

Limiting Distributions of the Likelihood Ratio Test Statistics for Independence of Normal Random Vectors

Mingyue Hu^{a,b}, Yongcheng Qi^a

^aDepartment of Mathematics and Statistics, University of Minnesota Duluth, 1117 University Drive, Duluth, MN 55812, USA

^bSchool of Mathematics and Statistics, Southwest University, 400715 Chongqing, China

Abstract. Consider the likelihood ratio test (LRT) statistics for the independence of sub-vectors from a p -variate normal random vector. We are devoted to deriving the limiting distributions of the LRT statistics based on a random sample of size n . It is well known that the limit is chi-square distribution when the dimension of the data or the number of the parameters are fixed. In a recent work by Qi, Wang and Zhang (Ann Inst Stat Math (2019) 71: 911–946), it was shown that the LRT statistics are asymptotically normal under condition that the lengths of the normal random sub-vectors are relatively balanced if the dimension p goes to infinity with the sample size n . In this paper, we investigate the limiting distributions of the LRT statistic under general conditions. We find out all types of limiting distributions and obtain the necessary and sufficient conditions for the LRT statistic to converge to a normal distribution when p goes to infinity. We also investigate the limiting distribution of the adjusted LRT test statistic proposed in Qi, Wang and Zhang (2019). Moreover, we present simulation results to compare the performance of classical chi-square approximation, normal and non-normal approximation to the LRT statistics, chi-square approximation to the adjusted test statistic, and some other test statistics.

Keywords: Likelihood ratio test; normal random vector; central limit theorem; chi-square approximation; non-normal limit; high dimension; independence

AMS 2020 Subject Classification: 62E20; 62H15

1 Introduction

In classical theory for inference, many statistical methods have been developed to test parametric hypotheses in the last few decades. One of the most popular methods is the so-called likelihood ratio test (LRT). It is known that the distribution functions of LRT statistics can be well approximated by a chi-square distribution under certain regularity conditions when the dimension of the data or the number of the parameters of interest are fixed. This means that one does not have to estimate the variance of the test statistics based on the likelihood ratio.

Many modern data sets such as financial data and modern manufacturing data are high-dimensional. Classical methods may not be adequate for high dimensional data anymore, especially when the dimension of data is relatively large compared with the sample size. Some recent papers have investigated the limiting distribution of the LRT statistics concerning the dependence structures of the multivariate normal distributions. It turns out that the chi-square approximation fails while the dimension of the data increases with the sample size. Instead, the normal approximation to the LRT statistics works well under high dimension setting. See, e.g., Bai et al. [?], Jiang et al. [?], Jiang and Yang [?], Jiang and Qi [?], Qi et al. [?], Dette and Dörnemann [?], and Guo and Qi [?]. Several approaches other than likelihood ratio method have been developed in the literature; See, e.g., Schott [?, ?, ?], Ledoit and Wolf [?], Bao et al. [?], Chen et al. [?], Srivastava and Reid [?], Jiang et al [?], Li et al. [?], and Bodnar et al. [?]. A very recent work by Dörnemann [?] also established the central limit theorems for some LRT statistics under non-normality.

In this paper, we consider a p -variate normal random vector and study the limiting distributions of the LRT for testing the independence of its grouped components based on a random sample of size n . For a p -dimensional multivariate normal distribution with mean vector μ and covariance matrix Σ , denoted by $N_p(\mu, \Sigma)$, we partition its p components into k subsets and test whether the k sub-vectors are mutually independent, or equivalently, we test whether the covariance matrix Σ is block diagonal. In this paper, both p and k can depend on n and diverge with the sample size. On the condition that the lengths of the k sub-vectors are relatively balanced, Qi et al. [?] proved the asymptotic normality and proposed an adjusted test statistic that usually has a chi-square limit when the dimension p goes to infinity with the sample size.

The aim of this paper is to give a complete description for the limiting distributions of the LRT statistic for independence for multivariate normal random vectors. We obtain all possible limiting distributions and give the necessary and sufficient conditions for the central limit theorem. We also investigate the limiting distributions of the adjusted test statistic proposed by Qi et al. [?], which performs better than the normal approximation and chi-square approximation to the LRT statistics in general.

The rest of the paper is organized as follows. In Section ??, we present our main results and establish the necessary and sufficient conditions for the central limit theorem as well as the conditions for non-normal

limits. In Section ??, we present some simulation results to compare the performance of four methods including the chi-square approximation to the LRT statistic, the normal and non-normal approximation to the LRT statistic and the chi-square approximation to the adjusted LRT statistic proposed in Qi et al. [?]. In Section ??, we present some preliminary lemmas which are used in Section ?? to prove the main results of the paper.

