Likelihood Ratio Tests for High-Dimensional Normal Distributions

Tiefeng Jiang[1] and Yongcheng Qi[2]
University of Minnesota

Abstract

In the paper by Jiang and Yang (2013), six classical Likelihood Ratio Test (LRT) statistics are studied under high-dimensional settings. Assuming that a random sample of size n is observed from a p-dimensional normal population, they derive the central limit theorems (CLTs) when $p/n \to y \in (0, 1]$, which are different from the classical chi-square limits as n goes to infinity while p remains fixed. In this paper, by developing a new tool, we prove that the above six CLTs hold in a more applicable setting: p goes to infinity and $p < n - c$ for some $1 \le c \le 4$. This is an almost sufficient and necessary condition for the CLTs. Simulations of histograms, comparisons on sizes and powers with those in the classical chi-square approximations and discussions are presented afterwards.

Keywords: Likelihood ratio test, central limit theorem, high-dimensional data, multivariate normal distribution, hypothesis test, covariance matrix, mean vector, multivariate Gamma function.
AMS 2000 Subject Classification: Primary 62H15; secondary 62H10.

---

[1]School of Statistics, University of Minnesota, 224 Church Street, S.E., MN55455, USA, jiang040@umn.edu.

The research of Tiefeng Jiang was supported in part by NSF Grant DMS-1208982.

[2]Department of Mathematics and Statistics, University of Minnesota Duluth, MN 55812, USA, yqi@d.umn.edu.

The research of Yongcheng Qi was supported in part by NSF Grant DMS-1005345.

1 Introduction

Analyzing data with large dimensionality of the population and large sample size is one of very active areas in mathematical sciences. This is particularly true in Statistics. If n points from a population with dimension p are sampled and they are put together, we then see an $n \times p$ matrix naturally. When the population is the multivariate normal distribution, the methodology of studying such data is elaborated in the field of Multivariate Analysis.

In the last decade, with the development and improvement of modern technologies such as the speed of computers, biology, Wall Street trading and weather forecast, the collected data have a common feature that both n and p are very large. Thus, renovating the old statistical methods and creating new methods are necessary. There are some recent literatures about these development. For example, in the field of multivariate analysis, Schott (2001, 2005, 2007), Ledoit and Wolf (2002), Bai et al. (2009), Chen et al. (2010), Jiang et al. (2012), and Jiang and Yang (2013) study the classical likelihood ratio tests when p is large. For literatures on large n and large p with other interests, see, for example, two book-length treatments by Serdobolskii (2000) and Fujikoshi et al. (2010).

In this paper we will investigate a problem asked by Jiang and Yang (2013). Our solutions become very applicable in practice. To make the problem clear to understand, let us take one example to illustrate. For a multivariate normal distribution $N_p(\mu, \mathbf{\Sigma})$, where $\mu \in \mathbb{R}^p$ is mean vector and $\mathbf{\Sigma}$ is $p \times p$ covariance matrix, consider the spherical test

$$H_0 : \mathbf{\Sigma} = \lambda \mathbf{I}_p \quad \text{vs} \quad H_a : \mathbf{\Sigma} \neq \lambda \mathbf{I}_p \tag{1.1}$$

with λ unspecified. Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be i.i.d. $\mathbb{R}^p$-valued random vectors with normal distribution $N_p(\mu, \mathbf{\Sigma})$. Let V_n be the likelihood ratio test (LRT) statistic (given in (2.2)). The traditional theory of the multivariate analysis says that $-n \log V_n$ goes to a chi-square distribution as n tends to infinity while p is fixed. Jiang and Yang (2013) prove that it is no longer true as $p \to \infty$. In fact, one of their results show that the central limit theorem (CLT) holds, that is, $(\log V_n - \mu_n)/\sigma_n$ actually converges to the standard normal distribution as $n \to \infty$ and $p/n \to y \in (0, 1]$, where μ_n and σ_n are explicit constants of n and p. Similar results on other five classical likelihood ratio tests are also obtained in their paper. By a comparison between the histograms of $(\log V_n - \mu_n)/\sigma_n$ and the standard normal curve, they observe that the above central limit theorem also holds even when p is large but not necessarily proportional to n, that is, the assumption $p/n \to y \in (0, 1]$ does not have to hold. See the comment in Problem 3 in Section 4 from their paper. Of course, if this is true the CLT will be very useful in practice since it is hard to judge whether p/n has a limit in $(0, 1]$ for real data.

We prove in this paper that the above CLT holds when p, which has to be less than n in the LRT, is large but not necessarily at the same scale of n. Other five classical likelihood ratio tests are also shown to have similar behaviors. Therefore, in the corresponding LRTs, the sizes of data are allowed to be more flexible. One does not need to concern if the value of p is large enough to be comparable of the sample size.

One of the key reasons to assume $p/n \to (0,1]$ in the previous studies relies on the fact that it is a very typical assumption in the field of Random Matrix Theory. In their enlightening work, Bai et al. (2009) study a test similar to (1.1) by using the central limit theorem of the eigenvalues of the Wishart matrices. Their work is based on the assumption $p/n \to y \in (0,1]$. The subsequent work aforementioned naturally use this condition. In this paper, we develop a machinery (Proposition 5.1) to deal with the case when p is much smaller than n. It is an expansion of the generalized Gamma function $\Gamma_p(z)$ (defined by (5.1)) which enables us to obtain the CLT as long as $p \to \infty$ regardless of the relative speed to n. Our starting step is the method of moment generating functions. When changing the point from the Random Matrix Theory to the method of the moment generating function, it is very interesting to see that the CLT actually holds for such a big range of n and p.

The difference between this work and Jiang and Yang (2013) is as follows. Jiang and Yang study six classical LRTs under the assumption $p/n \to y \in (0,1]$ or similar conditions. In this paper, we study the same six tests under the condition $n - c > p \to \infty$ for some $1 \le c \le 4$. So the earlier work is a special case of the current one. In fact, the assumption that "$n - c > p \to \infty$" is almost necessary: when p is finite the test statistic converges to a chi-square distribution by a classical LRT theorem; when $n - 1 < p$ the LRT does not exists (see Jiang and Yang (2013) for further details). Second, the new results are more applicable. Lastly, the derivation of our new tool of Proposition 5.1 is more challenging than Lemma 5.4 in Jiang and Yang (2013). Both are the core steps in the proofs appearing in the two papers. Readers are referred to their paper for more descriptions and narrations about the six tests.

The outline of the rest of this paper is given as follows. We present the six likelihood ratio tests in Sections 2.1 - 2.6. They are: (1) Testing covariance matrices of normal distributions proportional to identity matrix; (2) Testing independence of components of normal distributions; (3) Testing multiple normal distributions being identical; (4) Testing equality of several covariance matrices; (5) Testing specified values for mean vector and covariance matrix; (6) Testing complete independence of a normal distribution. The central limit theorems are presented in those sections. In Sections 3.1, we make pictures to compare the classical chi-square approximations and our CLTs. In 3.2, we give tables on sizes and powers of the tests. In Section 4, we provide a summary and lead some discussions. All of the theorems are proved in Section 5.

2 Main Results

Throughout the paper, $N_p(\mu, \mathbf{\Sigma})$ stands for the p-dimensional normal distribution with mean vector μ and covariance matrix $\mathbf{\Sigma}$, and $\mathbf{I}_p$ denotes the $p \times p$ identity matrix. For any given $\alpha \in (0,1)$, $\chi^2_{f,\alpha}$ denotes the α level critical value of χ^2_f, the chi-square random variable or the chi-square distribution with f degrees of freedom, and z_α denotes the α level critical value of the standard normal distribution $N(0,1)$. The notation $|\mathbf{A}|$ or $\det(\mathbf{A})$ stands for the determinant of the square matrix $\mathbf{A}$.

In this section we present the central limit theorems of six classical LRT statistics. The six theorems are stated in six subsections.

2.1 Testing Covariance Matrices of Normal Distributions Proportional to Identity Matrix

Consider a normal distribution $N_p(\mu, \mathbf{\Sigma})$. The spherical test is given by

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

with λ unspecified. Suppose $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are i.i.d. $\mathbb{R}^p$-valued random vectors with normal distribution $N_p(\mu, \mathbf{\Sigma})$. As usual, set

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

Mauchly (1940) shows that the likelihood ratio test statistic of (2.1) is as follows:

$$V_n = |\mathbf{S}| \cdot \Big(\frac{\mathrm{tr}(\mathbf{S})}{p}\Big)^{-p}. \tag{2.2}$$

In this paper, we have the following result about V_n.

THEOREM 1 *Let $p = p_n$ such that $n > p+1$ for all $n \geq 3$ and V_n be as in (2.2). Assume $\lim_{n\to\infty} p_n = \infty$, 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*

$$\mu_n = -p - \big(n - p - \frac{3}{2}\big)\log(1 - \frac{p}{n-1}) \quad \textit{and}$$
$$\sigma_n^2 = -2\Big[\frac{p}{n-1} + \log(1 - \frac{p}{n-1})\Big].$$

Jiang and Yang (2013) prove the above theorem for the special case $p/n \to y \in (0,1]$. Theorem 1 says that p/n does not need to have a limit. Further, if $\lim p/n = y$ exists, the theorem holds for $y = 0$. Theorem 1 holds as long as $p \to \infty$ and $p < n-1$. It is explained in Jiang and Yang (2013) that the LRT statistic does not exist if $p > n-1$.

If p is fixed, the classical chi-square approximation says that

$$-(n-1)\rho \log V_n \quad \text{converges to} \quad \chi^2_f \tag{2.3}$$

in distribution as $n \to \infty$, where

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

See, e.g., Muirhead (1982) or the summary in Jiang and Yang (2013).

Let $\alpha \in (0,1)$ be any given number. Recall that a likelihood ratio test of size-α rejects the null hypotheses if the likelihood ratio (or any of its monotone increasing functions) is smaller than a constant c_α which is chosen in such a way that the size or type I error is (approximately) equal to the given α. Therefore, the rejection region of likelihood ratio test of (2.1) is $V_n \le c_\alpha$. Based on the chi-square approximation in (2.3), an approximate size-α rejection region is $-(n-1)\rho\log V_n \ge \chi^2_{f,\alpha}$. Based on the normal approximation in Theorem 1, the rejection region is $(\log V_n - \mu_n)/\sigma_n \le -z_\alpha$.

2.2 Testing Independence of Components of Normal Distributions

For $k \ge 2$, let $p_1, \cdots, p_k$ be k positive integers. Denote $p = p_1 + \cdots + p_k$ and let

$$\mathbf{\Sigma} = (\mathbf{\Sigma}_{ij})_{1\le i,j\le k}$$

be a positive definite matrix, where $\mathbf{\Sigma}_{ij}$ is a $p_i \times p_j$ sub-matrix for all $1 \le i,j \le k$. Assume ξ_i is a p_i-variate normal random vector for each $1 \le i \le k$, and $(\xi_1', \cdots, \xi_k')'$ has the distribution $N_p(\mu, \mathbf{\Sigma})$. We are interested in testing the independence of $\xi_1, \cdots, \xi_k$, or equivalently testing

$$H_0 : \mathbf{\Sigma}_{ij} = \mathbf{0} \text{ for all } 1 \le i < j \le k \quad \text{vs} \quad H_a : H_0 \text{ is not true.} \tag{2.4}$$

Assume that $\mathbf{x}_1, \cdots, \mathbf{x}_N$ are i.i.d. from distribution $N_p(\mu, \mathbf{\Sigma})$. Set $n = N-1$. Define

$$\mathbf{A} = \sum_{i=1}^{n+1}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \quad \text{with } \bar{\mathbf{x}} = \frac{1}{n+1}\sum_{i=1}^{n+1}\mathbf{x}_i,$$

and partition it as follows

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

where $\mathbf{A}_{ij}$ is a $p_i \times p_j$ matrix. The likelihood ratio statistic for testing (2.4) is given by

$$\Lambda_n = \frac{|\mathbf{A}|^{(n+1)/2}}{\prod_{i=1}^k |\mathbf{A}_{ii}|^{(n+1)/2}} := (W_n)^{(n+1)/2}, \tag{2.5}$$

see Wilks (1935) or Theorem 11.2.1 from Muirhead (1982).

Since $\log \Lambda_n = \frac{n+1}{2}\log W_n$ from (2.5), it suffices to deal with the limiting distribution of $\log W_n$. Assume $p_i := p_{i,n}$ for all $1 \le i \le k$. We have the following result for $\log W_n$.

THEOREM 2 *Suppose that $n > p$ for all large n and there exists $\delta \in (0,1)$ satisfying $\delta \leq p_i/p_j \leq \delta^{-1}$ for all $1 \leq i,j \leq k$ and all large n. Recall W_n as defined in (2.5). If $\min_{1\leq i\leq k} p_i \to \infty$ as $n \to \infty$, then, under H_0 in (2.4), $(\log W_n - \mu_n)/\sigma_n$ converges in distribution to $N(0,1)$ where*

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

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 $1 \leq i \leq k$.

The assumption "$\delta \leq p_i/p_j \leq \delta^{-1}$ for all $1 \leq i,j \leq n$ and all n" requires that the sizes of the components p_i's are comparable. This rules out the unusual situation that some of the p_i's are much larger than the others. As is pointed out by Jiang and Yang (2013), the LRT fails if $p > N = n+1$ since the matrix $\mathbf{A}$ is not of full rank. Jiang and Yang (2013) prove the above theorem under condition that $\lim_{n\to\infty} p_i/n = y_i \in (0,1]$ for $1 \leq i \leq k$. When $p_1, p_2, \cdots, p_k$ are fixed as n goes to infinity, the classical LRT statistic of (2.4) has a chi-square limit:

$$-2\rho \log \Lambda_n \text{ converges to } \chi_f^2 \tag{2.6}$$

in distribution, where

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

see, e.g., Theorem 11.2.5 in Muirhead (1982).

Let $\alpha \in (0,1)$ be any given number. Based on the chi-square approximation in (2.6), the LRT rejects the null hypothesis in (2.4) if $-2\rho\log\Lambda_n \geq \chi^2_{f,\alpha}$. Based on the normal approximation in Theorem 2, the rejection region is $(\log W_n - \mu_n)/\sigma_n \leq -z_\alpha$.

2.3 Testing that Multiple Normal Distributions Are Identical

Consider normal distributions $N_p(\mu_i, \mathbf{\Sigma}_i)$, $i = 1, 2, \cdots, k$, where $k \geq 2$. We are interested in testing whether the k distributions are identical, that is,

$$H_0: \mu_1 = \cdots = \mu_k,\ \mathbf{\Sigma}_1 = \cdots = \mathbf{\Sigma}_k \ \text{ vs } H_a: H_0 \ \text{ is not true.} \tag{2.7}$$

Assume $\{\mathbf{y}_{ij};\ 1 \leq i \leq k,\ 1 \leq j \leq n_i\}$ are independent p-dimensional random vectors, and for each $i = 1, 2, \cdots, k$, $\{\mathbf{y}_{ij};\ 1 \leq j \leq n_i\}$ are i.i.d. from $N(\mu_i, \mathbf{\Sigma}_i)$. Define

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

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

where

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

The likelihood ratio test statistic for (2.7) is first derived by Wilks (1932) as follows

$$\Lambda_n = \frac{\prod_{i=1}^{k}|\mathbf{B}_i|^{n_i/2}}{|\mathbf{A}+\mathbf{B}|^{n/2}} \cdot \frac{n^{pn/2}}{\prod_{i=1}^{k} n_i^{pn_i/2}}. \tag{2.8}$$

See also Theorem 10.8.1 from Muirhead (1982). It is noted in Jiang and Yang (2013) that when $p \geq n_i$ for any $i = 1, \cdots, k$, the determinant of the matrix $\mathbf{B}_i$ is zero since $\mathbf{B}_i$ is not of full rank, and consequently, the likelihood ratio statistic Λ_n is zero. Thus, the condition $p < \min\{n_i;\, 1 \leq i \leq k\}$ is required to ensure that the LRT statistic for the test (2.7) is nondegenerate.

