1} \Gamma\left(\frac{j}{2} + t + \frac{1}{2}\right).$$

It follows that

$$\frac{\Gamma_p(n/2+t)}{\Gamma_p(n/2)} = \prod_{j=n-p}^{n-1} \frac{\Gamma(j/2+t+1/2)}{\Gamma(j/2+1/2)} = \prod_{j=n-p+1}^{n} \frac{\Gamma(j/2+t)}{\Gamma(j/2)}. \tag{5.9}$$

This implies

$$\frac{\Gamma_p(n/2+t)}{\Gamma_p(n/2)} = \frac{\Gamma(n/2+t)}{\Gamma(n/2)} \cdot \frac{\Gamma((n-p)/2)}{\Gamma((n-p)/2+t)} \cdot \prod_{j=n-p}^{n-1} \frac{\Gamma(j/2+t)}{\Gamma(j/2)}.$$

Now, replacing "t" in Proposition 5.1 with "$-t$" we then obtain

$$\log \prod_{j=n-p}^{n-1} \frac{\Gamma(j/2+t)}{\Gamma(j/2)} = pt(\log n - 1 - \log 2) + r_n^2\big(t^2 - (p-n+1.5)t\big) + o(1)$$

as $n \to \infty$. On the other hand, from Lemma 5.3,

$$\log\left[\frac{\Gamma(n/2+t)}{\Gamma(n/2)} \cdot \frac{\Gamma((n-p)/2)}{\Gamma((n-p)/2+t)}\right] = r_n^2 t + o(1)$$

as $n \to \infty$. Combining the last three equalities, we have

$$\log \frac{\Gamma_p(n/2+t)}{\Gamma_p(n/2)} = pt(\log n - 1 - \log 2) + r_n^2\left(t^2 - \left(p-n+\frac{1}{2}\right)t\right) + o(1)$$

as $n \to \infty$. Similarly,

$$\log \frac{\Gamma_p(n/2+s)}{\Gamma_p(n/2)} = ps(\log n - 1 - \log 2) + r_n^2\left[s^2 - \left(p-n+\frac{1}{2}\right)s\right] + o(1)$$

as $n \to \infty$. Taking the difference of the above two assertions, we obtain the desired conclusion. □

5.2. *Proof of Theorem 1.*

LEMMA 5.5 [Corollary 8.3.6 from Muirhead (1982)]. *Assume $n > p$. Let V_n be as in (2.3). Then, under H_0 in (2.1), we have*

$$E(V_n^h) = p^{ph} \frac{\Gamma(mp/2)}{\Gamma(mp/2+ph)} \cdot \frac{\Gamma_p(m/2+h)}{\Gamma_p(m/2)}$$

for $h > -\frac{1}{2}$ where $m = n-1$.

PROOF OF THEOREM 1. Recall that a sequence of random variables $\{Z_n; n \geq 1\}$ converges to Z in distribution as $n \to \infty$ if

$$\lim_{n\to\infty} Ee^{hZ_n} = Ee^{hZ} < \infty \tag{5.10}$$

for all $h \in (-h_0, h_0)$, where $h_0 > 0$ is a constant; see, for example, page 408 from Billingsley (1986). Thus, to prove the theorem, it suffices to show that there exists $\delta_0 > 0$ such that

$$E\exp\left\{\frac{\log V_n - \mu_n}{\sigma_n} s\right\} \to e^{s^2/2} \tag{5.11}$$

as $n \to \infty$ for all $|s| < \delta_0$.

Set $m = n - 1$ and $r_x := (-\log(1 - \frac{p}{x}))^{1/2}$ for $x > p$. By the fact that $x + \log(1 - x) < 0$ for all $x \in (0, 1)$, we know that $\sigma_n^2 > 0$ for all $n \geq 3$, and $\lim_{n\to\infty} \sigma_n^2 = -2y - 2\log(1 - y) > 0$ for $y \in (0, 1)$, and $\lim_{n\to\infty} \sigma_n^2 = +\infty$ for $y = 1$. Therefore,

$$\delta_0 := \inf\{\sigma_n; n \geq 3\} > 0.$$

Fix $|s| < \frac{\delta_0}{2}$. Set $t = t_n = \frac{s}{\sigma_n}$. Then $\{t_n; n \geq 3\}$ is bounded and $|t_n| < \frac{1}{2}$ for all $n \geq 3$. By Lemma 5.5,

$$Ee^{t \log V_n} = EV_n^t = p^{pt} \frac{\Gamma(mp/2)}{\Gamma(mp/2 + pt)} \cdot \frac{\Gamma_p(m/2 + t)}{\Gamma_p(m/2)} \tag{5.12}$$

for all $n \geq 3$. By Lemma 5.1 for the first case and the assumption $p/m \to y \in (0, 1]$,

$$\begin{aligned} \log \frac{\Gamma(mp/2)}{\Gamma(mp/2 + pt)} &= -\log \frac{\Gamma(mp/2 + pt)}{\Gamma(mp/2)} \\ &= -pt \log \frac{mp}{2} - \frac{p^2t^2 - pt}{mp} + O\left(\frac{1}{m}\right) \\ &= -pt \log \frac{mp}{2} - \frac{pt^2}{m} + O\left(\frac{1}{n}\right) \end{aligned} \tag{5.13}$$

as $n \to \infty$. Notice

$$\begin{aligned} & t^2 \cdot \left(-\log\left(1 - \frac{p}{m}\right)\right) \\ &= \frac{s^2}{\sigma_n^2} \cdot \left(-\log\left(1 - \frac{p}{m}\right)\right) \\ &\to \begin{cases} \frac{s^2}{2} \cdot \frac{\log(1 - y)}{y + \log(1 - y)}, & \text{if } y \in (0, 1); \\ \frac{s^2}{2}, & \text{if } y = 1 \end{cases} \end{aligned}$$

as $n \to \infty$. Thus, $t = O(1/r_m)$ as $n \to \infty$. By Lemma 5.4,

$$\log \frac{\Gamma_p(m/2 + t)}{\Gamma_p(m/2)} = pt(\log m - 1 - \log 2) + r_m^2\left(t^2 - \left(p - m + \frac{1}{2}\right)t\right) + o(1)$$

as $n \to \infty$. This together with (5.12) and (5.13) gives that

$$\begin{aligned} \log Ee^{t \log V_n} &= pt \log p + \log \frac{\Gamma(mp/2)}{\Gamma(mp/2 + pt)} + \log \frac{\Gamma_p(m/2 + t)}{\Gamma_p(m/2)} \\ &= pt \log p - pt \log \frac{mp}{2} - \frac{pt^2}{m} + pt(\log m - 1 - \log 2) \end{aligned}$$

$$+ r_m^2\left(t^2 - \left(p - m + \frac{1}{2}\right)t\right) + o(1)$$

$$= 2\left(r_m^2 - \frac{p}{m}\right)\frac{t^2}{2} + \left[-p + \left(n - p - \frac{3}{2}\right)r_m^2\right]t + o(1)$$

as $n \to \infty$. Reviewing the notation μ_n, σ_n and $t = t_n = \frac{s}{\sigma_n}$, the above indicates that

$$\log E \exp\left\{\frac{\log V_n}{\sigma_n} s\right\} = \log E e^{t \log V_n} = \frac{\sigma_n^2 t^2}{2} + \mu_n t + o(1) = \frac{s^2}{2} + \frac{\mu_n}{\sigma_n} s + o(1)$$

as $n \to \infty$ for all $|s| < \frac{\delta_0}{2}$. This implies (5.11). The proof is complete. □

5.3. *Proof of Theorem 2.*

LEMMA 5.6 [Theorem 11.2.3 from Muirhead (1982)]. *Let p and W_n be as in (2.9). Then, under H_0 in (2.8),*

$$E W_n^t = \frac{\Gamma_p(n/2 + t)}{\Gamma_p(n/2)} \cdot \prod_{i=1}^{k} \frac{\Gamma_{p_i}(n/2)}{\Gamma_{p_i}(n/2 + t)} \tag{5.14}$$

for any $t > -1/2$, where $\Gamma_p(z)$ is as in (5.8).

PROOF OF THEOREM 2. We need to prove

$$\frac{\log W_n - \mu_n}{\sigma_n} \text{ converges in distribution to } N(0, 1) \tag{5.15}$$

as $n \to \infty$, where

$$\mu_n = -r_n^2\left(p - n + \frac{1}{2}\right) + \sum_{i=1}^{k} r_{n,i}^2\left(p_i - n + \frac{1}{2}\right) \quad \text{and}$$

$$\sigma_n^2 = 2r_n^2 - 2\sum_{i=1}^{k} r_{n,i}^2.$$

First, since $n > p = p_1 + \cdots + p_k$ and $\lim_{n\to\infty} p_i/n = y_i$ for each $1 \le i \le k$, we know

$$\frac{p}{n} = \sum_{i=1}^{k} \frac{p_i}{n} \to \sum_{i=1}^{k} y_i := y \in (0, 1] \tag{5.16}$$

as $n \to \infty$. Second, it is known $\prod_{i=1}^{k}(1 - x_i) > 1 - \sum_{i=1}^{k} x_i$ for all $x_i \in (0, 1)$, $1 \le i \le k$, see, for example, page 60 from Hardy, Littlewood and Pólya (1988). Taking the logarithm on both sides and then taking $x_i = p_i/n$, we see that

$$\frac{1}{2}\sigma_n^2 = r_n^2 - \sum_{i=1}^{k} r_{n,i}^2 = \sum_{i=1}^{k} \log\left(1 - \frac{p_i}{n}\right) - \log\left(1 - \frac{p}{n}\right) > 0 \tag{5.17}$$

for all $n \geq 2$. Now, by the assumptions and (5.16), it is easy to see

$$\lim_{n\to\infty} \sigma_n^2 = \begin{cases} -2\log(1-y) + 2\sum_{i=1}^{k} \log(1-y_i), & \text{if } y < 1; \\ +\infty, & \text{if } y = 1. \end{cases}$$

By the same argument as in the last inequality in (5.17), we know the limit above is always positive. Reviewing that $p = p_n$ and $n > p$, we then have

$$\delta_0 := \inf\{\sigma_n; n \geq 2\} > 0.$$

Fix $|s| < \delta_0/2$. Set $t = t_n = s/\sigma_n$. Then $\{t_n; n \geq 2\}$ is a bounded sequence of numbers satisfying $|t_n| < 1/2$ for all $n \geq 2$. In particular, as $n \to \infty$, we have

$$t = t_n = O\left(\frac{1}{r_{n,i}}\right), \qquad 1 \leq i \leq k, \tag{5.18}$$

thanks to that $\lim_{p\to\infty} r_{n,i} = -\log(1-y_i) \in (0,\infty)$ for $1 \leq i \leq k$. On the other hand, notice

$$\sum_{i=1}^{k} r_{n,i}^2 = -\sum_{i=1}^{k} \log\left(1 - \frac{p_i}{n}\right) \to -\sum_{i=1}^{k} \log(1-y_i)$$

as $n \to \infty$. It follows from (5.16) that

$$\lim_{n\to\infty} \frac{r_n^2}{\sigma_n^2} = \begin{cases} \frac{1}{2} \cdot \dfrac{\log(1-y)}{\log(1-y) - \sum_{i=1}^{k} \log(1-y_i)}, & \text{if } y \in (0,1); \\ \frac{1}{2}, & \text{if } y = 1. \end{cases}$$

This implies that

$$t = \frac{s}{\sigma_n} = O\left(\frac{1}{r_n}\right) \tag{5.19}$$

as $n \to \infty$. From Lemma 5.6,

$$Ee^{t\log W_n} = EW_n^t = \frac{\Gamma_p(n/2+t)}{\Gamma_p(n/2)} \cdot \prod_{i=1}^{k} \frac{\Gamma_{p_i}(n/2)}{\Gamma_{p_i}(n/2+t)} \tag{5.20}$$

since $|t| = |t_n| < 1/2$. By Lemma 5.4 and (5.19),

$$\begin{aligned} &\log \frac{\Gamma_p(n/2+t)}{\Gamma_p(n/2)} \\ &\quad = pt(\log n - 1 - \log 2) + r_n^2\left[t^2 - \left(p - n + \frac{1}{2}\right)t\right] + o(1) \end{aligned} \tag{5.21}$$

as $n \to \infty$. Similarly, by Lemma 5.4 and (5.18),

$$\begin{aligned}(5.22)\qquad &\log \frac{\Gamma_{p_i}(n/2+t)}{\Gamma_{p_i}(n/2)} \\ &\quad = p_i t(\log n - 1 - \log 2) + r_{n,i}^2\left[t^2 - \left(p_i - n + \frac{1}{2}\right)t\right] + o(1)\end{aligned}$$

as $n \to \infty$ for $1 \le i \le k$. Therefore, use the identity $p = p_1 + \cdots + p_k$ to have

$$\begin{aligned}&\log \prod_{i=1}^{k} \frac{\Gamma_{p_i}(n/2+t)}{\Gamma_{p_i}(n/2)} \\ &\quad = \sum_{i=1}^{k} \log \frac{\Gamma_{p_i}(n/2+t)}{\Gamma_{p_i}(n/2)} \\ &\quad = pt(\log n - 1 - \log 2) + t^2 \sum_{i=1}^{k} r_{n,i}^2 - t \sum_{i=1}^{k} r_{n,i}^2 \left(p_i - n + \frac{1}{2}\right) + o(1)\end{aligned}$$

as $n \to \infty$. This together with (5.20) and (5.21) gives

$$\begin{aligned}\log E W_n^t &= t^2\left(r_n^2 - \sum_{i=1}^{k} r_{n,i}^2\right) + t\left[-r_n^2\left(p - n + \frac{1}{2}\right) + \sum_{i=1}^{k} r_{n,i}^2\left(p_i - n + \frac{1}{2}\right)\right] \\ &\quad + o(1) \\ &= \frac{s^2}{2} + \frac{\mu_n}{\sigma_n} s + o(1)\end{aligned}$$

as $n \to \infty$ by the definitions of μ_n and σ_n as well as the fact $t = s/\sigma_n$. We then arrive at

$$E \exp\left\{\frac{\log W_n - \mu_n}{\sigma_n} s\right\} = e^{-\mu_n s/\sigma_n} \cdot E W_n^t \to e^{s^2/2}$$

as $n \to \infty$ for all $|s| < \delta_0/2$. This implies (5.15) by using the moment generating function method stated in (5.10). □

5.4. *Proof of Theorem 3.* Consider

$$\lambda_n = \frac{\prod_{i=1}^{k} |\mathbf{B}_i|^{n_i/2}}{|\mathbf{A} + \mathbf{B}|^{n/2}}. \tag{5.23}$$

LEMMA 5.7 [Corollary 10.8.3 from Muirhead (1982)]. *Let $n_i > p$ for $i = 1, 2 \ldots, k$. Let λ_n be as in (5.23). Then, under H_0 in (2.11),*

$$E(\lambda_n^t) = \frac{\Gamma_p((1/2)(n-1))}{\Gamma_p((1/2)n(1+t) - 1/2)} \cdot \prod_{i=1}^{k} \frac{\Gamma_p((1/2)n_i(1+t) - 1/2)}{\Gamma_p((1/2)(n_i - 1))}$$

for all $t > \max_{1 \le i \le k}\{\frac{p}{n_i}\} - 1$, where $\Gamma_p(z)$ is as in (5.8).

The restriction $t > \max_{1\le i\le k}\{\frac{p}{n_i}\} - 1$ comes from the restriction in (5.8).

PROOF OF THEOREM 3. Review (2.12) and (5.23). Notice

$$\log \Lambda_n = \log \lambda_n + \frac{1}{2} pn \log n - \frac{1}{2}\sum_{i=1}^{k} pn_i \log n_i. \tag{5.24}$$

Evidently,

$$\frac{p}{n} = \frac{1}{\sum_{i=1}^{k} n_i/p} \to \frac{1}{\sum_{i=1}^{k} 1/y_i} := y \in (0,1) \tag{5.25}$$

as $p \to \infty$. As a consequence,

$$r_n^2 \to -\log(1-y) \in (0,\infty) \tag{5.26}$$

as $p \to \infty$.