We have the following result for the limiting distribution of Λ_n defined in (2.8).

Theorem 3 *Let $n_i = n_i(p) > p+1$ for all p and there exists $\delta \in (0,1)$ such that $\delta < n_i/n_j < \delta^{-1}$ for all $1 \leq i, j \leq k$. Let Λ_n be as in (2.8). Then, under H_0 in (2.7),*

$$\frac{\log \Lambda_n - \mu_n}{n\sigma_n} \quad \textit{converges to } N(0,1)$$

in distribution as $p \to \infty$, where

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

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

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

If the dimension p is fixed and the null hypothesis in (2.7) is true, it follows from Theorem 10.8.4 in Muirhead (1982) that

$$-2\rho\log\Lambda_n \quad \text{converges to} \quad \chi_f^2 \tag{2.9}$$

in distribution as $\min_{1\leq i\leq k} n_i \to \infty$, where

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

When p grows with the same rate of n_i, namely, $\lim_{p\to\infty} p/n_i = y_i \in (0,1]$ for $1 \leq i \leq k$, the above theorem is proved by Jiang and Yang (2013). We should mention that μ_n in Theorem

3 of Jiang and Yang (2013) is defined slightly differently: the counterpart of $\sum_{i=1}^k y_i$ in their result is $\sum_{i=1}^k \frac{p}{n_i}$ in the above theorem. Note that in our Theorem 3 we do not assume the limits of $\frac{p}{n_i}$ exist. This substitution does not change the limiting distribution since both $\sum_{i=1}^k \frac{p}{n_i}$ and $\sum_{i=1}^k y_i$ are bounded by k, and therefore they are negligible compared with $n\sigma_n$, which converges to infinity as shown in the proof of Theorem 3.

Let $\alpha \in (0,1)$ be any given number. Based on the chi-square approximation in (2.9), the LRT rejects the null hypothesis in (2.7) if $-2\rho \log \Lambda_n \geq \chi^2_{f,\alpha}$. Based on our normal approximation, the rejection region is $(\log \Lambda_n - \mu_n)/(n\sigma_n) \leq -z_\alpha$.

2.4 Testing Equality of Several Covariance Matrices

Let $k \geq 2$ be an integer. Consider p-dimensional normal distributions $N_p(\mu_i, \mathbf{\Sigma}_i)$, $1 \leq i \leq k$, where μ_i and $\mathbf{\Sigma}_i$ are unknown. We are interested in testing

$$H_0 : \mathbf{\Sigma}_1 = \cdots = \mathbf{\Sigma}_k \quad \text{vs} \quad H_a : H_0 \text{ is not true.} \tag{2.10}$$

For $1 \leq i \leq k$, let $\mathbf{x}_{i1}, \cdots, \mathbf{x}_{in_i}$ be i.i.d. $N_p(\mu_i, \mathbf{\Sigma}_i)$-distributed random vectors. Define

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

and

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

The likelihood ratio test statistic for (2.10), derived in Wilks (1932), is given by

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

The test rejects the null hypothesis H_0 when $\Lambda_n \leq c_\alpha$, where c_α is selected such that the test has the significance level of α. The test statistic Λ_n is non-degenerate only if all determinants of $\mathbf{A}_i$ are nonzero, and hence the condition that $p < n_i$ for all $i = 1, \ldots, k$ is required. We are interested in the following modified likelihood ratio test statistic Λ_n^*, suggested by Bartlett (1937):

$$\Lambda_n^* = \frac{\prod_{i=1}^k (|\mathbf{A}_i|)^{(n_i-1)/2}}{(|\mathbf{A}|)^{(n-k)/2}} \cdot \frac{(n-k)^{(n-k)p/2}}{\prod_{i=1}^k (n_i - 1)^{(n_i-1)p/2}}. \tag{2.11}$$

This modified likelihood ratio test has been proved to be unbiased; see, e.g., Sugiura and Nagao (1968) and Perlman (1980). In this paper, we will prove the following CLT for $\log \Lambda_n^*$.

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 there exists* $\delta \in (0,1)$ *such that* $\delta < n_i/n_j < \delta^{-1}$ *for all* i, j. *Let* Λ_n^* *be as in (2.11). Then, under* H_0 *in (2.10),* $(\log \Lambda_n^* - \mu_n)/((n-k)\sigma_n)$ *converges weakly to* $N(0,1)$ *as* $p \to \infty$, *where*

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

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

The CLT for $\log \Lambda_n^*$ has also been studied by Bai et al. (2009), Jiang et al. (2012), and Jiang and Yang (2013). Jiang and Yang (2013) prove Theorem 4 under more restrictive condition that $p/n_i \to y_i \in (0,1]$ for $i = 1, \cdots, k$. When p is fixed, the classical chi-square approximation is obtained by Box (1949). Under the null hypothesis of (2.10), Box (1949) shows that

$$-2\rho \log \Lambda_n^* \ \text{ converges to } \ \chi_f^2 \tag{2.12}$$

in distribution as $\min_{1\le i\le k} n_i \to \infty$, where

$$f = \frac{1}{2}p(p+1)(k-1) \ \text{ and } \ \rho = 1 - \frac{2p^2+3p-1}{6(p+1)(k-1)(n-k)}\Big(\sum_{i=1}^{k}\frac{n-k}{n_i-1} - 1\Big).$$

Let $\alpha \in (0,1)$ be any given number. Based on the chi-square approximation, the LRT rejects the null hypothesis in (2.10) if $-2\rho\log\Lambda_n^* \ge \chi^2_{f,\alpha}$. Based on the normal approximation in Theorem 4, the rejection region is $(\log \Lambda_n^* - \mu_n)/((n-k)\sigma_n) \le -z_\alpha$.

2.5 Testing Specified Values for Mean Vector and Covariance Matrix

Consider a normal distribution $N_p(\mu, \mathbf{\Sigma})$, where $\mu \in \mathbb{R}^p$ is the mean vector and $\mathbf{\Sigma}$ is the $p \times p$ covariance matrix. Based on n i.i.d. random vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$ from the normal distribution $N_p(\mu, \mathbf{\Sigma})$, we test

$$H_0: \quad \mu = \mu_0 \text{ and } \mathbf{\Sigma} = \mathbf{\Sigma}_0 \ \text{ vs } \ H_a: H_0 \text{ is not true},$$

where μ_0 is a given vector in $\mathbb{R}^p$ and $\mathbf{\Sigma}_0$ is a given $p \times p$ non-singular matrix. Through the data transformation $\tilde{\mathbf{x}}_i = \mathbf{\Sigma}_0^{-1/2}(\mathbf{x}_i - \mu_0)$, the above hypothesis test is equivalent to the test

$$H_0: \quad \mu = \mathbf{0} \text{ and } \mathbf{\Sigma} = \mathbf{I}_p \ \text{ vs } \ H_a: H_0 \text{ is not true}. \tag{2.13}$$

Set

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i, \quad \mathbf{A} = \sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'. \tag{2.14}$$

The likelihood ratio test statistic of (2.13) is given by

$$\Lambda_n = \left(\frac{e}{n}\right)^{np/2}|\mathbf{A}|^{n/2}e^{-\mathrm{tr}(\mathbf{A})/2}e^{-n\bar{\mathbf{x}}'\bar{\mathbf{x}}/2}; \tag{2.15}$$

see, for example, Theorem 8.5.1 from Muirhead (1982). The condition that $p < n$ is required to ensure that Λ_n is non-degenerate. We have the following CLT for $\log \Lambda_n$.

Theorem 5 *Assume that $p := p_n$ such that $n > 1 + p$ for all $n \geq 3$ and $p \to \infty$ as n goes to infinity. Let Λ_n be defined as in (2.15). Then under the null hypothesis of* (2.13), $(\log \Lambda_n - \mu_n)/(n\sigma_n)$ *converges in distribution to $N(0,1)$ as $n \to \infty$, where*

$$\mu_n = -\frac{1}{4}\Big[n(2n-2p-3)\log(1-\frac{p}{n-1}) + 2(n+1)p\Big] \quad \textit{and}$$
$$\sigma_n^2 = -\frac{1}{2}\Big(\frac{p}{n-1} + \log(1-\frac{p}{n-1})\Big) > 0.$$

Jiang and Yang (2013) show Theorem 5 under a stronger condition, $p/n \to y \in (0,1]$. When p is fixed, it follows from Theorem 8.5.5 in Muirhead (1982) that as $n \to \infty$

$$-2\rho \log \Lambda_n \ \text{ converges to } \ \chi^2_f \tag{2.16}$$

under the null hypothesis of (2.13), where

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

Let $\alpha \in (0,1)$ be any given number. Based on the chi-square approximation in (2.16), the LRT rejects the null hypothesis in (2.13) if $-2\rho\log\Lambda_n \geq \chi^2_{f,\alpha}$. Based on the normal approximation in Theorem 5, the rejection region is $(\log \Lambda_n - \mu_n)/\sigma_n \leq -z_\alpha$.

2.6 Testing Complete Independence

Assume that a p-dimensional random vector $\mathbf{x} = (x_1, \cdots, x_p)'$ has a distribution $N_p(\mu, \mathbf{\Sigma})$. We are interested in testing that the p components $x_1, x_2, \cdots, x_p$ are independent or equivalently testing that the covariance matrix $\mathbf{\Sigma}$ is diagonal. Let $\mathbf{R} = (r_{ij})_{p\times p}$ be the correlation matrix generated from $N_p(\mu, \mathbf{\Sigma})$. Then the test can be written as

$$H_0 : \mathbf{R} = \mathbf{I}_p \ \text{ vs } \ H_a : \mathbf{R} \neq \mathbf{I}_p. \tag{2.17}$$

To obtain the asymptotic distribution for the test statistic of the LRT for (2.17), we will consider a larger class of distributions namely, *spherical distributions*. Recall that a random vector $\mathbf{y} \in \mathbb{R}^n$ has a *spherical distribution* if $\mathbf{O}\mathbf{y}$ and $\mathbf{y}$ have the same probability distribution for all $n \times n$ orthogonal matrix $\mathbf{O}$. Obviously, any n-dimensional normal random vector with distribution $N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$ has a *spherical distribution* for any $\sigma > 0$.

Let $\mathbf{X} = (x_{ij})_{n\times p} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)' = (\mathbf{y}_1, \cdots, \mathbf{y}_p)$ be an $n \times p$ matrix such that $\mathbf{y}_1, \cdots, \mathbf{y}_p$ are independent random vectors with n-variate spherical distributions and $P(\mathbf{y}_i = \mathbf{0}) = 0$ for all $1 \le i \le p$ (these distributions may be different). For $1 \le i, j \le p$, let $\hat{r}_{ij}$ denote the Pearson correlation coefficient between $(x_{1i}, \cdots, x_{ni})$ and $(x_{1j}, \cdots, x_{nj})$, given by

$$\hat{r}_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \cdot \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \tag{2.18}$$

where $\bar{x}_i = \frac{1}{n}\sum_{k=1}^n x_{ki}$ and $\bar{x}_j = \frac{1}{n}\sum_{k=1}^n x_{kj}$. Then

$$\hat{\mathbf{R}}_n := (\hat{r}_{ij})_{p\times p} \tag{2.19}$$

is the sample correlation matrix based on the p-dimensional random vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$. A natural requirement for non-singularity of $\hat{\mathbf{R}}_n$ is $n > p$. From Theorem 5.1.3 in Muirhead (1982), the density of $|\hat{\mathbf{R}}_n|$ is given by

$$\text{Constant} \cdot |\mathbf{R}_n|^{(n-p-2)/2} d\mathbf{R}_n. \tag{2.20}$$

We first present the limiting distribution concerning the determinant of $\hat{\mathbf{R}}_n$.

Theorem 6 *Let $p = p_n$ satisfy that $n > p + 4$ and $p \to \infty$. Let $\mathbf{X} = (\mathbf{y}_1, \cdots, \mathbf{y}_p)$ be an $n \times p$ matrix such that $\mathbf{y}_1, \cdots, \mathbf{y}_p$ 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). Then $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n$ converges in distribution to $N(0, 1)$ as $n \to \infty$, where*

$$\mu_n = (p - n + \frac{3}{2})\log(1 - \frac{p}{n-1}) - \frac{n-2}{n-1}p,$$
$$\sigma_n^2 = -2\Big[\frac{p}{n-1} + \log\Big(1 - \frac{p}{n-1}\Big)\Big].$$

The theorem is proved by Jiang and Yang (2013) under the condition that $n > p + 4$ and $p/n \to y \in (0, 1]$.

Now we return to the LRT of (2.17). Assume 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 a correlation matrix $\mathbf{R}$. Write $\mathbf{x}_i = (x_{i1}, \cdots, x_{ip})'$ for $1 \le i \le n$, and define the sample correlation matrix $\hat{\mathbf{R}}_n = (\hat{r}_{ij})_{p\times p}$ as in (2.18). From Morrison (2004), page 40, the rejection region of the likelihood ratio test for (2.17) is

$$|\hat{\mathbf{R}}_n|^{n/2} \le c_\alpha$$

where c_α is determined so that the test has significance level of α. To derive the asymptotic distribution of $\log|\hat{\mathbf{R}}_n|$, write $\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)' = (x_{ij})_{n\times p} = (\mathbf{y}_1, \cdots, \mathbf{y}_p)$. Then, under the null hypothesis of (2.17), $\mathbf{y}_1, \cdots, \mathbf{y}_p$ are independent random vectors from n-variate normal distributions (these normal distributions may differ by their covariance matrices), and $P(\mathbf{y}_i = 0) = 0$ for all $1 \le i \le p$. Therefore, we have the following corollary to Theorem 6.

COROLLARY 1 *Assume that $p := p_n$ satisfy that $n > p+4$ and $p \to \infty$. Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be i.i.d. from $N_p(\mu, \mathbf{\Sigma})$ with the Pearson sample correlation matrix $\hat{\mathbf{R}}_n$ as defined in (2.19). Then, under H_0 in (2.17), $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n$ converges in distribution to $N(0,1)$ as $n \to \infty$, where*

$$\mu_n = \Big(p - n + \frac{3}{2}\Big)\log\Big(1 - \frac{p}{n-1}\Big) - \frac{n-2}{n-1}p\,;$$
$$\sigma_n^2 = -2\left[\frac{p}{n-1} + \log\Big(1 - \frac{p}{n-1}\Big)\right] > 0.$$

When p is fixed, the following chi-square approximation holds under the null hypothesis of test (2.17): as $n \to \infty$,

$$-\Big(n - 1 - \frac{2p+5}{6}\Big)\log|\hat{\mathbf{R}}_n| \text{ converges to } \chi^2_{p(p-1)/2} \tag{2.21}$$

in distribution. See, e.g., Bartlett (1954) or p. 40 from Morrison (2005) for this.