Step 1. We show $\sigma_n^2 > 0$ for all $\min_{1\le i\le k} n_i \ge p+1$ and $p \ge 1$. In fact, let $h(x) = -\log(1-x)$ for $x \in [0,1)$. Then $h(x)$ is convex on $[0,1)$. Take $x_i = \frac{p}{n_i}$, $\lambda_i = \frac{n_i^2}{n^2}$ for $1 \le i \le k$, and $x_{k+1} = 0$ and $\lambda_{k+1} = 1 - \frac{\sum_{i=1}^{k} n_i^2}{n^2}$. Since $n_i' = n_i - 1$ and r_x is decreasing in $x > p$, by convexity,

$$\sum_{i=1}^{k} \frac{n_i^2}{n^2} r_{n_i'}^2 \ge \sum_{i=1}^{k} \frac{n_i^2}{n^2} r_{n_i}^2 = \sum_{i=1}^{k+1} \lambda_i h(x_i) > h\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = h\left(\frac{p}{n}\right) = r_n^2, \tag{5.27}$$

where ">," instead of "≥," comes from the fact that $h(x)$ is strictly convex and $x_1 \ne x_{k+1}$. This says that $\sigma_n^2 > 0$ for all $\min_{1\le i\le k} n_i \ge p+1$ and $p \ge 1$. Second, we claim

$$\begin{aligned} 2\sigma_n^2 &\to 2\sigma^2 \\ &= \begin{cases} \log(1-y) - \sum_{i=1}^{k} \dfrac{y^2}{y_i^2} \log(1-y_i) > 0, & \text{if } \max_{1\le i\le k} y_i < 1; \\ +\infty, & \text{if } \max_{1\le i\le k} y_i = 1 \end{cases} \end{aligned} \tag{5.28}$$

as $p \to \infty$. In fact, for the second case, noticing $y \in (0,1)$ by (5.25), the limit is obviously $+\infty$ since $\lim_{x\to 1-} -\log(1-x) = +\infty$. On the other hand, by (5.25), $\lim \frac{n_i}{n} = \frac{y}{y_i}$ for all $i = 1, 2, \ldots, k$. Thus, the statement for the case $\max_{1\le i\le k} y_i < 1$ in (5.28) follows. Moreover, replacing $\frac{p}{n}$ with y, $\frac{p}{n_i}$ with y_i and $\frac{n_i^2}{n^2}$ with $\frac{y^2}{y_i^2}$ in (5.27), respectively, we know that the first limit in (5.28) is positive.

Step 2. In this step we collect some facts that will be used later. Fix s such that $|s| < \frac{\sigma}{2y}$. Set $t = t_n = \frac{s}{n\sigma_n}$. We claim that

$$t > \max_{1\le i\le k}\left\{\frac{p}{n_i}\right\} - 1 \qquad \text{provided } p \text{ is sufficiently large;} \tag{5.29}$$

$$\left|\frac{nt-1}{2}\right| \vee \frac{1}{2} = O\left(\frac{1}{r_n}\right) \quad \text{and} \quad \frac{n_i t}{2} = O\left(\frac{1}{r_{n_i-1}}\right) \tag{5.30}$$

as $p \to \infty$ for $i = 1, 2 \ldots, k$.

First, the assumption $\min_{1\le i\le k} n_i > p + 1$ implies that

$$\max_{1\le i\le k}\left\{\frac{p}{n_i}\right\} - 1 < \frac{p}{p+1} - 1 = -\frac{1}{p+1}. \tag{5.31}$$

Further, for $|s| < \frac{\sigma}{2y}$, we know $s > -\frac{\sigma}{2y}$. Moreover, $-\frac{n}{p+1}\sigma_n \to -\frac{\sigma}{y}$ by (5.25) and (5.28). These imply that, as p is sufficiently large, $-\frac{n}{p+1}\sigma_n < s$, or $t = \frac{s}{n\sigma_n} > -\frac{1}{p+1}$. This together with (5.31) concludes (5.29).

Second, by (5.26),

$$\left|\frac{nt-1}{2}\right| \vee \frac{1}{2} = O(1+nt) = O\left(1 + \frac{1}{\sigma_n}\right) = O\left(\frac{1}{r_n}\right) \tag{5.32}$$

as $p \to \infty$ by (5.26) and (5.28). We obtain the first identity in (5.30). Moreover, noticing $n_i \le n$ and $t = s/(n\sigma_n)$ we have

$$\frac{n_i t}{2} = O(nt) = O\left(\frac{1}{\sigma_n}\right) \tag{5.33}$$

as $p \to \infty$. By the definition of σ_n^2, (5.26) and the fact that $\lim_{p\to\infty} \frac{n_i}{n} \in (0, \infty)$ again, we know $r_{n_i'}^2 = O(\sigma_n^2)$ as $p \to \infty$. Then $\frac{1}{\sigma_n} = O(\frac{1}{r_{n_i'}})$ as $p \to \infty$. This together with (5.33) gives that

$$\frac{n_i t}{2} = O\left(\frac{1}{r_{n_i'}}\right) \tag{5.34}$$

as $p \to \infty$. This completes the second identity in (5.30).

Step 3. To prove the theorem, it is enough to prove

$$E\exp\left\{\frac{\log \Lambda_n - \mu_n}{n\sigma_n} s\right\} \to e^{s^2/2} \tag{5.35}$$

as $p \to \infty$ for all $|s| < \frac{\sigma}{2y}$.

Recalling $t = \frac{s}{n\sigma_n}$, by Lemma 5.7 and (5.29) we have

$$Ee^{t\log\lambda_n} = \frac{\Gamma_p((1/2)(n-1))}{\Gamma_p((1/2)n(1+t) - 1/2)} \cdot \prod_{i=1}^{k} \frac{\Gamma_p((1/2)n_i(1+t) - 1/2)}{\Gamma_p((1/2)(n_i - 1))}. \tag{5.36}$$

Now, replacing t by $\frac{1}{2}(nt-1)$ and taking $s = -\frac{1}{2}$ in Lemma 5.4, by the first assertion in (5.30) we get

$$\begin{aligned} &\log \frac{\Gamma_p((1/2)(n-1))}{\Gamma_p((1/2)n(1+t) - 1/2)} \\ &= -\log \frac{\Gamma_p(n/2 + (nt-1)/2)}{\Gamma_p(n/2 - 1/2)} \end{aligned} \tag{5.37}$$

$$
\begin{aligned}
&= -\frac{npt}{2}(\log n - 1 - \log 2) - r_n^2\left[\frac{(nt-1)^2}{4} - \frac{1}{4} - \left(p - n + \frac{1}{2}\right)\frac{nt}{2}\right] \\
&\quad + o(1) \\
&= \frac{npt}{2}(1 + \log 2 - \log n) + \frac{r_n^2}{4}\left(-n^2t^2 + (2p - 2n + 3)nt\right) + o(1)
\end{aligned}
$$

as $p \to \infty$. Recall $n_i' = n_i - 1$ for $1 \le i \le k$. By Lemma 5.4 and the second identity in (5.30), we get

$$
\begin{aligned}
\log \frac{\Gamma_p((1/2)n_i(1+t) - 1/2)}{\Gamma_p((1/2)(n_i - 1))} &= \log \frac{\Gamma_p((1/2)n_i' + n_i t/2)}{\Gamma_p((1/2)n_i')} \\
&= \frac{n_i pt}{2}(\log n_i' - 1 - \log 2) \\
&\quad + \frac{r_{n_i'}^2}{4}\left(n_i^2 t^2 - (2p - 2n_i + 3)n_i t\right) + o(1)
\end{aligned}
$$