Based on the chi-square approximation in (2.21), the LRT rejects the null hypothesis in (2.17) if $-\Big(n-1-\frac{2p+5}{6}\Big)\log|\hat{\mathbf{R}}_n| \ge \chi^2_{p(p-1)/2,\alpha}$. According to Corollary 1, the rejection region based on the normal approximation is $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n \le -z_\alpha$.

3 Simulation Study

In this section, we compare the performance of the chi-square approximation and the normal approximation for all six likelihood ratio tests in Sections 2.1-2.6 through a finite sample simulation study. We plot the histograms for the six chi-square statistics which are used for the chi-square approximations specified in (2.3), (2.6), (2.9), (2.12), (2.16), and (2.21), and compare with their corresponding limiting chi-square curves. Similarly, we plot the histograms of the six statistics which are used for the normal approximations given in Theorems 1-6, and compare with the standard normal curve. We also report estimated sizes and powers for the six LRTs based on their chi-square approximations and the normal approximations. All simulations have been done by using software **R**, and the histograms, estimates of the sizes and powers are based on 10,000 replicates.

3.1 Comparison of Histograms

Six figures, Figures 1 to 6, are reported, and each of them corresponds to the two approximation methods stated in one of Sections 2.1-2.6. These figures are self-evident: when the sample sizes are small, the classical chi-square approximations are good; Our central limit theorems outperform the chi-square approximations when the sample sizes are large. Even the data dimensions are large but are small relative to the sample sizes, the fits are still quite well. These simulations are consistent with Theorems 1-6.

3.2 Simulation Study: Sizes and Powers

In this part, for each of the six LRTs treated earlier, we simulate the sizes and the powers for the normal approximation and for the classical chi-square approximation, and the simulation results are listed in six tables, Tables 1–6, which are self-explained. The notation $\mathbf{J}_p$ stands for the $p \times p$ matrix whose entries are all equal to 1 and $\lfloor x \rfloor$ stands for the integer part of $x > 0$.

From the six tables, we can see that when p is small, the chi-square approximation works better. A common feature is that for very small values of p, the LRTs based on the normal approximation have a slightly larger sizes than the nominal level 0.05. With the increase of p, the sizes of the LRTs based on the normal approximation method are very close to 0.05 while the sizes of the chi-square approximations are significantly higher than 0.05. We see that the sizes for the normal approximation are quite stable over the different choices of p.

For comparison of the powers reported in the six tables, we note that the larger the estimated sizes, the larger the corresponding estimated powers. When p is relatively large, the powers for the chi-square approximation are larger than those for the normal approximation, however, the sizes from the chi-square approximation are seriously higher than the nominal level 0.05. To understand this phenomenon, one should be aware of the fact that the two approximation methods use the same test statistics but they result in different cutoff values for rejection regions. For illustrating purpose, we can look at the spherical test in Section 2.1 with the LRT statistic V_n defined in (2.2). From the last paragraph in Section 2.1, the rejection region is $V_n \leq c_\alpha$, where $\alpha \in (0, 1)$ is the type I error. Since c_α is unknown, the actual cutoffs used to approximate c_α are $c_{\alpha,1} = \exp(-\chi^2_{f,\alpha}/(n-1)\rho)$ from the chi-square approximation and $c_{\alpha,2} = \exp(\mu_n - z_\alpha \sigma_n)$ from the normal approximation, which result in two different rejection regions $\{V_n \leq c_{\alpha,1}\}$ and $\{V_n \leq c_{\alpha,2}\}$, respectively. The two rejection regions are nested, that is, one is a subset of the other. Therefore, the larger rejection region has a larger size and a larger power. In other words, the larger powers for the chi-square approximation come from the sacrifice of the accuracy in type I errors or sizes of the test. The same relation holds true for other five LRTs. This explains what we have observed in the six tables for the powers.

In what follows, we provide more explanations on the simulation results for the sizes in the six tables.

(1) Table 1. We consider the spherical test $H_0 : \mathbf{\Sigma} = \lambda \mathbf{I}_p$ with λ unspecified, as given in (2.1). With a fixed sample size $n = 100$, we choose $p = 5, 20, 40$ and 60 for the values of the dimension p. When p is small (5 and 20), the chi-square and the normal approximation methods are comparable but the chi-square approximation is slightly better than the normal approximation in terms of the accuracy in the size of the test. When $p = 60$, the size for the chi-square approximation is 0.3342, much larger than the nominal level 0.05, while the size for the normal approximation is 0.0570.

(2) Table 2. This table reports the comparison results for the sizes and powers of the two tests for the hypotheses given in (2.4). We choose $k = 3$ for the simulation study, that is, a normal random vector is divided into three sub-vectors with dimensions p_1, p_2 and p_3, where p_1, p_2 and p_3 are specified in the table, and we test the independence of the three sub-vectors. For small p_i's the chi-square approximation performs better than the normal approximation, but the normal approximation seems not too bad with size 0.0659 compared with the nominal level 0.05 even when p_i's are as small as $p_1 = 2$, $p_2 = 2$ and $p_3 = 1$. The normal approximation improves as p_i's grow, and eventually the chi-square approximation yields a much larger size than the nominal level.

(3) Table 3. This table is concerning the simulation on testing the equality of k p-dimensional normal distributions as given in (2.7). We consider $k = 3$ normal populations, and sample sizes are fixed as $n_1 = n_2 = n_3 = 100$. Then four cases when $p = 5, 20, 40$ and 60 are investigated in the simulation, and the distribution under null hypothesis used in the simulation is $N_p(\mathbf{0}, \mathbf{I}_p)$. Although the size at $p = 5$ is 0.0567, a little bit larger than the nominal level, the size for the normal approximation is quite stable in general. The size of the test based the chi-square approximation is reasonably close to the nominal level only for very small p.

(4) Table 4. For hypotheses in (2.10), i.e., testing the equality of k covariance matrices of normal random vectors, we report the sizes and powers of the two tests. We choose $k = 3$, and fix the sizes of the samples for the three populations as $n_1 = n_2 = n_3 = 100$ with 4 different dimension choices $p = 5, 20, 40, 60$. The table shows that the sizes for the normal approximation are very close to the nominal level 0.05 except the case when $p = 5$. The chi-square approximation works well only for $p = 5$, and its size grows drastically fast as p increases. For example, the size for the chi-square approximation at $p = 60$ is as high as 0.9126, that is, about 91% of time, the test rejects the true null hypotheses.

(5) Table 5. Considering the test of hypothesis that the underlying distribution is a specific multivariate normal distribution or equivalently the test described in (2.13), the table compares the performance of the two different approximation methods. We fix the sample size $n = 100$ and consider four different values for the dimension, $p = 5, 20, 40$

and 60. For small p, the chi-square approximation works pretty well in terms of accuracy in the size but it becomes worse as p gets larger. The normal approximation works very well for all reasonably large p.

(6) Table 6. This table is about the test of independence of all components from a normal random vector, see (2.17) or equivalently, the covariance matrix is diagonal. In the simulation, the sample size is chosen as $n = 100$, and the dimension p has four choices, $p = 5, 20, 40$ and 60. From the table, all four sizes from the normal approximation are close to 0.05, while the chi-square approximation results in a reasonable size only for $p = 5$ or 20.

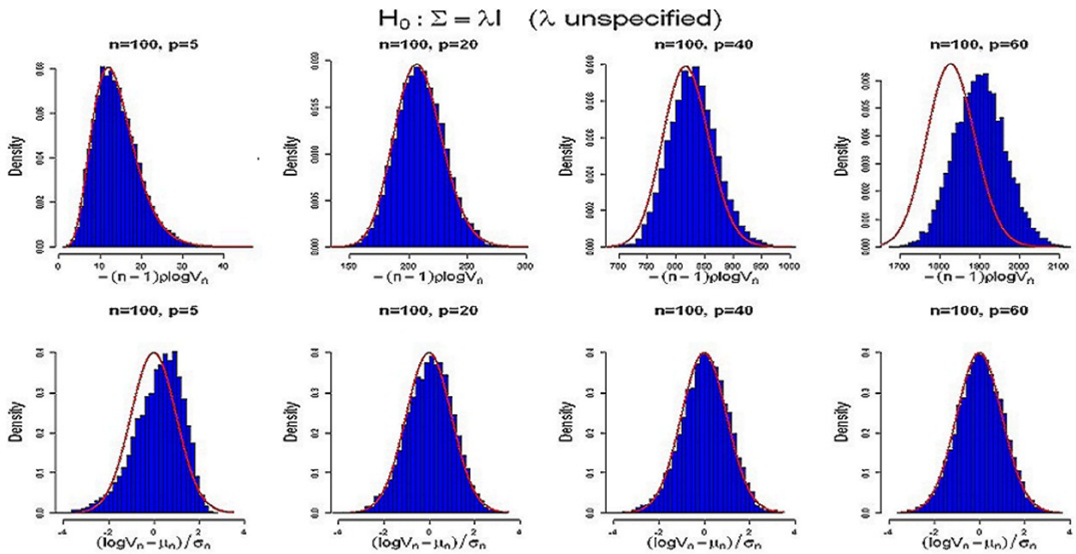

Figure 1: *Comparison between Theorem 1 and (2.3). We choose* $n = 100$ *with* $p = 5, 20, 40, 60$*. The pictures in the top row show that the* χ^2 *curves stay away farther gradually from the histogram of* $-(n-1)\rho \log V_n$ *when* p *grows. The bottom row shows that the* $N(0,1)$*-curve fits the histogram of* $(\log V_n - \mu_n)/\sigma_n$ *better as* p *becomes larger.*

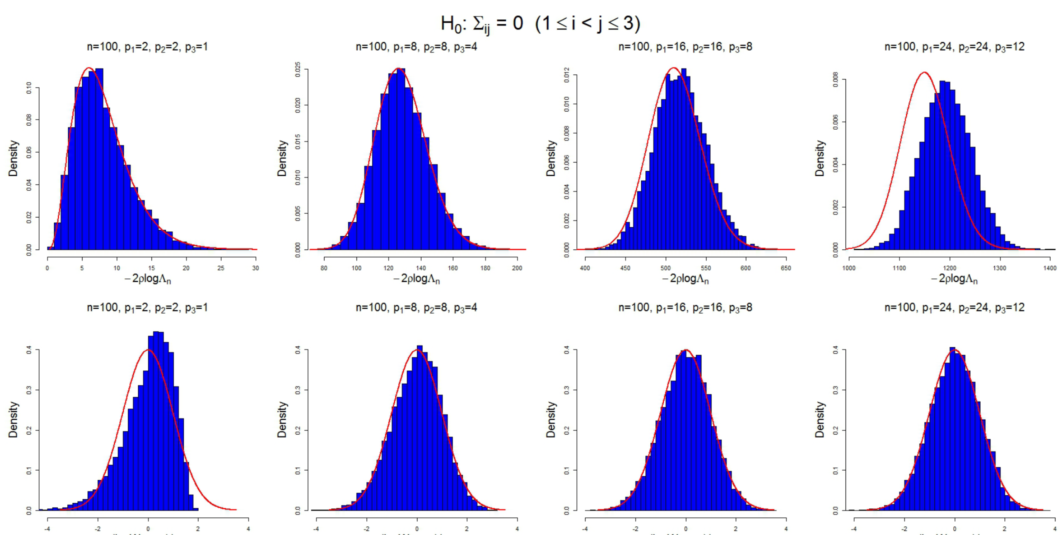

Figure 2: *Comparison between Theorem 2 and (2.6). We choose* $k = 3$*,* $n = 100$ *and* $p = 5, 20, 40, 60$ *with* $p_1 : p_2 : p_3 = 2 : 2 : 1$*. The pictures in the top row show that the histogram of* $-2\rho \log \Lambda_n$ *move away gradually from* χ^2 *curve when* p *grows. The pictures in the bottom row indicate that* $(\log W_n - \mu_n)/\sigma_n$ *and* $N(0,1)$*-curve match better as* p *becomes larger.*

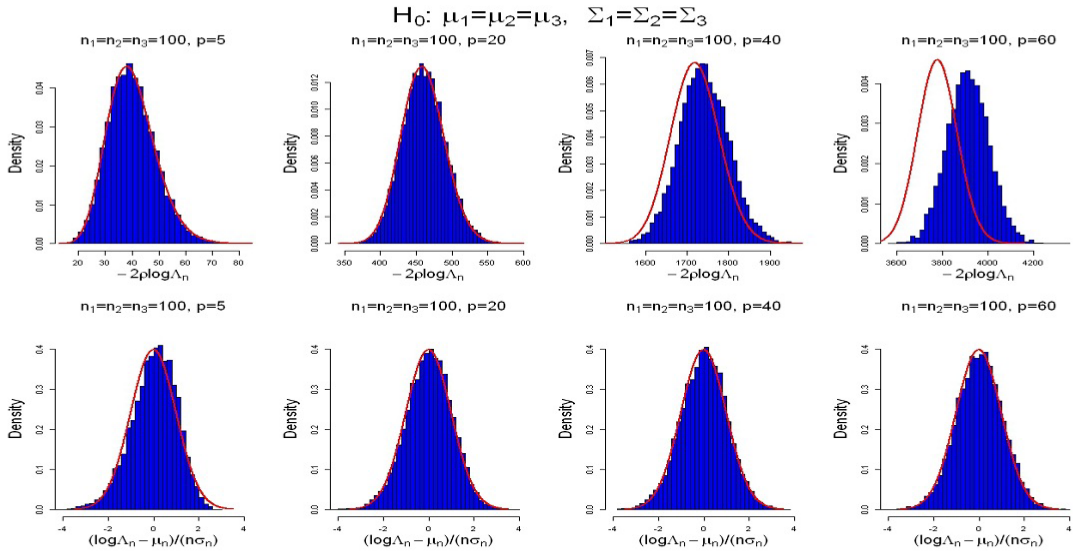

Figure 3: *Comparison between Theorem 3 and (2.9). We choose* $n_1 = n_2 = n_3 = 100$ *with* $p = 5, 20, 40, 60$*. The pictures in the top row show that the* χ^2 *curves stay away farther gradually from the histogram of* $-2\rho \log \Lambda_n$ *when* p *grows. The pictures in the bottom row show that the* $N(0,1)$*-curve fits the histogram of* $(\log \Lambda_n - \mu_n)/(n\sigma_n)$ *very well as* p *grows.*