as $p \to \infty$. Therefore, this, (5.36) and (5.37) say that

$$
\begin{aligned}
\log E e^{t \log \lambda_n} &= \frac{t^2}{4}\left(\sum_{i=1}^k n_i^2 r_{n_i'}^2 - n^2 r_n^2\right) \\
&\quad + \frac{t}{4}\Bigg[2\left(\sum_{i=1}^k n_i \log n_i' - n \log n\right)p \\
&\qquad + n r_n^2 (2p - 2n + 3) - \sum_{i=1}^k n_i r_{n_i'}^2 (2p - 2n_i + 3)\Bigg] \\
&\quad + o(1)
\end{aligned}
$$

as $p \to \infty$. Combining with (5.24), we obtain that

$$
\begin{aligned}
\log E e^{t \log \Lambda_n} &= \frac{t^2}{4}\left(\sum_{i=1}^k n_i^2 r_{n_i'}^2 - n^2 r_n^2\right) \\
&\quad + \frac{t}{4}\Bigg[2\left(\sum_{i=1}^k n_i \log n_i' - n_i \log n_i\right)p + n r_n^2 (2p - 2n + 3) \\
&\qquad - \sum_{i=1}^k n_i r_{n_i'}^2 (2p - 2n_i + 3)\Bigg] \\
&\quad + o(1)
\end{aligned} \tag{5.38}
$$

as $p \to \infty$. Observe $n_i \log n_i' - n_i \log n_i = n_i \log(1 - \frac{1}{n_i})$ and

$$n_i \log\left(1 - \frac{1}{n_i}\right) = n_i\left(-\frac{1}{n_i} - \frac{1}{2}\frac{1}{n_i^2} + O\left(\frac{1}{n_i^3}\right)\right) = -1 - \frac{1}{2n_i} + O\left(\frac{1}{n_i^2}\right)$$

as $p \to \infty$. Thus

$$p\sum_{i=1}^{k}(n_i \log n_i' - n_i \log n_i) = -kp - \frac{1}{2}\sum_{i=1}^{k} y_i + o(1)$$

as $p \to \infty$. Joining this with (5.38), we arrive at

$$\begin{aligned}\log E e^{t \log \Lambda_n} &= \frac{t^2}{4}\left(\sum_{i=1}^{k} n_i^2 r_{n_i'}^2 - n^2 r_n^2\right) + \frac{t}{4}\Bigg[-2kp - \sum_{i=1}^{k} y_i + n r_n^2(2p - 2n + 3) \\ &\qquad - \sum_{i=1}^{k} n_i r_{n_i'}^2 (2p - 2n_i + 3)\Bigg] \\ &\quad + o(1)\end{aligned}$$

as $p \to \infty$. By the definitions of μ_n and σ_n, the above implies

$$\log E e^{t \log \Lambda_n} = \frac{n^2 t^2}{2}\sigma_n^2 + \mu_n t + o(1)$$

as $p \to \infty$, which is equivalent to

$$\log E e^{t(\log \Lambda_n - \mu_n)} = \frac{n^2 t^2}{2}\sigma_n^2 + o(1) = \frac{s^2}{2} + o(1)$$

as $p \to \infty$ for any $|s| < \frac{\sigma}{2y}$. This leads to (5.35) since $t = \frac{s}{n\sigma_n}$. □

5.5. *Proof of Theorem 4.* Let Λ_n^* be as in (2.17). Set

$$W_n = \frac{\prod_{i=1}^{k} |\mathbf{A}_i|^{(n_i-1)/2}}{|\mathbf{A}|^{(n-k)/2}} = \Lambda_n^* \cdot \frac{\prod_{i=1}^{k}(n_i - 1)^{(n_i-1)p/2}}{(n-k)^{(n-k)p/2}}. \tag{5.39}$$

We have the following result.

LEMMA 5.8 [Page 302 from Muirhead (1982)]. *Assume $n_i > p$ for $1 \le i \le k$. Under H_0 in (2.15),*

$$E\big[(W_n)^h\big] = \frac{\Gamma_p((1/2)(n-k))}{\Gamma_p((1/2)(n-k)(1+h))} \cdot \prod_{i=1}^{k} \frac{\Gamma_p((1/2)(n_i-1)(1+h))}{\Gamma_p((1/2)(n_i-1))} \tag{5.40}$$

for all $h > \max_{1\le i\le k} \frac{p-1}{n_i-1} - 1$, where $\Gamma_p(x)$ is defined as in (5.8).

The condition "$h > \max_{1\le i\le k} \frac{p-1}{n_i-1} - 1$" is imposed in the above lemma because, by the definition of (5.8), the following inequalities are needed:

$$\frac{(n-k)(1+h)}{2} > \frac{p-1}{2} \quad \text{and} \quad \frac{n-k}{2} > \frac{p-1}{2},$$
$$\frac{(n_i-1)(1+h)}{2} > \frac{p-1}{2} \quad \text{and} \quad \frac{(n_i-1)}{2} > \frac{p-1}{2}$$

for each $1 \le i \le k$. These are obviously satisfied if $n_i > p$ for $1 \le i \le k$ and $h > \max_{1\le i\le k} \frac{p-1}{n_i-1} - 1$ (noting that $n = \sum_{i=1}^k n_i > n_i$ for each i).

PROOF OF THEOREM 4. According to (5.39), write

$$\log \Lambda_n^* = \log W_n + \frac{(n-k)p}{2}\log(n-k) - \sum_{i=1}^k \frac{(n_i-1)p}{2}\log(n_i-1).$$

To prove the theorem, it is enough to show

$$\frac{\log W_n - \mu_n'}{(n-k)\sigma_n} \text{ converges in distribution to } N(0,1) \tag{5.41}$$

as $p \to \infty$, where

$$\mu_n' = \mu_n + \sum_{i=1}^k \frac{(n_i-1)p}{2}\log(n_i-1) - \frac{(n-k)p}{2}\log(n-k). \tag{5.42}$$

Equation (5.41) can be proved through the following three steps:

Step 1. Let

$$y = \frac{1}{\sum_{i=1}^k y_i^{-1}}. \tag{5.43}$$

Then, $y \in (0,1)$ and

$$\frac{p}{n-k} = \frac{p}{\sum_{i=1}^k (n_i-1)} \to \frac{1}{\sum_{i=1}^k y_i^{-1}} = y. \tag{5.44}$$

We first show $\sigma_n^2 > 0$. In fact, let $\eta(x) = -\log(1-x)$ for $x \in [0,1)$. Then $\eta(x)$ is strictly convex on $[0,1)$. Recall that $n = n_1 + \cdots + n_k$. Take $x_i = p/(n_i-1)$ and $\lambda_i = (n_i-1)^2/(n-k)^2$ for $1 \le i \le k$ and $x_{k+1} = 0$ and $\lambda_{k+1} = 1 - \sum_{i=1}^k \lambda_i$. Then, by the strict convexity of $\eta(x)$,

$$\begin{aligned} -\sum_{i=1}^k \left(\frac{n_i-1}{n-k}\right)^2 \log\left(1 - \frac{p}{n_i-1}\right) &= \sum_{i=1}^{k+1} \lambda_i \eta(x_i) > \eta\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) \\ &= \eta\left(\frac{p}{n-k}\right) = -\log\left(1 - \frac{p}{n-k}\right), \end{aligned} \tag{5.45}$$

where the ">" holds since $x_{k+1} \neq x_1$. This says that $\sigma_n^2 > 0$ for all $\min_{1\le i\le k} n_i > 1 + p$ and $p \ge 1$. Second, we claim

$$2\sigma_n^2 \to 2\sigma^2 = \begin{cases} \log(1-y) - \sum_{i=1}^{k} \dfrac{y^2}{y_i^2} \log(1-y_i) > 0, & \text{if } \max\limits_{1\le i\le k} y_i < 1; \\ +\infty, & \text{if } \max\limits_{1\le i\le k} y_i = 1 \end{cases} \tag{5.46}$$

as $p \to \infty$. In fact, the limit in (5.46) for the case $\max_{1\le i\le k} y_i < 1$ follows since $n_i/n \to y/y_i$ for all $i = 1, 2, \ldots, k$ and $-\log(1-x)$ is a continuous function for $x \in (0, 1)$. Moreover, replacing $(n_i - 1)^2/(n-k)^2$ with y^2/y_i^2, $-\log[1 - p/(n_i - 1)]$ with $-\log(1 - y_i)$, and $p/(n-k)$ with y in (5.45), respectively, we obtain $\sigma^2 > 0$ as $\max_{1\le i\le k} y_i < 1$. For the second case, we know that one of the y_i's is equal to 1 and $y \in (0, 1)$. Hence the limit is obviously $+\infty$.

Step 2. We will make some preparation for the key part in step 3. Fix s such that $|s| < \sigma/(2y)$. Set $t = t_n = s/[(n-k)\sigma_n]$. We claim that

$$t > \max_{1\le i\le k} \left\{ \frac{p-1}{n_i - 1} \right\} - 1 \qquad \text{as } p \text{ is sufficiently large,} \tag{5.47}$$

$$\frac{(n-k)t}{2} = O\left(\frac{1}{r_{n-k}}\right) \quad \text{and} \quad \frac{(n_i-1)t}{2} = O\left(\frac{1}{r_{n_i-1}}\right) \tag{5.48}$$

as $p \to \infty$ for $1 \le i \le k$, where $r_{n-k} = \{-\log(1 - \frac{p}{n-k})\}^{1/2}$ and $r_{n_i-1} = \{-\log(1 - \frac{p}{n_i-1})\}^{1/2}$. First, the assumption $\min_{1\le i\le k} n_i > 1 + p$ implies that

$$\max_{1\le i\le k} \left\{ \frac{p-1}{n_i - 1} \right\} - 1 < \frac{p-1}{p} - 1 = -\frac{1}{p}. \tag{5.49}$$

Further, since $|s| < \sigma/(2y)$, we see $s > -\sigma/(2y)$. Moreover, $-(n-k)\sigma_n/p \to -\sigma/y$ as $p \to \infty$ by (5.44) and (5.46). These imply that, as p is sufficiently large, $-(n-k)\sigma_n/p < s$, or $t = s/[(n-k)\sigma_n] > -1/p$. This together with (5.49) completes (5.47).

Second, since $t = t_n = s/[(n-k)\sigma_n]$, we know from (5.46) that $\{(n-k)t_n\}$ is bounded. Then the first assertion in (5.48) follows since $r_{n-k}^2 \to -\log(1-y) \in (0, \infty)$ as $p \to \infty$. Now, fix an $i \in \{1, \ldots, k\}$. Easily, by (5.46),

$$\frac{r_{n_i-1}^2}{\sigma_n^2} \to \begin{cases} \dfrac{-\log(1-y_i)}{\sigma^2}, & \text{if } \max\limits_{1\le j\le k} y_j < 1; \\ 0, & \text{if } \max\limits_{1\le j\le k} y_j = 1 \text{ and } y_i < 1 \end{cases} \tag{5.50}$$

as $p \to \infty$. Now assume $y_i = 1$ for some $1 \le i \le k$. By the definition of σ_n^2, we see that

$$2\sigma_n^2 \ge -r_{n-k}^2 + \left(\frac{n_i - 1}{n-k}\right)^2 r_{n_i-1}^2 \to +\infty$$

as $p \to \infty$. Therefore, use the facts that $r_{n-k}^2 \to -\log(1-y) \in (0,\infty)$ and $(n_i - 1)/(n-k) \to y/y_i \in (0,\infty)$ to have

$$\frac{(n_i-1)t}{2} = \frac{s}{2}\cdot\frac{n_i-1}{n-k}\cdot\frac{1}{\sigma_n} = O\biggl(\frac{1}{r_{n_i-1}}\biggr)$$

as $p \to \infty$. Combining this with (5.50) we see that

$$\frac{(n_i-1)t}{2} = O\biggl(\frac{1}{r_{n_i-1}}\biggr)$$

as $p \to \infty$ for any $y_i \in (0,1]$. This gives the second assertion in (5.48).

Step 3. To prove the theorem, from (5.10) and (5.41) it suffices to prove

$$E\exp\biggl\{\frac{\log W_n - \mu_n'}{(n-k)\sigma_n}s\biggr\} \to e^{s^2/2} \tag{5.51}$$

as $p \to \infty$ for all $|s| < \sigma/(2y)$. Recall $t = t_n = s/[(n-k)\sigma_n]$. By Lemma 5.8 and (5.47),

$$\log E\exp\biggl\{\frac{\log W_n}{(n-k)\sigma_n}s\biggr\} = \log E\bigl[W_n^t\bigr] = \log\frac{\Gamma_p[(1/2)(n-k)]}{\Gamma_p[(1/2)(n-k)(1+t)]} + \sum_{i=1}^{k}\log\frac{\Gamma_p[(1/2)(n_i-1)(1+t)]}{\Gamma_p[(1/2)(n_i-1)]}$$

as p is sufficiently large. Using Lemma 5.4 and the first assertion of (5.48), we obtain

$$\begin{aligned}
&\log\frac{\Gamma_p[(1/2)(n-k)]}{\Gamma_p[(1/2)(n-k)(1+t)]}\\
&\quad= \log\frac{\Gamma_p[(1/2)(n-k)]}{\Gamma_p[(1/2)(n-k)+(1/2)(n-k)t]}\\
&\quad= -\frac{(n-k)pt}{2}\bigl[\log(n-k) - 1 - \log 2\bigr]\\
&\qquad - r_{n-k}^2\biggl[\frac{(n-k)^2t^2}{4} - \frac{(p-n+k+0.5)(n-k)t}{2}\biggr] + o(1)
\end{aligned}$$

as $p \to \infty$. Similarly, by the Lemma 5.4 and the second assertion of (5.48), we have

$$\begin{aligned}
&\log\frac{\Gamma_p[(1/2)(n_i-1)(1+t)]}{\Gamma_p[(1/2)(n_i-1)]}\\
&\quad= \frac{(n_i-1)pt}{2}\bigl[\log(n_i-1) - 1 - \log 2\bigr]\\
&\qquad + r_{n_i-1}^2\biggl[\frac{(n_i-1)^2t^2}{4} - \frac{(p-n_i+1.5)(n_i-1)t}{2}\biggr] + o(1)
\end{aligned}$$

as $p \to \infty$. Take sum over all i to have

$$\begin{aligned}
&\sum_{i=1}^{k} \log \frac{\Gamma_p[(1/2)(n_i-1)(1+t)]}{\Gamma_p[(1/2)(n_i-1)]} \\
&\quad = \frac{pt}{2}\left[\sum_{i=1}^{k}(n_i-1)\log(n_i-1)\right] - \frac{(n-k)pt}{2}(1+\log 2) \\
&\qquad + \frac{t^2}{4}\left[\sum_{i=1}^{k}(n_i-1)^2 r_{n_i-1}^2\right] \\
&\qquad - \frac{t}{2}\sum_{i=1}^{k}(p-n_i+1.5)(n_i-1)r_{n_i-1}^2 + o(1)
\end{aligned}$$

as $p \to \infty$. Therefore,

$$\begin{aligned}
&\log E e^{t \log W_n} \\
&\quad = \frac{t^2}{4}\left[\sum_{i=1}^{k}(n_i-1)^2 r_{n_i-1}^2 - (n-k)^2 r_{n-k}^2\right] \\
&\qquad + \frac{t}{2}\left[(p-n+k+0.5)(n-k)r_{n-k}^2 - \sum_{i=1}^{k}(p-n_i+1.5)(n_i-1)r_{n_i-1}^2\right] \\
&\qquad + \frac{pt}{2}\left[\sum_{i=1}^{k}(n_i-1)\log(n_i-1) - (n-k)\log(n-k)\right] + o(1) \\
&\quad = \frac{t^2}{2}(n-k)^2\sigma_n^2 + \mu_n' t + o(1)
\end{aligned}$$

as $p \to \infty$ where μ_n' is as in (5.42). Since $t = t_n = s/[(n-k)\sigma_n]$, we know

$$\log E \exp\left\{\frac{\log W_n - \mu_n'}{(n-k)\sigma_n}s\right\} = \log E e^{t \log W_n} - \mu_n' t \to \frac{s^2}{2}$$

as $p \to \infty$ for all $|s| < \sigma/(2y)$. This leads to (5.51). □

5.6. *Proof of Theorem 5.*

LEMMA 5.9 [Theorems 8.5.1 and 8.5.2 and Corollary 8.5.4 from Muirhead (1982)]. *Assume $n > p$. Then the LRT statistic for testing H_0 in (2.19) given by*

$$\Lambda_n = \left(\frac{e}{n}\right)^{np/2} |\mathbf{A}|^{n/2} \cdot \exp\left\{-\frac{1}{2}\operatorname{tr}(\mathbf{A}) - \frac{1}{2}n\overline{\mathbf{x}}'\overline{\mathbf{x}}\right\}$$

is unbiased, where $\mathbf{A}$ *is as in* (2.20). *Further, assuming* H_0 *in* (2.19), *we have*

$$E(\Lambda_n^t) = \left(\frac{2e}{n}\right)^{npt/2} (1+t)^{-np(1+t)/2} \frac{\Gamma_p((n(1+t)-1)/2)}{\Gamma_p((n-1)/2)}$$

for any $t > \frac{p}{n} - 1$.