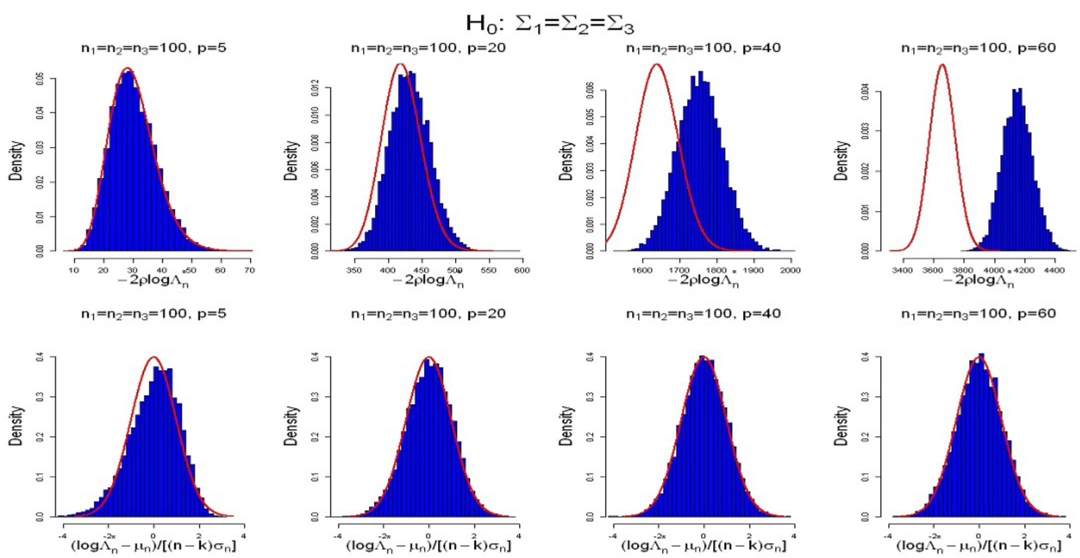

Figure 4: *Comparison between Theorem 4 and (2.12). We chose* $n_1 = n_2 = n_3 = 100$ *with* $p = 5, 20, 40, 60$*. The pictures in the top row show that the* χ^2 *curves goes away quickly from the histogram of* $-2\rho \log \Lambda_n^*$ *as* p *grows. The pictures in the second row show that the* $N(0,1)$*-curve fits the histogram of* $(\log \Lambda_n^* - \mu_n)/[(n-k)\sigma_n]$ *better as* p *grows.*

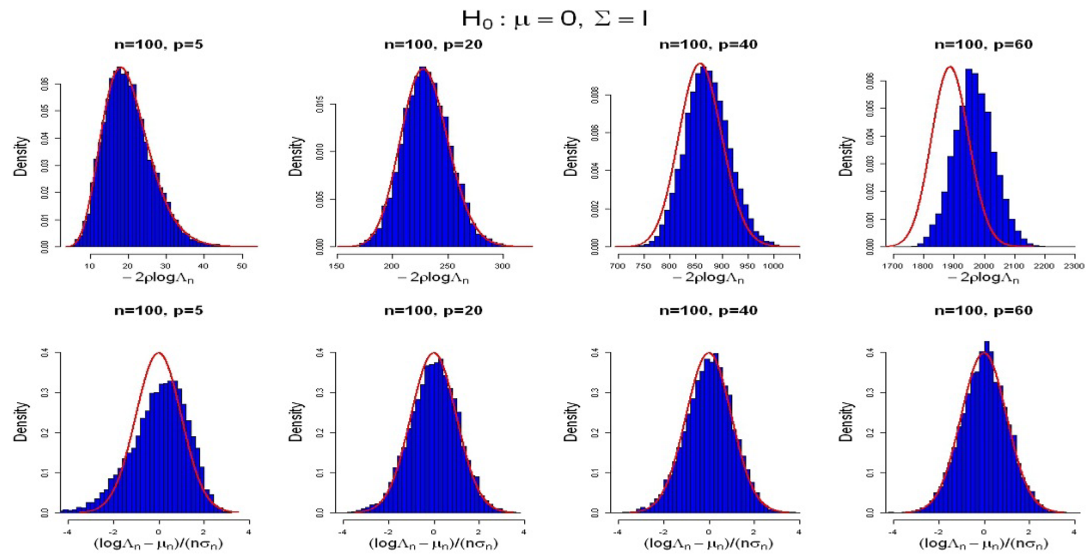

Figure 5: *Comparison between Theorem 5 and (2.16). We choose* $n = 100$ *with* $p = 5, 20, 40, 60$. *The pictures in the first row show that, as* p *is large, the* χ^2*-curve fits the histogram of* $-2\rho \log \Lambda_n$ *poorly. Those in the second row indicate that the N(0,1)-curve fits the histogram of* $(\log \Lambda_n - \mu_n)/(n\sigma_n)$ *very well as* p *is large.*

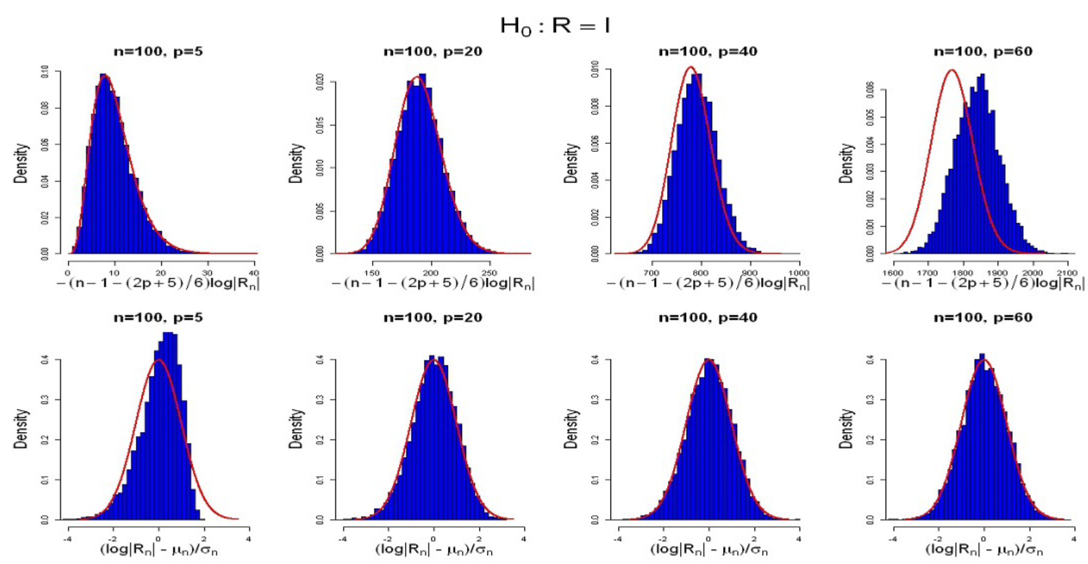

Figure 6: *Comparison between Corollary 1 and (2.21). We choose* $n = 100$ *with* $p = 5, 20, 40, 60$. *The pictures in the first row show that, as* p *is large, the* χ^2*-curve fits the histogram of* $-(n - 1 - \frac{2p+5}{6}) \log |\hat{\mathbf{R}}_n|$ *poorly. Those in the second row indicate that the N(0,1)-curve fits the histogram of* $(\log |\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n$ *very well as* p *is large.*

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.
$n = 100, p = 5$	0.0629	0.0548	0.7519	0.7340
$n = 100, p = 20$	0.0557	0.0546	0.8757	0.8735
$n = 100, p = 40$	0.0529	0.0868	0.8529	0.9022
$n = 100, p = 60$	0.0570	0.3342	0.7887	0.9750

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* $\mathbf{\Sigma} = diag(1.69, \cdots, 1.69, 1, \cdots, 1)$, *where the number of 1.69 appearing in the diagonal is equal to* $\lfloor p/2 \rfloor$.

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.
$n = 100, p_1 = p_2 = 2, p_3 = 1$	0.0659	0.0510	0.7611	0.7199
$n = 100, p_1 = p_2 = 8, p_3 = 4$	0.0570	0.0516	0.9787	0.9767
$n = 100, p_1 = p_2 = 16, p_3 = 8$	0.0508	0.0699	0.9590	0.9730
$n = 100, p_1 = p_2 = 24, p_3 = 12$	0.0539	0.2204	0.8593	0.9714

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* $\mathbf{\Sigma} = 0.15\mathbf{J}_p + 0.85\mathbf{I}_p$.

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.
$n_1 = n_2 = n_3 = 100, p = 5$	0.0567	0.0476	0.5857	0.5499
$n_1 = n_2 = n_3 = 100, p = 20$	0.0493	0.0494	0.7448	0.7455
$n_1 = n_2 = n_3 = 100, p = 40$	0.0519	0.0997	0.6645	0.7751
$n_1 = n_2 = n_3 = 100, p = 60$	0.0495	0.4491	0.5134	0.9400

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* $\mu_1 = (0, \ldots, 0)'$, $\mathbf{\Sigma}_1 = 0.5\mathbf{J}_p + 0.5\mathbf{I}_p$; $\mu_2 = (0.1, \ldots, 0.1)'$, $\mathbf{\Sigma}_2 = 0.6\mathbf{J}_p + 0.4\mathbf{I}_p$; $\mu_3 = (0.1, \ldots, 0.1)'$, $\mathbf{\Sigma}_3 = 0.5\mathbf{J}_p + 0.31\mathbf{I}_p$.

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.
$n_1 = n_2 = n_3 = 100, p = 5$	0.0814	0.0540	0.7239	0.6551
$n_1 = n_2 = n_3 = 100, p = 20$	0.0565	0.0531	0.7247	0.7808
$n_1 = n_2 = n_3 = 100, p = 40$	0.0554	0.0984	0.6111	0.7161
$n_1 = n_2 = n_3 = 100, p = 60$	0.0526	0.4366	0.4649	0.9126

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* $\mathbf{\Sigma}_1 = \mathbf{I}_p$, $\mathbf{\Sigma}_2 = 1.21\mathbf{I}_p$, *and* $\mathbf{\Sigma}_3 = 0.81\mathbf{I}_p$.

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.
$n = 100, p = 5$	0.1086	0.0525	0.5318	0.3872
$n = 100, p = 20$	0.0664	0.0542	0.7684	0.7375
$n = 100, p = 40$	0.0593	0.0905	0.7737	0.8297
$n = 100, p = 60$	0.0610	0.3478	0.7112	0.9552

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* $\mu = (0.1, \ldots, 0.1, 0, \ldots, 0)'$ *where the number of* 0.1 *is equal to* $\lfloor p/2 \rfloor$ *and* $\mathbf{\Sigma} = (\sigma_{ij})_{p\times p}$ *where* $\sigma_{ij} = 1$ *for* $i = j$, $\sigma_{ij} = 0.1$ *for* $0 < |i - j| \leq 3$, *and* $\sigma_{ij} = 0$ *for* $|i - j| > 3$.

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.
$n = 100, p = 5$	0.0581	0.0550	0.5318	0.3872
$n = 100, p = 20$	0.0552	0.0558	0.7684	0.7375
$n = 100, p = 40$	0.0512	0.0870	0.7737	0.8297
$n = 100, p = 60$	0.0555	0.3163	0.7112	0.9552

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* $\mathbf{R} = (r_{ij})_{p\times p}$ *where* $r_{ij} = 1$ *for* $i = j$, $r_{ij} = 0.1$ *for* $0 < |i - j| \leq 3$, *and* $r_{ij} = 0$ *for* $|i - j| > 3$.

4 Conclusions and Discussions

We study six likelihood ratio tests in this paper. The central limit theorems of the six statistics are derived under the assumption that the population dimension $p \to \infty$ and $p < n - c$ for some constant c with $1 \le c \le 4$. Jiang and Yang (2013) show that the CLTs hold only at $p/n \to y \in (0, 1]$. In this paper we get the same CLTs under almost most relaxed conditions. Precisely, if p is finite the LRT statistics converge weakly to chi-square distributions; if $p > n - 1$ the LRT statistics do no exist. The only "sacrifice" is that p is not allowed to be too close to n such as $p = n - 1$ in some cases. Ignoring these small technical losses, we get the sufficient and necessary conditions for the CLTs. These give us almost the maximum flexibilities to use them in practice.

The strategies of our proofs are based on the moment generating functions of targeted LRT statistics. We develop a new tool in Proposition 5.1 which is the key part in the proofs. This very technically involved tool is different from those used in Bai et al. (2009), Jiang et al. (2012) and Jiang and Yang (2013).

Finally, we make some comments as follows.

(1) Let $\mathbf{X}_1, \cdots, \mathbf{X}_n$ be i.i.d. random vectors with a probability density function or probability mass function $p(x|\theta)$ where $\theta \in \Theta$. Consider the hypothesis test $H_0 : \theta \in \Theta_0$ vs $H_a : \theta \in \Theta \backslash \Theta_0$. Let Λ_n be the likelihood ratio test statistic. The Wilks theorem (Wilks (1938) or van der Vaart (1998)) says that $-2 \log \Lambda_n \to \chi^2_d$ as $n \to \infty$ under H_0, where d, the difference between the dimensions of Θ and Θ_0, is fixed. If $d = d_n$ depends on n and goes to infinity but at a very slow rate, we can see that the distributions of $-2 \log \Lambda_n$ and χ^2_d are very close. Meanwhile, by the standard central limit theorem, $(\chi^2_d - d)/\sqrt{2d} \to N(0, 1)$ whenever $d \to \infty$. So, heuristically, $(2 \log \Lambda_n + d)/\sqrt{2d} \to N(0, 1)$ when $d = d_n$ goes to infinity very slowly as $n \to \infty$, it may not be true when d goes to infinity too fast. This process is similar to the exchange of the limits in the Real and Complex Analysis, which is usually non-trivial. The six theorems we present in the paper can help clarify the situation when d goes to infinity, that is, even though the central limit theorems hold for $-2 \log \Lambda_n$, the asymptotic mean and variance may not be d and $2d$ anymore. The proofs of our Theorems 1-6 are based on the analysis of the moments of Λ_n which are available thanks to the normal assumptions. In general, without the the normal assumption on $\mathbf{X}_i$'s, we expect the above heuristic to work also under some technical conditions but details can be very complicated.

(2) In Theorems 2, 3 and 4, the assumption "$\delta < p_i/p_j < \delta^{-1}$" or "$\delta < n_i/n_j < \delta^{-1}$" says that the population distribution dimensions or the sample sizes are comparable. We impose these assumptions in the theorems only for the purpose to simplify the proofs. It is possible that the conditions can be relaxed. And it will be interesting to see how less constrained among the p_i's or n_i's to make the three theorems hold.

(3) Recently some authors study similar problems under the nonparametric setting, see, e.g., Cai et al. (2013), Cai and Ma (2012), Chen et al. (2010), Li and Chen (2012), Qiu and Chen (2012) and Xiao and Wu (2013).

(4) If the normality assumptions are slightly altered, how much the corresponding central limit theorems are affected is the problem of robustness. See a detailed discussion at Comment 4 at Section 4 from Jiang and Yang (2013).

(5) The central limit theorems in this paper are derived under null hypothesis. The analogues under alternative hypotheses are related to the zonal polynomials. See a further discussion at Comment 1 in Section 4 from Jiang and Yang (2013). Here we consider the scenarios when p is smaller than n such that either p is at the same scale of n or p is much smaller than n. These are the necessary situations to study the likelihood ratio tests because the tests do not exist otherwise. In a bigger picture, experts consider tests with the dimensionality of data p being larger or much larger than the sample size n. Readers are referred to the papers, for example, by Ledoit and Wolf (2002) and Chen et al. (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. Onatski et al. study the powers for the sphericity test. A projection method is used to investigate the two-sample test by Lopes et al.

5 Proofs

This section is divided into seven subsections. We first develop some tools, and then in each of the subsequent subsections, we prove one theorem introduced in Section 2. The following are some standard notation.

For two sequences of numbers $\{a_n;\, n \ge 1\}$ and $\{b_n;\, n \ge 1\}$, 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 $\lim_{n\to\infty} a_n/b_n = 0$, and the symbol $a_n \sim b_n$ stands for $\lim_{n\to\infty} a_n/b_n = 1$. For two functions $f(x)$ and $g(x)$, the notation $f(x) = O(g(x))$, $f(x) = o(g(x))$ and $f(x) \sim g(x)$ as $x \to x_0 \in [-\infty, \infty]$ are similarly interpreted.