The range "$t > \frac{p}{n} - 1$" follows from the definition of $\Gamma_p(z)$ in (5.8).

PROOF OF THEOREM 5. First, since $\log(1-x) < -x$ for all $x < 1$, we know $\sigma_n^2 > 0$ for all $n \geq 3$. Now, by assumption, it is easy to see that

$$\lim_{n\to\infty} \sigma_n^2 = \begin{cases} -\frac{1}{2}[y + \log(1-y)], & \text{if } y < 1; \\ +\infty, & \text{if } y = 1. \end{cases} \tag{5.52}$$

Easily, the limit is always positive. Hence,

$$\delta_0 := \inf\{\sigma_n; n \geq 3\} > 0.$$

Fix a number h with $|h| < \delta_0$, then $h > -\delta_0 \geq -\sigma_n$ for all $n \geq 3$. It follows that

$$\frac{p}{n} - 1 \leq \frac{n-1}{n} - 1 = -\frac{1}{n} < \frac{h}{n\sigma_n}$$

for all $n \geq 3$. Set $t = t_n = \frac{h}{n\sigma_n}$. Then the above says that

$$t > \frac{p}{n} - 1 \tag{5.53}$$

for all $n \geq 3$. From (5.52) we know that $\{t_n; n \geq 3\}$ is bounded. By Lemma 5.9 and (5.53),

$$\begin{aligned} E e^{t \log \Lambda_n} &= E(\Lambda_n^t) \\ &= \left(\frac{2e}{n}\right)^{npt/2} (1+t)^{-np(1+t)/2} \frac{\Gamma_p((n(1+t)-1)/2)}{\Gamma_p((n-1)/2)} \end{aligned} \tag{5.54}$$

for all $n \geq 3$. To prove the theorem, we only need to show

$$E \exp\left\{\frac{\log \Lambda_n - \mu_n}{n\sigma_n} h\right\} \to e^{h^2/2} \tag{5.55}$$

as $n \to \infty$ for all h with $|h| < \delta_0$. Let $r_n = (-\log(1 - \frac{p}{n}))^{1/2}$ for all $n \geq 3$. From the definition of σ_n^2 and (5.52), it is evident that

$$\frac{nt}{2} = \frac{h}{2\sigma_n} = O\left(\frac{1}{r_{n-1}}\right) \tag{5.56}$$

as $n \to \infty$. Set $m = n - 1$. Take $s = 0$ and replace t with $\frac{nt}{2}$ in Lemma 5.4. We obtain from (5.56) that

$$\begin{aligned}
&\log \frac{\Gamma_p((n(1+t)-1)/2)}{\Gamma_p((n-1)/2)} \\
&\quad = \log \frac{\Gamma_p(m/2 + nt/2)}{\Gamma_p(m/2)} \\
&\quad = \frac{npt}{2}(\log m - 1 - \log 2) + r_m^2\left[\frac{n^2t^2}{4} - \frac{1}{2}\left(p - m + \frac{1}{2}\right)nt\right] + o(1) \\
&\quad = \frac{npt}{2}(\log m - 1 - \log 2) + \frac{n^2 r_m^2}{4}t^2 + \frac{1}{4}(2n - 2p - 3)nr_m^2 t + o(1)
\end{aligned} \tag{5.57}$$

as $n \to \infty$. Use the fact $\log(1+s) = s - \frac{s^2}{2} + o(s^3)$ as $s \to 0$ to have

$$(1+t)\log(1+t) = (1+t)\left(t - \frac{t^2}{2} + o(t^3)\right) = t + \frac{t^2}{2} + O(t^3)$$

as $n \to \infty$ since $\lim_{n\to\infty} t = \lim_{n\to\infty} t_n = 0$. It follows that

$$\log(1+t)^{-np(1+t)/2} = -\frac{1}{2}np(1+t)\log(1+t) = -\frac{1}{2}np\left(t + \frac{t^2}{2}\right) + O(npt^3)$$

as $n \to \infty$. Thus, by (5.54),

$$\begin{aligned}
&\log E(\Lambda_n^t) \\
&\quad = \frac{npt}{2}\log\frac{2e}{n} - \frac{1}{2}np\left(t + \frac{t^2}{2}\right) + \frac{npt}{2}(\log m - 1 - \log 2) \\
&\qquad + \frac{n^2 r_m^2}{4}t^2 + \frac{1}{4}(2n - 2p - 3)nr_m^2 t + O(npt^3) + o(1)
\end{aligned} \tag{5.58}$$

$$\begin{aligned}
&= \frac{1}{4}(n^2 r_m^2 - np)t^2 + \frac{npt}{2}\log\left(1 - \frac{1}{n}\right) + \frac{1}{4}[(2n - 2p - 3)nr_m^2 - 2np]t \\
&\qquad + o(1),
\end{aligned} \tag{5.59}$$

where the sum of the first and third terms in (5.58) gives the second term in (5.59), and $O(npt^3) = o(1)$ as $n \to \infty$ by the definition of t and (5.52). Now,

$$\begin{aligned}
\frac{1}{4}(n^2 r_m^2 - np)t^2 &= \frac{1}{4}n^2\left(r_m^2 - \frac{p}{m}\right)t^2 + \frac{1}{4}n^2\left(\frac{p}{m} - \frac{p}{n}\right)t^2 \\
&= \frac{1}{4}n^2\left(r_m^2 - \frac{p}{m}\right)t^2 + o(1)
\end{aligned}$$

since $\frac{1}{4}n^2(\frac{p}{m} - \frac{p}{n})t^2 = O(pt^2) = o(1)$ as $n \to \infty$ by the definition of t and (5.52). Also,

$$\frac{npt}{2}\log\left(1 - \frac{1}{n}\right) = \frac{npt}{2}\left(-\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right) = -\frac{pt}{2} + o(1)$$

as $n \to \infty$. Joining the above two assertions and (5.59), recalling the definitions of μ_n and σ_n, we get

$$\log E(\Lambda_n^t) = \frac{1}{2}n^2t^2 \cdot \frac{1}{2}\left(r_m^2 - \frac{p}{m}\right) + \frac{1}{4}[(2n-2p-3)nr_m^2 - 2(n+1)p]t + o(1)$$
$$= \frac{n^2\sigma_n^2t^2}{2} + \mu_n t + o(1) = \frac{h^2}{2} + \mu_n t + o(1)$$

as $n \to \infty$ for all h with $|h| < \delta_0$ since $t = t_n = \frac{h}{n\sigma_n}$. Therefore, we eventually conclude

$$\log E \exp\left\{\frac{\log \Lambda_n - \mu_n}{n\sigma_n}h\right\} = \log E(\Lambda_n^t) - \mu_n t \to \frac{h^2}{2}$$

as $n \to \infty$ for all $|h| < \delta_0$, which is (5.55). The proof is complete. □

5.7. *Proof of Theorem 6.*

LEMMA 5.10. *Let $\mathbf{R}_n$ be the correlation matrix with the density function as in (2.26). Assume $n - 5 \geq p \geq 2$. Then,*

$$E[|\mathbf{R}_n|^t] = \left[\frac{\Gamma((n-1)/2)}{\Gamma((n-1)/2+t)}\right]^p \cdot \frac{\Gamma_p((n-1)/2+t)}{\Gamma_p((n-1)/2)} \tag{5.60}$$

for all $t \geq -1$.

In the literature such as Muirhead (1982) and Wilks (1932), the above formula is only valid for integer $t \geq -1$. The above lemma says that it is actually true for all real numbers $t \geq -1$.

PROOF OF LEMMA 5.10. Recall (2.25), $\mathbf{R}_n$ is a $p \times p$ nonnegative definite matrix and each of its entries takes value in $[-1, 1]$, thus the determinant $|\mathbf{R}_n| \leq p!$. By (9) on page 150 from Muirhead (1982) or (48) on page 492 from Wilks (1932),

$$E[|\mathbf{R}_n|^k] = \left[\frac{\Gamma((n-1)/2)}{\Gamma((n-1)/2+k)}\right]^p \cdot \frac{\Gamma_p((1/2)(n-1)+k)}{\Gamma_p((1/2)(n-1))} \tag{5.61}$$

for any integer k such that $\frac{n-1}{2} + k > \frac{p-1}{2}$ by (5.8), which is equivalent to that $n - p > -2k$. By the given condition, $n - p \geq 5$. Thus, (5.61) holds for all $k \geq -2$. In particular, $E[|\mathbf{R}_n|^{-2}] < \infty$. Since $|\mathbf{R}_n|$ is bounded, this implies that

$$E[|\mathbf{R}_n|^{-1}] < \infty \quad \text{and} \quad E[|\mathbf{R}_n|^{-1}|\log|\mathbf{R}_n||] < \infty. \tag{5.62}$$

Now set $Z = -\log(|\mathbf{R}_n|/p!)$. Then $P(Z > 0) = 1$, $E(Ze^Z) < \infty$ by (5.62) and

$$h_1(z) := (p!)^{-(z-1)} \cdot E[|\mathbf{R}_n|^{z-1}] = Ee^{-(z-1)Z} \tag{5.63}$$

for all $\mathrm{Re}(z) \geq 0$. It is not difficult to check that $\frac{d}{dz}(Ee^{-(z-1)Z}) = -E[Ze^{-(z-1)Z}]$ for all $\mathrm{Re}(z) \geq 0$. Further, by (5.62) again, $|h_1(z)| \leq p! \cdot E[|\mathbf{R}_n|^{-1}|] < \infty$ on $\{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\}$. Therefore, $h_1(z)$ is analytic and bounded on $\{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\}$. Define

$$h_2(z) = (p!)^{-(z-1)} \cdot \left[\frac{\Gamma((n-1)/2)}{\Gamma((n-1)/2+z-1)}\right]^p \times \frac{\Gamma_p((n-1)/2+z-1)}{\Gamma_p((1/2)(n-1))} \tag{5.64}$$

for $\mathrm{Re}(z) \geq 0$. By the Carlson uniqueness theorem [see, e.g., Theorem 2.8.1 on page 110 from Andrews, Askey and Roy (1999)], if we know that $h_2(z)$ is also bounded and analytic on $\{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\}$, since $h_1(z) = h_2(z)$ for all $z = 0, 1, 2, \ldots,$ we obtain that $h_1(z) = h_2(z)$ for all $\mathrm{Re}(z) \geq 0$. This implies our desired conclusion. Thus, we only need to check that $h_2(z)$ is bounded and analytic on $\{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\}$. To do so, reviewing (5.8), it suffices to show

$$h_3(z) := \prod_{i=2}^{p} \frac{\Gamma((n-i)/2+z-1)}{\Gamma((n-1)/2+z-1)} \tag{5.65}$$

is bounded and analytic on $\{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\}$. Noticing $\frac{3}{2} \leq \frac{n-i}{2} - 1 \leq \frac{n-2}{2} - 1$ for all $2 \leq i \leq p$, to show that, it is enough to prove

$$h(z) := \frac{\Gamma(\alpha+z)}{\Gamma(\beta+z)} \tag{5.66}$$

is bounded and analytic on $\{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\}$ for all fixed $\beta > \alpha > 0$. This is confirmed by Lemma A.1 in the Appendix. □

PROOF OF THEOREM 6. First, since $\log(1-x) < -x$ for all $x < 1$, we know $\sigma_n^2 > 0$ for all $n \geq 3$ since $n - 1 > p \geq 1$ for all $n \geq 3$ by the assumption. Now, from the given condition, it is easy to see

$$\lim_{n\to\infty} \sigma_n^2 = \begin{cases} -2[y + \log(1-y)], & \text{if } y < 1; \\ +\infty, & \text{if } y = 1. \end{cases} \tag{5.67}$$

Trivially, the limit is always positive for $y \in (0, 1]$. Consequently,

$$\delta_0 := \inf\{\sigma_n; n \geq 3\} > 0.$$

To finish the proof, by (5.10) it is enough to show that

$$E \exp\left\{\frac{\log|\mathbf{R}_n| - \mu_n}{\sigma_n} s\right\} \to e^{s^2/2} \tag{5.68}$$

as $n \to \infty$ for all s such that $|s| < \delta_0/2$.

Fix s such that $|s| < \delta_0/2$. Set $t = s/\sigma_n$. Then $|t| < 1/2$ for all $n \geq 3$. Thus, by Lemma 5.1 for the second case,

$$\begin{aligned}\log\left[\frac{\Gamma((n-1)/2)}{\Gamma((n-1)/2+t)}\right]^p &= -p \cdot \log\frac{\Gamma((n-1)/2+t)}{\Gamma((n-1)/2)} \\ &= -p\left(t\log\frac{n-1}{2} + \frac{t^2-t}{n-1} + O\left(\frac{1}{n^2}\right)\right) \\ &= -pt\big[\log(n-1) - \log 2\big] - \frac{p}{n-1}\big(t^2-t\big) + o(1)\end{aligned} \tag{5.69}$$

as $n \to \infty$. Second, it is ease to see that

$$t^2 r_{n-1}^2 = \frac{r_{n-1}^2}{\sigma_n^2}s^2 \to \begin{cases} \dfrac{s^2}{2}\dfrac{\log(1-y)}{y+\log(1-y)}, & \text{if } y \in (0,1); \\ \dfrac{s^2}{2}, & \text{if } y = 1 \end{cases}$$

as $n \to \infty$, where $r_n = (-\log(1-\frac{p}{n}))^{1/2}$ for all $n \geq 3$. In particular, $t = O(1/r_{n-1})$ as $n \to \infty$. Therefore, by Lemma 5.4,

$$\begin{aligned}&\log\frac{\Gamma_p((n-1)/2+t)}{\Gamma_p((n-1)/2)} \\ &\quad = pt\big(\log(n-1) - 1 - \log 2\big) + r_{n-1}^2\left(t^2 - \left(p-n+\frac{3}{2}\right)t\right) + o(1)\end{aligned}$$

as $n \to \infty$. By the given condition, $\mathbf{R}_n$ has the density function as in (2.26). Therefore, from (5.60) and (5.69) we conclude that

$$\begin{aligned}\log E\big[|\mathbf{R}_n|^t\big] &= \left(r_{n-1}^2 - \frac{p}{n-1}\right)t^2 - \left[p - \frac{p}{n-1} + \left(p-n+\frac{3}{2}\right)r_{n-1}^2\right]t + o(1) \\ &= \frac{s^2}{2} + \mu_n t + o(1)\end{aligned}$$

as $n \to \infty$ since $t = s/\sigma_n$ and $\sigma_n^2 = 2(r_{n-1}^2 - \frac{p}{n-1})$. This implies that

$$E\exp\left\{\frac{\log|\mathbf{R}_n| - \mu_n}{\sigma_n}s\right\} = e^{-\mu_n t}E\big[|\mathbf{R}_n|^t\big] \to e^{s^2/2}$$

as $n \to \infty$ for any $|s| < \delta_0/2$. We get (5.68). □

APPENDIX

In this section we give a lemma on the complex analysis needed to prove Lemma 5.10.