5.1 A Preparation

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

$$\Gamma_p(z) := \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\Big(z - \frac{1}{2}(i-1)\Big) \tag{5.1}$$

for complex number z with $\mathrm{Re}(z) > \frac{1}{2}(p-1)$. See p. 62 from Muirhead (1982). The key tool to prove the main theorems is the following analysis of $\Gamma_p(z)$.

PROPOSITION 5.1 *Let $\{p = p_n \in \mathbb{N};\ n \geq 1\}$, $\{m = m_n \in \mathbb{N};\ n \geq 1\}$ and $\{t_n \in \mathbb{R};\ n \geq 1\}$ satisfy that (i) $p_n \to \infty$ and $p_n = o(n)$; (ii) there exists $\epsilon \in (0,1)$ such that $\epsilon \leq m_n/n \leq \epsilon^{-1}$ for large n; (iii) $t = t_n = O(n/p)$. Then, as $n \to \infty$,*

$$\log \frac{\Gamma_p(\frac{m-1}{2}+t)}{\Gamma_p(\frac{m-1}{2})} = \alpha_n t + \beta_n t^2 + \gamma_n(t) + o(1)$$

where

$$\alpha_n = -\big[2p + (m-p-\frac{3}{2})\log(1-\frac{p}{m-1})\big]; \quad \beta_n = -\big[\frac{p}{m-1} + \log\big(1-\frac{p}{m-1}\big)\big];$$
$$\gamma_n(t) = p\Big[\big(\frac{m-1}{2}+t\big)\log\big(\frac{m-1}{2}+t\big) - \frac{m-1}{2}\log\frac{m-1}{2}\Big].$$

In addition, in the proofs of the main theorems, we will repeatedly use the so-called subsequence argument, that is, to prove that a sequence of random variables converges in distribution to $N(0,1)$, it suffices to show that every subsequence has a further subsequence that converges in distribution to the standard normal. This further subsequence is selected in a way that the subsequential limits of some bounded quantities exist. Therefore, we only need to verify the theorems by assuming that the limits for these bounded quantities (to be specified in each proof below) exist.

Proposition 5.1 will be proved by using a series of lemmas given below.

LEMMA 5.1 *As $x \to +\infty$,*

$$\log\frac{\Gamma(x+b)}{\Gamma(x)} = (x+b)\log(x+b) - x\log x - b - \frac{b}{2x} + O\big(\frac{b^2+1}{x^2}\big) \tag{5.2}$$

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

Proof. Recall the Stirling formula (see, e.g., p. 368 from Gamelin (2001) or (37) on p. 204 from Ahlfors (1979)):

$$\log\Gamma(x) = \big(x-\frac{1}{2}\big)\log x - x + \log\sqrt{2\pi} + \frac{1}{12x} + O\left(\frac{1}{x^3}\right) \tag{5.3}$$

as $x \to +\infty$. For any fixed $\delta \in (0,1)$, we have that

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

uniformly on $b \in (-\delta x, \delta x)$ as $x \to \infty$. Then (5.2) follows from the facts that

$$\log(1+\frac{b}{x}) = \frac{b}{x} + O(\frac{b^2}{x^2}) \qquad \text{and} \qquad \frac{b}{x(x+b)} = O(\frac{|b|}{x^2}) = O(\frac{b^2+1}{x^2})$$

uniformly on $b \in (-\delta x, \delta x)$ as $x \to \infty$. ■

Lemma 5.2 *Let $p = p_n$ be such that $1 \le p < n$, $p \to \infty$ and $p/n \to 0$. Then,*

$$\sum_{i=1}^{p}\Big(\frac{1}{n-i}-\frac{1}{n-1}\Big) = \frac{\sigma_n^2}{2}\Big(1+O\big(\frac{1}{p}+\frac{p}{n}\big)\Big); \tag{5.4}$$

$$\sum_{i=1}^{p}(\log(n-1)-\log(n-i)) = -\mu_n + O(\sigma_n^2) \tag{5.5}$$

as $n \to \infty$, where $\sigma_n^2 = -2\big[\frac{p}{n-1}+\log\big(1-\frac{p}{n-1}\big)\big]$ and

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

Proof. First, note that $\sigma_n^2 = (\frac{p}{n-1})^2(1+O(\frac{p}{n-1}))$, which implies

$$\frac{1}{2}(\frac{p}{n-1})^2 = \frac{\sigma_n^2}{2}(1+O(\frac{p}{n})) \tag{5.6}$$

as $n \to \infty$. Observe that $|\frac{i-1}{n-i}-\frac{i-1}{n-1}| = \frac{(i-1)^2}{(n-1)(n-i)} \le 2\frac{(i-1)^2}{(n-1)^2}$ for all $1 \le i \le p$ as n is sufficiently large. Then,

$$\begin{aligned}\sum_{i=1}^{p}(\frac{1}{n-i}-\frac{1}{n-1}) = \frac{1}{n-1}\sum_{i=1}^{p}\frac{i-1}{n-i} &= \frac{1}{n-1}\sum_{i=1}^{p}\frac{i-1}{n-1}+\epsilon_n \\ &= \frac{p(p-1)}{2(n-1)^2}+\epsilon_n\end{aligned}$$

where $|\epsilon_n| \le 2(n-1)^{-3}\sum_{i=1}^{p}(i-1)^2 \le 3(p/n)^3$ as $n \to \infty$. Evidently, $\frac{p(p-1)}{2(n-1)^2} = \frac{p^2}{2(n-1)^2} + O(pn^{-2})$. Then

$$\begin{aligned}\sum_{i=1}^{p}(\frac{1}{n-i}-\frac{1}{n-1}) &= \frac{p^2}{2(n-1)^2}+O\big((\frac{p}{n})^3+\frac{p}{n^2}\big) \\ &= \frac{p^2}{2(n-1)^2}\Big(1+O\big(\frac{p}{n}+\frac{1}{p}\big)\Big) = \frac{\sigma_n^2}{2}\Big(1+O\big(\frac{p}{n}+\frac{1}{p}\big)\Big)\end{aligned}$$

by (5.6). This concludes (5.4).

Now we show (5.5). Apply the Stirling formula (5.3) to $x = n-1$ and $x = n-p-1$ and take difference to get

$$\begin{aligned}&\log\Gamma(n-1)-\log\Gamma(n-p-1) \\ =\ & (n-\frac{3}{2})\log(n-1)-(n-p-\frac{3}{2})\log(n-p-1)-p \\ &\qquad + \frac{1}{12}(\frac{1}{n-1}-\frac{1}{n-p-1})+O(\frac{1}{n^3}) \\ =\ & (n-\frac{3}{2})\log(n-1)-(n-p-\frac{3}{2})\log(n-p-1)-p+O(\frac{p}{n^2})\end{aligned}$$

as $n\to\infty$. Note that for any integer $k\geq 1$, $\Gamma(k)=(k-1)!=\prod_{i=1}^{k-1} i$. Then we have

$$\begin{aligned}
&\sum_{i=1}^{p}(\log(n-1)-\log(n-i))\\
=&\ \ p\log(n-1)-(\log\Gamma(n-1)-\log\Gamma(n-p-1))+\log(1-\frac{p}{n-1})\\
=&\ \ -(p-n+\frac{3}{2})\log(1-\frac{p}{n-1})+p+\log(1-\frac{p}{n-1})+O(\frac{p}{n^2})\\
=&\ \ -(p-n+\frac{3}{2})\log(1-\frac{p}{n-1})+p-\frac{p}{n-1}+O((\frac{p}{n})^2)\\
=&\ \ -(p-n+\frac{3}{2})\log(1-\frac{p}{n-1})+\frac{n-2}{n-1}p+O((\frac{p}{n})^2)
\end{aligned}$$

as $n\to\infty$. Then (5.5) follows by noticing $\frac{n-2}{n-1}p=\frac{n-1}{n}p+O(pn^{-2})$. ■

Lemma 5.3 *Let $p=p_m$ be such that $1\leq p<m$, $p\to\infty$ and $p/m\to 0$ as $m\to\infty$. Define*

$$g_i(x)=\big(\frac{m-i}{2}+x\big)\log\big(\frac{m-i}{2}+x\big)-\big(\frac{m-1}{2}+x\big)\log\big(\frac{m-1}{2}+x\big)$$

for $1\leq i\leq p$ and $x>-(m-p)/2$. Let μ_m and $\sigma_m>0$ be as in Lemma 5.2. Given $t=t_m=O(m/p)$, we have that, as $m\to\infty$,

$$\sum_{i=1}^{p}\big(g_i(t)-g_i(0)\big)=\mu_m t+\frac{\sigma_m^2}{2}t^2+o(1).$$

Proof. Evidently,

$$\begin{aligned}
g_i'(x)&=\log\big(\frac{m-i}{2}+x\big)-\log\big(\frac{m-1}{2}+x\big),\\
g_i''(x)&=\frac{1}{\frac{m-i}{2}+x}-\frac{1}{\frac{m-1}{2}+x}=\frac{\frac{i-1}{2}}{(\frac{m-i}{2}+x)(\frac{m-1}{2}+x)},\\
g_i^{(3)}(x)&=-\frac{1}{(\frac{m-i}{2}+x)^2}+\frac{1}{(\frac{m-1}{2}+x)^2}=-\frac{\frac{i-1}{2}\cdot\frac{2m-i-1}{2}+(i-1)x}{(\frac{m-i}{2}+x)^2(\frac{m-1}{2}+x)^2}
\end{aligned}$$

for all $1\leq i\leq p$. Easily, $\sigma_m\sim\frac{p}{m}$. Then

$$\sup_{|x|\leq|t|,\ 1\leq i\leq p}|g_i^{(3)}(x)|\leq C\frac{p}{m^3}$$

where C is a constant not depending on t,p or m. By Taylor's expansion,

$$\begin{aligned}
g_i(t)-g_i(0)&=g_i'(0)t+\frac{t^2}{2}g_i''(0)+\frac{t^3}{6}g_i^{(3)}(\xi_i)\\
&=\big(\log(m-i)-\log(m-1)\big)t+\frac{i-1}{(m-1)(m-i)}t^2+\epsilon_i
\end{aligned}$$

where ξ_i is between 0 and t for all $1 \le i \le p$, and $\sup_{1\le i\le p}|\epsilon_i| = O(p^{-2})$. Also, $\sigma_m^2 t = O(p/m)$. By (5.5) and then (5.4), we obtain

$$\begin{aligned}\sum_{i=1}^{p}\big(g_i(t)-g_i(0)\big) &= \mu_m t + \frac{t^2}{m-1}\sum_{i=1}^{p}\frac{i-1}{m-i} + O\big(\frac{1}{p}+\frac{p}{m}\big)\\ &= \mu_m t + t^2\sum_{i=1}^{p}\Big(\frac{1}{m-i}-\frac{1}{m-1}\Big)+o(1)\\ &= \mu_m t + \frac{\sigma_m^2}{2}t^2 + o(1)\end{aligned}$$

as $m\to\infty$, proving the lemma. ■

Now we are ready to prove the key result in this subsection.
Proof of Proposition 5.1. From (5.1), we have

$$\log\frac{\Gamma_p(\frac{m-1}{2}+t)}{\Gamma_p(\frac{m-1}{2})} = \sum_{i=1}^{p}\log\frac{\Gamma(\frac{m-i}{2}+t)}{\Gamma(\frac{m-i}{2})}.$$

By the given condition, $\frac{t^2+|t|+1}{m^2} = O(p^{-2})$ as $n\to\infty$. Then, we have from (5.2) that

$$\begin{aligned}\log\frac{\Gamma(\frac{m-i}{2}+t)}{\Gamma(\frac{m-i}{2})} &= \big(\frac{m-i}{2}+t\big)\log\big(\frac{m-i}{2}+t\big) - \frac{m-i}{2}\log\frac{m-i}{2}\\ &\quad -t-\frac{t}{m-i}+O\big(\frac{1}{p^2}\big)\end{aligned}$$

as $n\to\infty$ uniformly for all $1\le i\le p$. Write $\frac{t}{m-i} = \frac{t}{m}+\frac{t}{m}\cdot\frac{i}{m-i}$. Easily, $\sum_{i=1}^{p}\frac{t}{m-i} = \frac{pt}{m}+O(\frac{p}{n})$ since $t = O(n/p)$. Then

$$\sum_{i=1}^{p}\Big(-t-\frac{t}{m-i}+O\big(\frac{1}{p^2}\big)\Big) = -pt-\frac{pt}{m}+O\big(\frac{1}{p}+\frac{p}{n}\big)$$

as $n\to\infty$. Therefore,

$$\begin{aligned}&\log\frac{\Gamma_p(\frac{m-1}{2}+t)}{\Gamma_p(\frac{m-1}{2})}\\ &= -\frac{(m+1)pt}{m}+\sum_{i=1}^{p}\Big[\big(\frac{m-i}{2}+t\big)\log\big(\frac{m-i}{2}+t\big)-\frac{m-i}{2}\log\frac{m-i}{2}\Big]+o(1) \qquad (5.7)\end{aligned}$$

as $n\to\infty$. Set

$$g_i(x) = \big(\frac{m-i}{2}+x\big)\log\big(\frac{m-i}{2}+x\big)-\big(\frac{m-1}{2}+x\big)\log\big(\frac{m-1}{2}+x\big)$$

for $1 \le i \le p$ and $x > -(m-p)/2$. Then, it is trivial to verify that the term of "$\sum$" in (5.7) is equal to

$$\begin{aligned} & p\Big[\big(\frac{m-1}{2}+t\big)\log\big(\frac{m-1}{2}+t\big)-\frac{m-1}{2}\log\frac{m-1}{2}\Big]+\sum_{i=1}^{p}\big(g_i(t)-g_i(0)\big) \\ = \ & p\Big[\big(\frac{m-1}{2}+t\big)\log\big(\frac{m-1}{2}+t\big)-\frac{m-1}{2}\log\frac{m-1}{2}\Big]+\mu_m t+\frac{1}{2}\sigma_m^2 t^2+o(1) \end{aligned}$$

by Lemma 5.3, where

$$\mu_m = (p-m+\frac{3}{2})\log(1-\frac{p}{m-1})-\frac{m-1}{m}p \ \text{ and } \ \sigma_m^2 = -2\big[\frac{p}{m-1}+\log\big(1-\frac{p}{m-1}\big)\big]$$

This joint with (5.7) proves the proposition by simply noticing

$$\begin{aligned} & \alpha_n = -\frac{(m+1)p}{m}+\mu_m, \quad \beta_n = \frac{1}{2}\sigma_m^2, \\ & \gamma_n(t) = p\Big[\big(\frac{m-1}{2}+t\big)\log\big(\frac{m-1}{2}+t\big)-\frac{m-1}{2}\log\frac{m-1}{2}\Big]. \end{aligned}$$

■

5.2 Proof of Theorem 1

Lemma 5.4 *(Corollary 8.3.6 from Muirhead (1982)) Assume $n > p$. Let V_n be as in (2.2). Then, under H_0 in (2.1), we have*

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

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

The restriction "$h > -\frac{n-p}{2}$" in the lemma comes from the definition of $\Gamma_p(z)$ as in (5.1).

Proof of Theorem 1. We need to prove

$$H_n := \frac{\log V_n - \mu_n}{\sigma_n} \text{ converges to } N(0,1) \tag{5.8}$$

in distribution as $n \to \infty$. Equivalently, it suffices to show that for any subsequence $\{n_k\}$, there is a further subsequence $\{n_{k_j}\}$ such that $H_{n_{k_j}}$ converges to $N(0,1)$ in distribution as $j \to \infty$. Now, noticing $p_n/n \in [0,1]$ for all n, for any subsequence n_k, take a further subsequence n_{k_j} such that $p_{n_{k_j}}/n_{k_j} \to y \in [0,1]$. So, without loss of generality, we only need to show (5.8) under the condition that $\lim_{n\to\infty} p_n/n = y \in [0,1]$.