LEMMA A.1. *Let $\Gamma(z)$ be the Gamma function defined on the complex plane $\mathbb{C}$. Let $\beta > \alpha > 0$ be two constants. Then*

$$h(z) := \frac{\Gamma(\alpha + z)}{\Gamma(\beta + z)}$$

is bounded and analytic on $\{z \in \mathbb{C}; \operatorname{Re}(z) \geq 0\}$.

PROOF. It is known that $\Gamma(z)$ is a meromorphic function, and all of its poles are simple poles at $z = 0, -1, -2, \ldots$ Also $\Gamma(z)$ has no zeros on the complex plane $\mathbb{C}$; see, for example, page 199 from Ahlfors (1979) or page 364 from Gamelin (2001). Thus, $h(z)$ is analytic for all $\operatorname{Re}(z) \geq 0$. On the other hand, by Euler's formula [see, e.g., page 199 from Ahlfors (1979) or page 363 from Gamelin (2001)],

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right) e^{-z/k} \tag{A.1}$$

for all $z \in \mathbb{C}$, where $\gamma = 0.5772, \ldots$ is the Euler constant. Hence,

$$h(z) = \frac{z + \beta}{z + \alpha} \cdot e^{\gamma(\beta - \alpha)} \cdot \prod_{k=1}^{\infty} \frac{k + z + \beta}{k + z + \alpha} \cdot e^{-(\beta - \alpha)/k} \tag{A.2}$$

for all $\operatorname{Re}(z) \geq 0$. Since $|k + z + \alpha| \geq k + \alpha$ for all $\operatorname{Re}(z) \geq 0$, we have

$$\left|\frac{k + z + \beta}{k + z + \alpha}\right| = \left|1 + \frac{\beta - \alpha}{k + z + \alpha}\right| \leq 1 + \frac{\beta - \alpha}{k + \alpha} \leq \exp\left\{\frac{\beta - \alpha}{k + \alpha}\right\}$$

for all $\operatorname{Re}(z) \geq 0$. Consequently,

$$\left|\prod_{k=1}^{\infty} \frac{k + z + \beta}{k + z + \alpha} \cdot e^{-(\beta - \alpha)/k}\right| \leq \exp\left\{-\sum_{k=1}^{\infty} \frac{(\beta - \alpha)\alpha}{k(k + \alpha)}\right\} \leq 1$$

for all $\operatorname{Re}(z) \geq 0$ since $\beta > \alpha > 0$. Obviously, $\frac{\alpha + z}{\beta + z}$ is bounded on $\{z \in \mathbb{C}; \operatorname{Re}(z) \geq 0\}$. By (A.2) and the first paragraph, we know $h(z)$ is bounded and analytic on $\{z \in \mathbb{C}; \operatorname{Re}(z) \geq 0\}$. □

Acknowledgments. We thank Drs. Danning Li and Xingyun Zeng and Professors Xue Ding, Feng Luo, Albert Marden, Yongcheng Qi and Yong Zhang very much for their help in discussing and checking the mathematical proofs in this paper. We also thank the referees and the Associate Editor for their valuable comments which improved the paper significantly.

REFERENCES

AHLFORS, L. V. (1979). *Complex Analysis*: *An Introduction to the Theory of Analytic Functions of One Complex Variable*, 3rd ed. McGraw-Hill, New York. MR0510197

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

ANDREWS, G. E., ASKEY, R. and ROY, R. (1999). *Special Functions*. *Encyclopedia of Mathematics and Its Applications* **71**. Cambridge Univ. Press, Cambridge. MR1688958

BAI, Z., JIANG, D., YAO, J.-F. and ZHENG, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann*. *Statist*. **37** 3822–3840. MR2572444

BARTLETT, M. S. (1937). Properties and sufficiency and statistical tests. *Proc*. *R*. *Soc*. *Lond*. *A* **160** 268–282.

BARTLETT, M. S. (1954). A note on the multipying factors for various χ^2 approximations. *J*. *R*. *Stat*. *Soc*. *Ser*. *B Stat*. *Methodol*. **16** 296–298. MR0069457

BILLINGSLEY, P. (1986). *Probability and Measure*, 2nd ed. Wiley, Hoboken.

BOX, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika* **36** 317–346. MR0033491

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

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

CAI, T., LIU, W. and XIA, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J*. *Amer*. *Statist*. *Assoc*. **108** 265–277.

CAI, T. and MA, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli*. To appear.

CHEN, S. X., ZHANG, L.-X. and ZHONG, P.-S. (2010). Tests for high-dimensional covariance matrices. *J*. *Amer*. *Statist*. *Assoc*. **105** 810–819. MR2724863

DAS GUPTA, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions. *Ann*. *Math*. *Statist*. **40** 697–701. MR0238434

DAVIS, A. W. (1971). Percentile approximations for a class of likelihood ratio criteria. *Biometrika* **58** 349–356. MR0324814

DONOHO, D. L. (2000). High-dimensional data analysis: The curses and blessings of dimensionality. Aide-Memoire of the lecture in AMS conference Math challenges of 21st Centrury. Available at http://www-stat.stanford.edu/~donoho/lectures.

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

GAMELIN, T. W. (2001). *Complex Analysis*, 1st ed. Springer, New York. MR1830078

GLESER, L. J. (1966). A note on the sphericity test. *Ann*. *Math*. *Statist*. **37** 464–467. MR0187329

HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1988). *Inequalities*. Cambridge Univ. Press, Cambridge. MR0944909

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

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

JIANG, D., JIANG, T. and YANG, F. (2012). Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. *J*. *Statist*. *Plann*. *Inference* **142** 2241–2256. MR2911842

JIANG, T. and MATSUMOTO, S. (2011). Moments of traces for circular beta-ensembles. Available at http://arxiv.org/pdf/1102.4123v1.pdf.

JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann*. *Statist*. **29** 295–327. MR1863961

LEDOIT, O. and WOLF, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30** 1081–1102. MR1926169

LI, J. and CHEN, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *Ann. Statist.* **40** 908–940. MR2985938

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

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

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

LIU, W.-D., LIN, Z. and SHAO, Q.-M. (2008). The asymptotic distribution and Berry–Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. *Ann. Appl. Probab.* **18** 2337–2366. MR2474539

MAUCHLY, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Statistics* **11** 204–209. MR0002084

MORRISON, D. F. (2004). *Multivariate Statistical Methods*, 4th ed. Cengage Learning, Stamford, CT.

MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York. MR0652932

NAGARSENKER, B. N. and PILLAI, K. C. S. (1973). The distribution of the sphericity test criterion. *J. Multivariate Anal.* **3** 226–235. MR0329126

ONATSKI, A., MOREIRA, M. J. and HALLIN, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Ann. Statist.* **41** 1204–1231.

PERLMAN, M. D. (1980). Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations. *Ann. Statist.* **8** 247–263. MR0560727

QIU, Y. and CHEN, S. X. (2012). Test for bandedness of high-dimensional covariance matrices and bandwidth estimation. *Ann. Statist.* **40** 1285–1314. MR3015026

SCHOTT, J. R. (2001). Some tests for the equality of covariance matrices. *J. Statist. Plann. Inference* **94** 25–36. MR1820169

SCHOTT, J. R. (2005). Testing for complete independence in high dimensions. *Biometrika* **92** 951–956. MR2234197

SCHOTT, J. R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Comput. Statist. Data Anal.* **51** 6535–6542. MR2408613

SRIVASTAVA, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *J. Japan Statist. Soc.* **35** 251–272. MR2328427

SUGIURA, N. and NAGAO, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *Ann. Math. Statist.* **39** 1686–1692. MR0237033

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

WILKS, S. S. (1935). On the independence of k sets of normally distributed statistical variables. *Econometrica* **3** 309–326.

XIAO, H. and WU, W. B. (2013). Asymptotic theory for maximum deviations of sample covariance matrix estimates. *Stochastic Process. Appl.* **123** 2899–2920. MR3054550

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
224 CHURCH STREET, S.E.
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: jiang040@umn.edu

BOSTON SCIENTIFIC
1 SCIMED PLACE
MAPLE GROVE, MINNESOTA 55311
USA
E-MAIL: yangf@bsci.com