By Theorem 1 from Jiang and Yang (2013), we know that the theorem is true for the case $\lim p/n = y \in (0,1]$. Thus, to prove (5.8), we only need to prove it for the case $\lim p_n/n = 0$. From now on, we assume $p = p_n \to \infty$ and $p = o(n)$. It is enough to show that

$$E\exp\Big\{\frac{\log V_n - \mu_n}{\sigma_n}\,s\Big\} \to e^{s^2/2}$$

as $n \to \infty$ for all $|s| \le 1$, or equivalently,

$$\log E(V_n^t) = \mu_n t + \frac{1}{2}s^2 + o(1) \tag{5.9}$$

as $n \to \infty$ for all $|s| \le 1$, where $t = t_n = \frac{s}{\sigma_n}$.

Set $m = n-1$. Obviously, $\sigma_n \sim \frac{p}{m}$ as $n \to \infty$. Thus, $t \sim \frac{m}{p}s \ge -\frac{m}{p} > -\frac{n-p}{2}$ as n is sufficiently large. By Lemma 5.4,

$$E(V_n^t) = p^{pt}\frac{\Gamma(\frac{mp}{2})}{\Gamma(\frac{mp}{2}+pt)} \cdot \frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})} \tag{5.10}$$

for all large n. Easily, $pt/(\frac{1}{2}mp) = 2s/(m\sigma_n) = O(1/p)$ and hence $(p^2t^2+p|t|+1)/(m^2p^2) = O(1/p^2)$. By (5.2),

$$\begin{aligned}\log\frac{\Gamma(\frac{mp}{2})}{\Gamma(\frac{mp}{2}+pt)} &= -\big(\frac{mp}{2}+pt\big)\log\big(\frac{mp}{2}+pt\big) + \frac{mp}{2}\log\frac{mp}{2} + pt + O\big(\frac{1}{p}\big) \\ &= -\gamma_n(t) + p(1-\log p)t + O\big(\frac{1}{p}\big) \qquad (5.11)\end{aligned}$$

as $n \to \infty$, where

$$\gamma_n(t) = p\Big[\big(\frac{m}{2}+t\big)\log\big(\frac{m}{2}+t\big) - \frac{m}{2}\log\frac{m}{2}\Big]$$

and the formula $\log(uv) = \log u + \log v$ is used in the last equality. Since $t \sim \frac{m}{p}s$, by Proposition 5.1,

$$\log\frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})} = \alpha_n t + \beta_n t^2 + \gamma_n(t) + o(1) \tag{5.12}$$

as n is sufficiently large, where

$$\alpha_n = -\big[2p + (n-p-\frac{3}{2})\log(1-\frac{p}{n-1})\big] \text{ and } \beta_n = -\big[\frac{p}{n-1} + \log\big(1-\frac{p}{n-1}\big)\big].$$

Combining (5.10)-(5.12), we get

$$\begin{aligned}\log E(V_n^t) &= pt\log p + \log\frac{\Gamma(\frac{mp}{2})}{\Gamma(\frac{mp}{2}+pt)} + \log\frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})} \\ &= (\alpha_n + p)t + \beta_n t^2 + o(1) = \mu_n t + \frac{s^2}{2} + o(1)\end{aligned}$$

as $n \to \infty$. Hence, we get (5.9) and complete the proof of the theorem. ■

5.3 Proof of Theorem 2

LEMMA 5.5 *(Theorem 11.2.3 from Muirhead (1982)) Let p and W_n be as in Section 2.2. Then, under H_0 in (2.4),*

$$EW_n^t = \frac{\Gamma_p(\frac{n}{2}+t)}{\Gamma_p(\frac{n}{2})} \cdot \prod_{i=1}^{k}\frac{\Gamma_{p_i}(\frac{n}{2})}{\Gamma_{p_i}(\frac{n}{2}+t)}$$

for any $t > (p-n-1)/2$, where $\Gamma_p(z)$ is as in (5.1).

Proof of Theorem 2. We need to prove

$$H_n := \frac{\log W_n - \mu_n}{\sigma_n} \text{ converges to } N(0,1) \tag{5.13}$$

in distribution as $n \to \infty$, where

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

(the fact $\sigma_n^2 > 0$ is shown in Theorem 2 from Jiang and Yang (2013)). By an argument similar to the first paragraph of the proof of Theorem 1, it suffices to show (5.13) under the condition $\lim_{n\to\infty} p_i/n = y_i \in [0,1]$ for each $1 \le i \le k$. From now on, we assume that this is true.

From the assumption $n > p = p_1 + \cdots + p_k$ and that $\delta \le p_i/p_j \le \delta^{-1}$ for all $1 \le i, j \le n$ and all n, we know that either $y_i \in (0,1)$ for all $1 \le i \le k$ or $y_1 = \cdots = y_k = 0$. The first case is proved in Theorem 2 from Jiang and Yang (2013). Now we study the second case, that is, $y_1 = \cdots = y_k = 0$.

Since $\lim_{n\to\infty} p_i/n = 0$ for each $1 \le i \le k$, it follows from the Taylor expansion that

$$\sigma_n^2 \sim \frac{1}{n^2}\big(p^2 - \sum_{i=1}^k p_i^2\big) \tag{5.14}$$

as $n \to \infty$. To complete the proof, it suffices to show that

$$E\exp\Big\{\frac{\log W_n - \mu_n}{\sigma_n}s\Big\} = \exp(-\frac{\mu_n s}{\sigma_n})E[W_n^{\frac{s}{\sigma_n}}] \to e^{s^2/2}$$

as $n \to \infty$ for all s such that $|s| \le 1$, or equivalently,

$$\log E[W_n^t] = \mu_n t + \frac{s^2}{2} + o(1) \tag{5.15}$$

as $n \to \infty$ for all $|s| \le 1$, where $t := s/\sigma_n$.

Now fix s with $|s| \le 1$. By (5.14), $\sigma_n^2 \ge p_1p_2/n^2$ for all large n since $p = p_1 + \cdots + p_k$. It follows that $|t| \le n/\sqrt{p_1p_2}$, and hence $t \ge -n/\sqrt{p_1p_2} > (p-n-1)/2$ as n is large enough. By Lemma 5.5,

$$\log E[W_n^t] = \log\frac{\Gamma_p(\frac{n}{2}+t)}{\Gamma_p(\frac{n}{2})} - \sum_{i=1}^k \log\frac{\Gamma_{p_i}(\frac{n}{2}+t)}{\Gamma_{p_i}(\frac{n}{2})}.$$

Now, from (5.14) we see that

$$\Big(\frac{t}{(n/p)}\Big)^2 \sim \frac{p^2}{p^2 - \sum_{i=1}^k p_i^2}s^2 \le \Big(1 - \frac{\sum_{i=1}^k p_i^2}{p^2}\Big)^{-1}.$$

Without loss of generality, assume $p_1 \geq \cdots \geq p_k$. By assumption, $p_i \geq \delta\, p_1$ for all $1 \leq i \leq k$. Thus,

$$\frac{\sum_{i=1}^k p_i^2}{p^2} \leq \frac{p_1}{\sum_{i=1}^k p_i} \leq \frac{1}{1+(k-1)\delta} < 1$$

which says $t = O(n/p)$ and hence $t = O(n/p_i)$. By Proposition 5.1, $\log E[W_n^t]$ is equal to

$$\begin{aligned}
&-\big[2p + (n-p-\frac{1}{2})\log(1-\frac{p}{n})\big]t + \sum_{i=1}^k \big[2p_i + (n-p_i-\frac{1}{2})\log(1-\frac{p_i}{n})\big]t \\
&-\big[\frac{p}{n} + \log\big(1-\frac{p}{n}\big)\big]t^2 + \sum_{i=1}^k \big[\frac{p_i}{n} + \log\big(1-\frac{p_i}{n}\big)\big]t^2 \\
&+(p-\sum_{i=1}^k p_i)\Big[\big(\frac{n}{2}+t\big)\log\big(\frac{n}{2}+t\big) - \frac{n}{2}\log\frac{n}{2}\Big] + o(1).
\end{aligned}$$

Use the fact that $p = \sum_{i=1}^k p_i$ to see that the "$2p$" and the "$2p_i$'s" in the first line above are canceled; the terms "$\frac{p}{n}$" and "$\frac{p_i}{n}$'s" in the second line are canceled; the whole third line is $0 + o(1) = o(1)$. Review the notation μ_n and σ_n in the theorem. These lead to

$$\log E[W_n^t] = \mu_n t + \frac{\sigma_n^2}{2}t^2 + o(1) = \mu_n t + \frac{s^2}{2} + o(1)$$

as $n \to \infty$, which is exactly (5.15). ■

5.4 Proof of Theorem 3

Review the notation in (2.8). Let

$$\lambda_n = \frac{\prod_{i=1}^k |\mathbf{B}_i|^{n_i/2}}{|\mathbf{A}+\mathbf{B}|^{n/2}}. \tag{5.16}$$

LEMMA 5.6 *(Corollary 10.8.3 from Muirhead (1982)) Let $n_i > p$ for $i = 1, 2\cdots, k$. Let λ_n be as in (5.16). Then, under H_0 in (2.7),*

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

for all $t > \max_{1\leq i\leq k}\{\frac{p}{n_i}\} - 1$, where $\Gamma_p(z)$ is as in (5.1).

The restriction $t > \max_{1\leq i\leq k}\{\frac{p}{n_i}\} - 1$ comes from the restriction in (5.1).

LEMMA 5.7 *Suppose the conditions in Theorem 3 hold. Let $t = t_p \sim C/p$ as $p \to \infty$ for some constant C. Define*

$$\rho_l(t) = p\Big[\big(\frac{l-1}{2} + \frac{lt}{2}\big)\log\big(\frac{l-1}{2} + \frac{lt}{2}\big) - \frac{l-1}{2}\log\frac{l-1}{2}\Big]$$

for $l \geq (1+t)^{-1}$. Then, as $p \to \infty$,

$$-\rho_n(t) + \sum_{i=1}^{k} \rho_{n_i}(t) = \frac{1}{2} p\Big[(1-k) - n\log n + \sum_{i=1}^{k} n_i \log n_i\Big] t + O(\frac{1}{p} + \frac{p}{n}).$$

Proof. Set $\psi(x) = x\log x$ for $x > 0$. Then, $\psi'(x) = 1 + \log x$ and $\psi''(x) = x^{-1}$. By the Taylor expansion at $x = \frac{n_i}{2} + \frac{n_i t}{2}$,

$$\begin{aligned}
& \big(\frac{n_i-1}{2} + \frac{n_i t}{2}\big)\log\big(\frac{n_i-1}{2} + \frac{n_i t}{2}\big) \\
= & \big(\frac{n_i}{2} + \frac{n_i t}{2}\big)\log\big(\frac{n_i}{2} + \frac{n_i t}{2}\big) - \frac{1}{2}\big(1 + \log\frac{n_i(1+t)}{2}\big) + O\big(\frac{1}{n}\big) \\
= & \frac{1}{2}(1+t)n_i \log\frac{n_i}{2} + \frac{1}{2} n_i (1+t)\log(1+t) - \frac{1}{2}\big(1 + \log\frac{n_i(1+t)}{2}\big) + O\big(\frac{1}{n}\big)
\end{aligned}$$

as $p \to \infty$. Use $\log\frac{n_i-1}{2} = \log\frac{n_i}{2} - \frac{1}{n_i} + O(\frac{1}{n_i^2})$ as $p \to \infty$ to have

$$\frac{n_i-1}{2}\log\frac{n_i-1}{2} = \frac{n_i}{2}\log\frac{n_i}{2} - \frac{1}{2}\big(1 + \log\frac{n_i}{2}\big) + O\big(\frac{1}{n}\big)$$

as $p \to \infty$. By subtracting the second identity from the first one, we have

$$\begin{aligned}
\frac{1}{p}\rho_{n_i}(t) &= \frac{t}{2}\Big(n_i \log\frac{n_i}{2}\Big) + n_i\Big(\frac{1+t}{2}\log(1+t)\Big) - \frac{1}{2}\log(1+t) + O\big(\frac{1}{n}\big) \\
&= \frac{t}{2}(-1 + n_i\log\frac{n_i}{2}) + n_i\Big(\frac{1+t}{2}\log(1+t)\Big) + O\big(\frac{1}{p^2} + \frac{1}{n}\big)
\end{aligned}$$

as $p \to \infty$. By the same argument,

$$\frac{1}{p}\rho_n(t) = \frac{t}{2}\Big(-1 + n\log\frac{n}{2}\Big) + n\Big(\frac{1+t}{2}\log(1+t)\Big) + O\big(\frac{1}{p^2} + \frac{1}{n}\big)$$

as $p \to \infty$. Use the fact $n = n_1 + \cdots + n_k$ to have that

$$-\rho_n(t) + \sum_{i=1}^{k} \rho_{n_i}(t) = \frac{1}{2} p\Big[(1-k) - n\log n + \sum_{i=1}^{k} n_i \log n_i\Big] t + O(\frac{1}{p} + \frac{p}{n})$$

as $p \to \infty$, proving the lemma. ■

Proof of Theorem 3. Since $0 \leq p/n_i < 1$ for each i, by the subsequence principle as in the first paragraph of the proof of Theorem 1, it suffices to prove the theorem provided $\lim p/n_i = y_i \in [0,1]$ for $1 \leq i \leq k$. The theorem is proved by Jiang and Yang (2013) when $y_i \in (0,1]$. Note that if one of y_i is zero, then all y_i's are zero. Now we prove it by assuming $y_1 = \cdots = y_k = 0$ through several steps.

Step 1. It is shown in Theorem 3 from Jiang and Yang (2013) that $\sigma_n^2 > 0$ for all $\min_{1\leq i\leq k} n_i > p+1$ and $p \geq 1$. Note that

$$\log\big(1 - \frac{p}{n_i - 1}\big) - \log\big(1 - \frac{p}{n_i}\big) = -\log\frac{n_i - 1}{n_i} + \log\frac{n_i - p - 1}{n_i - p}.$$

By the formula $\log(1-x) = -x - \frac{x^2}{2} + O(x^3)$ as $x \to 0$, we have that

$$\begin{aligned}
& \log\big(1 - \frac{p}{n_i - 1}\big) - \log\big(1 - \frac{p}{n_i}\big) \\
= \;& \frac{1}{n_i} + \frac{1}{2n_i^2} - \frac{1}{n_i - p} - \frac{1}{2(n_i-p)^2} + O(\frac{1}{n_i^3}) \\
= \;& -\frac{p}{n_i(n_i-p)} + O(\frac{p}{n^3}) \qquad (5.17)
\end{aligned}$$

as $p \to \infty$ for all $1 \le i \le k$ by using the assumption $\delta < n_i/n_j \le \delta^{-1}$ for all i, j and $n = \sum_{i=1}^k n_i$. Thus,

$$\sigma_n^2 = \frac{1}{2}\Big[\log\big(1 - \frac{p}{n}\big) - \sum_{i=1}^k \frac{n_i^2}{n^2}\log\big(1 - \frac{p}{n_i}\big)\Big] + O(\frac{p}{n^2})$$

as $p \to \infty$. By the Taylor expansion, $\log(1 - \frac{p}{n_i}) = -\frac{p}{n_i} - \frac{p^2}{2n_i^2} + O(\frac{p^3}{n^3})$ for $1 \le i \le k$, and $\log(1 - \frac{p}{n}) = -\frac{p}{n} - \frac{p^2}{2n^2} + O(\frac{p^3}{n^3})$. Then, we use the fact $n = \sum_{i=1}^k n_i$ to see

$$\sum_{i=1}^k \frac{n_i^2}{n^2}\log\big(1 - \frac{p}{n_i}\big) = -\frac{p}{n} - \frac{kp^2}{2n^2} + O(\frac{p^3}{n^3}).$$

Thus,

$$\sigma_n^2 \sim \frac{(k-1)p^2}{4n^2} \qquad (5.18)$$

as $p \to \infty$.

Step 2. In this step we collect some "little" facts for the main proof in step 3. Fix s such that $|s| \le 1$. Set $t = t_n = \frac{s}{n\sigma_n}$. Then

$$t \sim \frac{2s}{\sqrt{k-1}} \cdot \frac{1}{p} \qquad (5.19)$$

as $p \to \infty$. We claim

$$n^2 t^2 \log\big(1 - \frac{p}{n-1}\big) = n^2 t^2 \log\big(1 - \frac{p}{n}\big) + o(1) \qquad (5.20)$$

$$\Big[(2n - 2p - 3)n\log(1 - \frac{p}{n-1})\Big]t = \Big[(2n-2p-3)n\log(1 - \frac{p}{n})\Big]t - 2pt + o(1) \qquad (5.21)$$

as $p \to \infty$. Similar to (5.17) we have that

$$\log\big(1 - \frac{p}{n-1}\big) - \log\big(1 - \frac{p}{n}\big) = -\frac{p}{n(n-p)} + O(\frac{p}{n^3})$$

which joint with (5.19) gives (5.20). Further, the difference between the two sides of (5.21) is equal to

$$\begin{aligned}
& (2n - 2p - 3)nt\Big(-\frac{p}{n(n-p)} + O(\frac{p}{n^3})\Big) + 2pt + o(1) \\
= \;& -2pt + \frac{3pt}{n-p} + 2pt + O(\frac{1}{n}) + o(1) = o(1)
\end{aligned}$$

as $p \to \infty$. This concludes (5.21).

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

$$E \exp\Big\{\frac{\log \Lambda_n - \mu_n}{n\sigma_n} s\Big\} \to e^{s^2/2}$$

as $p \to \infty$ for all $|s| \le 1$. Review (2.8) and (5.16). We have

$$\log \Lambda_n = \log \lambda_n + z_n$$

where

$$z_n = \frac{1}{2} pn \log n - \frac{1}{2} \sum_{i=1}^{k} pn_i \log n_i. \tag{5.22}$$

Therefore, recalling the notation $t = \frac{s}{n\sigma_n}$, we only need to show

$$\log E(\lambda_n^t) = (\mu_n - z_n)t + \frac{s^2}{2} + o(1) \tag{5.23}$$

as $p \to \infty$. From (5.19) we know $t > \max_{1 \le i \le k}\{\frac{p}{n_i}\} - 1$ as p is sufficiently large. Thus, by Lemma 5.6,

$$E(\lambda_n^t) = \frac{\Gamma_p(\frac{1}{2}(n-1))}{\Gamma_p(\frac{1}{2}n(1+t) - \frac{1}{2})} \cdot \prod_{i=1}^{k} \frac{\Gamma_p(\frac{1}{2}n_i(1+t) - \frac{1}{2})}{\Gamma_p(\frac{1}{2}(n_i - 1))} \tag{5.24}$$

as p is large. Write $\frac{1}{2}n_i(1+t) - \frac{1}{2} = \frac{n_i - 1}{2} + \frac{n_i t}{2}$. We obtain from (5.19) and Proposition 5.1 that

$$\begin{aligned}
\log \frac{\Gamma_p(\frac{1}{2}n_i(1+t) - \frac{1}{2})}{\Gamma_p(\frac{1}{2}(n_i - 1))} &= -\frac{1}{4}\big[4pn_i + (2n_i - 2p - 3)n_i \log(1 - \frac{p}{n_i - 1})\big]t \\
&\quad -\frac{1}{4}n_i^2\big[\frac{p}{n_i - 1} + \log\big(1 - \frac{p}{n_i - 1}\big)\big]t^2 + \rho_{n_i}(t) + o(1)
\end{aligned}$$

where $\rho_l(t)$ is defined as in Lemma 5.7. Similarly,

$$\begin{aligned}
\log \frac{\Gamma_p(\frac{1}{2}n(1+t) - \frac{1}{2})}{\Gamma_p(\frac{1}{2}(n - 1))} &= -\frac{1}{4}\big[4pn + (2n - 2p - 3)n \log(1 - \frac{p}{n - 1})\big]t \\
&\quad -\frac{1}{4}n^2\big[\frac{p}{n - 1} + \log\big(1 - \frac{p}{n - 1}\big)\big]t^2 + \rho_n(t) + o(1) \\
&= -\frac{1}{4}\big[4pn - 2p + (2n - 2p - 3)n \log(1 - \frac{p}{n})\big]t \\
&\quad -\frac{1}{4}n^2\big[\frac{p}{n - 1} + \log\big(1 - \frac{p}{n}\big)\big]t^2 + \rho_n(t) + o(1),
\end{aligned}$$

where the second identity is obtained by (5.20) and (5.21). Note that $n_i^2 \cdot \frac{p}{n_i - 1} = (n_i + 1)p + O(n^{-1}p)$ for all $1 \le i \le k$ and the same holds when replacing n_i with n. We have from the fact $n = n_1 + \cdots + n_k$ that

$$\Big[\frac{1}{4}n^2 \cdot \frac{p}{n-1} - \sum_{i=1}^{k} \frac{1}{4} n_i^2 \cdot \frac{p}{n_i - 1}\Big]t^2 = \frac{1}{4}((1-k) + O(\frac{1}{n}))pt^2 = o(1)$$

as $p \to \infty$. Join (5.24), the above assertions and Lemma 5.7 to have

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

Reviewing the notation z_n in (5.22), the definitions of μ_n and σ_n^2, we finally see that

$$\begin{aligned}
&\log E(\lambda_n^t)\\
= \ & \frac{1}{4}\Big[-2kp-(2n-2p-3)nr_n^2+\sum_{i=1}^{k}(2n_i-2p-3)n_ir_{n_i'}^2\Big]t-z_nt\\
&+\frac{n^2\sigma_n^2}{2}t^2+o(1)=(\mu_n-z_n)t+\frac{1}{2}s^2+o(1)
\end{aligned}$$

as $p \to \infty$, where we use the fact that $(\sum_{i=1}^{k}\frac{p}{n_i})/(n\sigma_n)=O(\sum_{i=1}^{k}\frac{1}{n_i})=O(\frac{1}{n})$ by (5.18) in the last step. This leads to (5.23). ■

5.5 Proof of Theorem 4

Let Λ_n^* be as in (2.11). Set

$$W_n=\frac{\prod_{i=1}^{k}|\mathbf{A}_i|^{(n_i-1)/2}}{|\mathbf{A}|^{(n-k)/2}}=\Lambda_n^*\cdot\frac{\prod_{i=1}^{k}(n_i-1)^{(n_i-1)p/2}}{(n-k)^{(n-k)p/2}}. \tag{5.25}$$

We have the following result.

Lemma 5.8 *(p. 302 from Muirhead (1982)) Assume $n_i > p$ for $1 \le i \le k$. Under H_0 in (2.10),*

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

for all $t > \max_{1\le i\le k}\frac{p-1}{n_i-1}-1$, where $\Gamma_p(x)$ is defined as in (5.1).

Proof of Theorem 4. Since $0 \le p/n_i < 1$ for each i, by using an argument on subsequence principle similar to that in the first paragraph of the proof of Theorem 1, it suffices to prove the theorem provided $\lim p/n_i = y_i \in [0,1]$. The theorem is the same as Theorem 4 from Jiang and Yang (2013) when $y_i \in (0,1]$. So we only need to prove the theorem for the case $y_1=\cdots=y_k=0$.

First, according to (5.25), write

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

It has been shown that $\sigma_n > 0$ in Theorem 4 from Jiang and Yang (2013). To prove the theorem, it is enough to show

$$\frac{\log W_n - \mu_n'}{(n-k)\sigma_n} \quad \text{converges to} \quad N(0,1) \tag{5.26}$$

in distribution as $p \to \infty$, where

$$\mu_n' = \mu_n + \sum_{i=1}^{k} \frac{(n_i - 1)p}{2} \log(n_i - 1) - \frac{(n-k)p}{2} \log(n-k). \tag{5.27}$$

Assertion (5.26) is proved through the following steps.

Step 1. In this step we collect some useful facts for the main proof in Step 2. First, from the relation $n = n_1 + \cdots + n_k$ we see that

$$\begin{aligned} &\sum_{i=1}^{k} \Big(\frac{n_i - 1}{n-k}\Big)^2 \Big(\frac{p}{n_i - 1} + \frac{p^2}{2(n_i - 1)^2}\Big) \\ = &\frac{p}{(n-k)^2} \sum_{i=1}^{k} (n_i - 1) + \frac{kp^2}{2(n-k)^2} = \frac{p}{n-k} + \frac{kp^2}{2(n-k)^2}. \end{aligned}$$

Therefore, by using the fact $\log(1-x) = -x - \frac{x^2}{2} + O(x^3)$ as $x \to 0$, we have

$$\begin{aligned} 2\sigma_n^2 &= -\frac{p}{n-k} - \frac{p^2}{2(n-k)^2} + \sum_{i=1}^{k} \Big(\frac{n_i - 1}{n-k}\Big)^2 \Big(\frac{p}{n_i - 1} + \frac{p^2}{2(n_i - 1)^2}\Big) + O\big(\frac{p^3}{n^3}\big) \\ &= \frac{(k-1)p^2}{2(n-k)^2} + O\big(\frac{p^3}{n^3}\big). \end{aligned}$$

Since k is fixed,

$$\sigma_n \sim \frac{\sqrt{k-1}}{2} \cdot \frac{p}{n}$$

as $p \to \infty$. Second, fix s such that $|s| \le 1$. Set $t = t_n = \frac{s}{(n-k)\sigma_n}$. Then

$$t \sim \frac{2}{\sqrt{k-1}} \cdot \frac{s}{p}$$

as $p \to \infty$.

Step 2. To prove (5.26), it suffices to prove

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

as $p \to \infty$ for all $|s| \le 1$. Recall $t = \frac{s}{(n-k)\sigma_n}$. It is equivalent to showing that

$$\log E(W_n^t) = \mu_n' t + \frac{s^2}{2} + o(1) \tag{5.28}$$

as $p\to\infty$. By Lemma 5.8,

$$\log E(W_n^t) = -\log\frac{\Gamma_p\left(\frac{1}{2}(n-k)(1+t)\right)}{\Gamma_p\left(\frac{1}{2}(n-k)\right)} + \sum_{i=1}^{k}\log\frac{\Gamma_p\left(\frac{1}{2}(n_i-1)(1+t)\right)}{\Gamma_p\left(\frac{1}{2}(n_i-1)\right)}$$

as p is sufficiently large. Write $\frac{1}{2}(n_i-1)(1+t) = \frac{1}{2}(n_i-1)+\frac{1}{2}(n_i-1)t$. By Proposition 5.1,

$$\begin{aligned}\log\frac{\Gamma_p\left(\frac{1}{2}(n_i-1)(1+t)\right)}{\Gamma_p\left(\frac{1}{2}(n_i-1)\right)} &= -\frac{1}{2}\big[2p+(n_i-p-\frac{3}{2})\log(1-\frac{p}{n_i-1})\big](n_i-1)t\\ &\quad -\frac{1}{4}\big[\frac{p}{n_i-1}+\log\big(1-\frac{p}{n_i-1}\big)\big](n_i-1)^2t^2\\ &\quad +p\Big[\frac{(n_i-1)(1+t)}{2}\log\frac{(n_i-1)(1+t)}{2}-\frac{n_i-1}{2}\log\frac{n_i-1}{2}\Big]+o(1).\end{aligned}$$

By reorganizing the right hand side, we obtain

$$\begin{aligned}&\log\frac{\Gamma_p\left(\frac{1}{2}(n_i-1)(1+t)\right)}{\Gamma_p\left(\frac{1}{2}(n_i-1)\right)}\\ =\;& -t(n_i-1)p-\frac{t}{4}(n_i-1)(2n_i-2p-3)\log(1-\frac{p}{n_i-1})\\ & -\frac{t^2}{4}(n_i-1)p-\frac{t^2}{4}(n_i-1)^2\log\big(1-\frac{p}{n_i-1}\big)\\ & +(pt)\frac{n_i-1}{2}\log\frac{n_i-1}{2}+\frac{(n_i-1)p}{2}\cdot(1+t)\log(1+t)+o(1).\end{aligned}$$

Replace n_i with $n-k+1$ to get

$$\begin{aligned}&\log\frac{\Gamma_p\left(\frac{1}{2}(n-k)(1+t)\right)}{\Gamma_p\left(\frac{1}{2}(n-k)\right)}\\ =\;& -t(n-k)p-\frac{t}{4}(n-k)(2(n-k)-2p-1)\log(1-\frac{p}{n-k})\\ & -\frac{t^2}{4}(n-k)p-\frac{t^2}{4}(n-k)^2\log\big(1-\frac{p}{n-k}\big)\\ & +(pt)\frac{n-k}{2}\log\frac{n-k}{2}+\frac{(n-k)p}{2}\cdot(1+t)\log(1+t)+o(1).\end{aligned}$$

Combining the above two identities, we have from the fact $n=n_1+\cdots+n_k$ that

$$\begin{aligned}\log E(W_n^t) &= \mu_n t+\frac{t^2(n-k)^2\sigma_n^2}{2}-pt\Big[\frac{n-k}{2}\log\frac{n-k}{2}-\sum_{i=1}^{k}\frac{n_i-1}{2}\log\frac{n_i-1}{2}\Big]+o(1)\\ &= \Big[\mu_n-\frac{1}{2}(n-k)p\log(n-k)+\sum_{i=1}^{k}\frac{1}{2}(n_i-1)p\log(n_i-1)\Big]t+\frac{s^2}{2}+o(1)\\ &= \mu_n' t+\frac{s^2}{2}+o(1)\end{aligned}$$

as $p\to\infty$ by (5.27). This leads to (5.28). ■

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 (2.13), given by*

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

is unbiased, where $\overline{\mathbf{x}}$ and $\mathbf{A}$ are as in (2.14). Further, assuming H_0 in (2.13), we have

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

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

The range "$t > \frac{p}{n} - 1$" follows from the definition of $\Gamma_p(z)$ in (5.1).

Lemma 5.10 *Let $p = p_n \to \infty$ and $p = o(n)$. For $s \in \mathbb{R}$, let $t = t_n = O(1/p)$. Then*

$$\begin{aligned} \eta_n(t) : &= \big(\frac{n-1}{2} + \frac{nt}{2}\big) p \log\big(\frac{n-1}{2} + \frac{nt}{2}\big) - \frac{(n-1)p}{2} \log \frac{n-1}{2} \\ &= \frac{npt}{2} \log \frac{n}{2} + \frac{np(1+t)}{2} \log(1+t) - \frac{pt}{2} + o(1) \end{aligned}$$

as $n \to \infty$.

Proof. First,

$$\begin{aligned} \log\big(\frac{n-1}{2} + \frac{nt}{2}\big) &= \log \frac{n}{2} + \log(1+t) + \log\big(1 - \frac{1}{n(1+t)}\big) \\ &= \log \frac{n}{2} + \log(1+t) - \frac{1}{n(1+t)} + O\big(\frac{1}{n^2}\big) \end{aligned}$$

as $n \to \infty$. And $\log \frac{n-1}{2} = \log \frac{n}{2} - \frac{1}{n} + O(n^{-2})$. Write $\big(\frac{n-1}{2} + \frac{nt}{2}\big)p = \frac{np(1+t)}{2} - \frac{p}{2}$. Then,

$$\begin{aligned} \eta_n(t) &= \big(\frac{np(1+t)}{2} - \frac{p}{2}\big)\big(\log \frac{n}{2} + \log(1+t) - \frac{1}{n(1+t)}\big) - \frac{(n-1)p}{2} \log \frac{n}{2} + \frac{p}{2} + O\big(\frac{p}{n}\big) \\ &= \frac{npt}{2} \log \frac{n}{2} + \frac{np(1+t)}{2} \log(1+t) - \frac{p}{2} \log(1+t) + O\big(\frac{p}{n}\big) \\ &= \frac{npt}{2} \log \frac{n}{2} + \frac{np(1+t)}{2} \log(1+t) - \frac{pt}{2} + O\big(\frac{1}{p} + \frac{p}{n}\big) \end{aligned}$$

as $n \to \infty$ by the definition of t. ■

Proof of Theorem 5. By Theorem 5 from Jiang and Yang (2013), we know that the theorem is true if $\lim p/n = y \in (0, 1]$. By the subsequence argument as in the first paragraph of the proof of Theorem 1, to prove the theorem, it suffices to prove it for the case $\lim p/n = 0$. From now on, we assume $p \to \infty$ and $p = o(n)$.

In order to get the desired conclusion, it is enough to show that

$$E\exp\Big\{\frac{\log\Lambda_n-\mu_n}{n\sigma_n}\,s\Big\}\to e^{s^2/2}$$

as $n\to\infty$ for all $|s|\le 1$, or equivalently,

$$\log E(\Lambda_n^t)=\mu_n t+\frac{1}{2}s^2+o(1) \tag{5.29}$$

as $n\to\infty$ for all $|s|\le 1$, where $t=t_n=\frac{s}{n\sigma_n}$. Evidently, $\sigma_n\sim\frac{p}{2(n-1)}$, and hence $t\sim\frac{2s}{p}>\frac{p}{n}-1$ as n is large enough. It follows from Lemma 5.9 that

$$\log E(\Lambda_n^t)=\frac{npt}{2}\big(1+\log\frac{2}{n}\big)-\frac{np(1+t)}{2}\log(1+t)+\log\frac{\Gamma_p(\frac{n(1+t)-1}{2})}{\Gamma_p(\frac{n-1}{2})}$$

for large n. Now, write $\frac{n(1+t)-1}{2}=\frac{n-1}{2}+\frac{nt}{2}$. Then, by Proposition 5.1,

$$\log\frac{\Gamma_p(\frac{n(1+t)-1}{2})}{\Gamma_p(\frac{n-1}{2})}=\frac{1}{2}n\alpha_n t+\frac{1}{4}(n^2\beta_n)t^2+\eta_n(t)+o(1)$$

where

$$\alpha_n=-\big[2p+(n-p-\frac{3}{2})\log(1-\frac{p}{n-1})\big],\quad \beta_n=-\big[\frac{p}{n-1}+\log\big(1-\frac{p}{n-1}\big)\big],$$
$$\eta_n(t)=p\Big[\big(\frac{n-1}{2}+\frac{nt}{2}\big)\log\big(\frac{n-1}{2}+\frac{nt}{2}\big)-\frac{n-1}{2}\log\frac{n-1}{2}\Big].$$

Further, by Lemma 5.10,

$$\eta_n(t)=\frac{npt}{2}\log\frac{n}{2}+\frac{np(1+t)}{2}\log(1+t)-\frac{pt}{2}+o(1)$$

as $n\to\infty$. Joining all of the above equalities, we finally see that

$$\log E(\Lambda_n^t)=-\frac{1}{4}\big(-2n\alpha_n-2p(n-1)\big)t+\frac{1}{4}n^2\beta_n t^2+o(1)=\mu_n t+\frac{s^2}{2}+o(1)$$

as $n\to\infty$, which yields (5.29). ■

5.7 Proof of Theorem 6

Lemma 5.11 *Let $\hat{\mathbf{R}}_n$ be the sample correlation matrix with the density function as in (2.20). Assume $n-4>p\ge 2$. Then,*

$$E[|\hat{\mathbf{R}}_n|^t]=\Big[\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2}+t)}\Big]^p\cdot\frac{\Gamma_p(\frac{n-1}{2}+t)}{\Gamma_p(\frac{n-1}{2})} \tag{5.30}$$

for all $t\ge-\max\big\{1,\lfloor\frac{n-p}{2}\rfloor-2\big\}$.

This lemma is a refinement of Lemma 2.10 from Jiang and Yang (2013) who show that (5.30) holds for all $t \geq -1$.

Proof. By Lemma 2.10 from Jiang and Yang (2013), (5.30) is true for $t \geq -1$. Thus, we only need to prove that (5.30) holds for $t \geq m := -\lfloor \frac{n-p}{2} \rfloor + 2$. It is a key observation that $m \leq 0$ by the assumption $n - 4 > p$.

Recall (2.19), $\hat{\mathbf{R}}_n$ is a $p \times p$ non-negative definite matrix and each of its entries takes value in $[-1, 1]$, thus the determinant $|\hat{\mathbf{R}}_n| \leq p!$. Second, from (2.20) we see that the density function of $|\hat{\mathbf{R}}_n|$ exists, hence, $P(|\hat{\mathbf{R}}_n| = 0) = 0$. By (9) on p. 150 from Muirhead (1982) or (48) on p. 492 from Wilks (1932),

$$E\big[|\hat{\mathbf{R}}_n|^k\big] = \Big[\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2} + k)}\Big]^p \cdot \frac{\Gamma_p(\frac{1}{2}(n-1) + k)}{\Gamma_p(\frac{1}{2}(n-1))} \tag{5.31}$$

for any integer k such that $\frac{n-1}{2} + k > \frac{p-1}{2}$ by (5.1), which is equivalent to that $k > -(n-p)/2$. Thus, (5.31) holds for all $k \geq m - 1$. In particular, $E\big[|\hat{\mathbf{R}}_n|^{m-1}\big] < \infty$. Now set $U = -\log(|\hat{\mathbf{R}}_n|/p!)$. Then $U \geq 0$ *a.s.* and $Ee^{(1-m)U} < \infty$. Since $|e^{-(z+m)U}| = e^{-(\mathrm{Re}(z)+m)U}$ and $|Ue^{-(z+m)U}| = Ue^{-(\mathrm{Re}(z)+m)U}$, they imply that

$$Ee^{-(z+m)U} \quad \text{and} \quad E\big(Ue^{-(z+m)U}\big) \quad \text{are both finite}$$

for all $\mathrm{Re}(z) \geq 0$, where we use the inequalities $e^{-mu} \leq e^{(1-m)u}$ and $ue^{-mu} \leq e^{(1-m)u}$ for all $u \geq 0$ to get the second assertion. Define

$$h_1(z) := (p!)^{-(z+m)} \cdot E\big[|\hat{\mathbf{R}}_n|^{z+m}\big] = Ee^{-(z+m)U}$$

for all z with $\mathrm{Re}(z) \geq 0$. It is not difficult to check that $\frac{d}{dz}\left(Ee^{-(z+m)U}\right) = -E\big[Ue^{-(z+m)U}\big]$ for all $\mathrm{Re}(z) \geq 0$. Further, $\sup_{\mathrm{Re}(z)\geq 0} |h_1(z)| \leq Ee^{-mU} < \infty$. Therefore, $h_1(z)$ is analytic and bounded on $\{z \in \mathbb{C};\, \mathrm{Re}(z) \geq 0\}$. Define

$$h_2(z) = (p!)^{-(z+m)} \cdot \Big[\frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-1}{2} + z + m)}\Big]^p \cdot \frac{\Gamma_p(\frac{n-1}{2} + z + m)}{\Gamma_p(\frac{1}{2}(n-1))}$$

for $\mathrm{Re}(z) \geq 0$. By the Carlson uniqueness theorem (see, for example Theorem 2.8.1 on p. 110 from Andrews et al. (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 $z = 0, 1, 2, \cdots$, 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 (5.1), it suffices to show

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

is bounded and analytic on $\{z \in \mathbb{C};\, \mathrm{Re}(z) \geq 0\}$. Since $2 \leq \frac{n-i}{2} + m \leq \frac{n-2}{2} + m$ for all $2 \leq i \leq p$, the two properties then follow from the fact that $h(z) := \frac{\Gamma(\alpha+z)}{\Gamma(\beta+z)}$ is bounded and analytic on $\{z \in \mathbb{C}; \operatorname{Re}(z) \geq 0\}$ for all fixed $\beta > \alpha > 0$ by Lemma 3.1 from Jiang and Yang (2013). ■

Proof of Theorem 6. By using the subsequence argument as in the first paragraph of the proof of Theorem 1, we only need to show the theorem when $\lim_{n\to\infty} p/n = y \in [0, 1]$. The case for $y \in (0, 1]$ is proved by Jiang and Yang (2013). We will prove the theorem for the case $y = 0$ next.

To finish the proof, it suffices to show that

$$E\exp\Big\{\frac{\log|\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n}s\Big\} = \exp(-\frac{\mu_n s}{\sigma_n}) \cdot E[|\hat{\mathbf{R}}_n|^{\frac{s}{\sigma_n}}] \to e^{s^2/2}$$

as $n \to \infty$ for all s with $|s| \leq 1$, or equivalently,

$$\log E[|\hat{\mathbf{R}}_n|^t] = \mu_n t + \frac{s^2}{2} + o(1) \tag{5.32}$$

as $n \to \infty$ for all $|s| \leq 1$, where $t := \frac{s}{\sigma_n}$. It is easy to see that

$$\sigma_n \sim \frac{p}{n} \quad \text{and} \quad t \sim \frac{n}{p}s$$

as $n \to \infty$. In particular, $t \geq -\max\big\{1, \lfloor (n-p)/2 \rfloor - 2\big\}$ as n is large enough. By Lemma 5.11,

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

if n is sufficiently large. By (5.2), we have

$$p\,\log\Big[\frac{\Gamma(\frac{n-1}{2}+t)}{\Gamma(\frac{n-1}{2})}\Big] = \gamma_n(t) - \frac{npt}{n-1} + O(\frac{1}{p})$$

where

$$\gamma_n(t) = p\Big[\big(\frac{n-1}{2}+t\big)\log\big(\frac{n-1}{2}+t\big) - \frac{n-1}{2}\log\frac{n-1}{2}\Big].$$

By the fact $t \sim \frac{n}{p}s$ and Proposition 5.1,

$$\log\frac{\Gamma_p(\frac{n-1}{2}+t)}{\Gamma_p(\frac{n-1}{2})} = -\big[2p + (n-p-\frac{3}{2})\log(1-\frac{p}{n-1})\big]t + \frac{\sigma_n^2 t^2}{2} + \gamma_n(t) + o(1)$$

as $n \to \infty$. Connecting the above assertions, we arrive at

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

as $n \to \infty$. So we get (5.32) and complete the proof. ■

Acknowledgements. We thank Fan Yang and Lin Zhang very much for their help in checking the mathematical proofs in this paper.

References

[1] Ahlfors, L. V. (1979). Complex Analysis. *McGraw-Hill, Inc.*, 3rd Ed.

[2] Andrews, G. E., Askey, R. and Roy, R. (1999). Special Functions. *Cambridge University Press.*

[3] Bai, Z., Jiang, D., Yao, J. and Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Stat.* 37, 3822-3840.

[4] Bartlett, M. S. (1937). Properties and sufficiency and statistical tests. *Proc. R. Soc. Lond. A* 160, 268-282.

[5] Bartlett, M. S. (1954). A note on multiplying factors for various chi-squared approximations. *J. Royal Stat. Soc., Ser. B* 16, 296-298.

[6] Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika* 36, 317-346.

[7] Cai, T., Liu, W. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. American Statistical Association* 108, 265-277.

[8] Cai, T. and Ma, Z. (2012). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli*, to appear.

[9] Chen, S., Zhang, L., and Zhong, P. (2010). Tests for high dimensional covariance matrices. *J. Amer. Stat. Assoc.* 105, 810-819.

[10] Fujikoshi, Y., Ulyanov, V. V. and Shimizu, R. (2010). Multivariate Statistics: High-dimensional and large-sample approximations. *Wiley.*

[11] Gamelin, T. W. (2001). Complex Analysis. *Springer*, 1st Ed.

[12] Jiang, D., Jiang, T. and Yang, F. (2012). Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. *J. Stat. Plann. Inference* 142, 2241-2256.

[13] Jiang, T. and Yang, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. Stat.* 41, 2029-2074.

[14] Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Stat.* 30, 1081-1102.

[15] Li, J. and Chen, S. X. (2012). Two sample tests for high dimensional covariance matrices. *Ann. Stat.* 40, 908-940.

[16] Lopes, M. E., Jacob, L. J. and Wainwright, M. J. A more powerful two-sample test in high dimensions using random projection. *http://arxiv.org/abs/1108.2401.*

[17] Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Stat.* 11, 204-209.

[18] Morrison, D. F. (2004). Multivariate Statistical Methods. *Duxbury Press*, 4th Ed.

[19] Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory. *Wiley, New York.*

[20] Onatski, A., Moreira, M. J. and Hallin, M. Asymptotic power of sphericity tests for high-dimensional data. *http://www.econ.cam.ac.uk/faculty/onatski/pubs/WPOnatskiMoreira.pdf.*

[21] Perlman, M. D. (1980). Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations. *Ann. Stat.* 8, 247-263.

[22] Qiu, Y-M and Chen, S. X. (2012). Test for bandedness of high dimensional covariance matrices with bandwidth estimation. *Ann. Stat.* 40, 1285-1314.

[23] Schott, J. R. (2001). Some tests for the equality of covariance matrices. *J. Stat. Plann. Inference* 94, 25-36.

[24] Schott, J. R. (2005). Testing for complete independence in high dimensions. *Biometrika* 92, 951-956.

[25] 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.

[26] Serdobolskii, V. I. (2000). Multivariate Statistical Analysis: A High-Dimensional Approach. *Springer.*

[27] Srivastava, M. S. (2005). Some tests concerning covariance matrix in high dimensonal data. *J. Japan Statist. Soc.* 35, 251-272.

[28] Sugiura, N. and Nagao, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *Ann. Math. Stat.* 39, 1686-1692.

[29] van der Vaart, A. W. (1998). Asymptotic Statistics. *Cambridge University Press.*

[30] Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika* 24, 471-494.

[31] Wilks, S. S. (1935). On the independence of k sets of normally distributed statistical variables. *Econometrica* 3, 309-326.

[32] Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.* 9, 60-62.

[33] Xiao, H. and Wu, W. (2013). Asymptotic theory for maximum deviations of sample covariance matrix estimates. *Stochastic Processes and their Applications* 123, 2899-2920.