2 Main results

Let χ_f^2 denote the random variables following chi-square distribution with f degrees of freedom and $N(0, 1)$ the standard normal variables.

For $k \geq 2$, let $q_1, \dots, q_k$ be k positive integers. Denote $p = q_1 + q_2 + \dots + q_k$ and let

$$\Sigma = (\Sigma_{ij})_{1 \leq i, j \leq k}$$

be a positive definite matrix, where Σ_{ij} is a $q_i \times q_j$ sub-matrix. Assume ξ_i is a q_i -dimensional normal random vector for each $1 \leq i \leq k$, and the p -dimensional random vector $(\xi'_1, \dots, \xi'_k)'$ has a multivariate normal distribution $N_p(\mu, \Sigma)$. We are interested in testing the independence of k random vectors $\xi_1, \dots, \xi_k$, or equivalently the following hypotheses

$$H_0 : \Sigma_{ij} = 0 \text{ for all } 1 \leq i < j \leq k \quad \text{vs} \quad H_1 : H_0 \text{ is not true.} \quad (2.1)$$

Assume that $\mathbf{x}_1, \dots, \mathbf{x}_n$ are n independent and identically distributed random vectors from distribution $N_p(\mu, \Sigma)$. Define

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

and partition $\mathbf{A}$ as follows

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

where $\mathbf{A}_{ij}$ is a $q_i \times q_j$ matrix. According to Theorem 11.2.3 from Muirhead [?], the likelihood ratio statistic for testing (??) is given by

$$\Lambda_n = \frac{|\mathbf{A}|^{\frac{n}{2}}}{\prod_{i=1}^k |\mathbf{A}_{ii}|^{\frac{n}{2}}} := (W_n)^{\frac{n}{2}}. \quad (2.2)$$

Note that the likelihood ratio statistic Λ_n is well defined only if $p < n$. When $p \geq n$, the determinant $|\mathbf{A}|$ is zero since $\mathbf{A}$ is singular. Therefore, we can only consider the case $p < n$ in the paper.

We introduce some notations before we give the main results. Let g be any function defined over $(0, \infty)$. For integers q and n with $1 \leq q < n$, define

$$\Delta_{g,n,q}(t) = \sum_{i=1}^q g\left(\frac{n-i}{2} + t\right), \quad t > -\frac{n-q}{2}. \quad (2.3)$$

Let $\Gamma(x)$ denote the Gamma function, given by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0,$$

and define the digamma function ψ

$$\psi(x) = \frac{d \log \Gamma(x)}{dx} = \frac{\Gamma'(x)}{\Gamma(x)}, \quad x > 0. \quad (2.4)$$

Theorem 2.1. Assume $p = p_n$ satisfy $2 \leq p < n$ and $p_n \rightarrow \infty$ as $n \rightarrow \infty$. Assume that $q_1, q_2, \dots, q_k$ are k positive integers such that $p = \sum_{i=1}^k q_i$, where $k = k_n$ may depend on n . Set $q_{\max} = \max\{q_1, \dots, q_k\}$ and assume $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$. Λ_n is the Wilks likelihood ratio statistic defined in (??). Then, under the null hypothesis in (??)

$$T_0 := \frac{-2 \log \Lambda_n - \mu_n}{\sigma_n} \xrightarrow{d} N(0, 1) \quad (2.5)$$

as $n \rightarrow \infty$, where

$$\mu_n = -n \left(\Delta_{\psi,n,p}(0) - \sum_{j=1}^k \Delta_{\psi,n,q_j}(0) \right), \quad (2.6)$$

$$\sigma_n^2 = 2n^2 \left(\sum_{j=1}^p \frac{1}{n-j} - \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{n-j} \right) + 2n^2 \left(b(n,p) - \sum_{i=1}^k b(n, q_i) \right), \quad (2.7)$$

$$b(n, q) = \sum_{j=1}^q \frac{1}{(n-j)^2}, \quad 1 \leq q < n, \quad (2.8)$$

and symbol $\xrightarrow{d}$ denotes convergence in distribution.

Now we consider the situation when $n - q_{\max}$ is bounded. In this case, both $n - p$ and $p - q_{\max}$ are bounded because $n - p + p - q_{\max} = n - q_{\max}$. The following theorem gives non-normal limits for $-2 \log \Lambda_n$ when both $n - p$ and $p - q_{\max}$ are fixed integers.

Theorem 2.2. Assume $p = p_n$ satisfy $2 \leq p < n$. Assume $q_1, q_2, \dots, q_k$ are k positive integers such that $p = \sum_{i=1}^k q_i$, where $k = k_n$ is an integer that may depend on n . Set $q_{\max} = \max\{q_1, \dots, q_k\}$ and assume $p - q_{\max} = r$ and $n - p = v$ for some fixed integers $r \geq 1$ and $v \geq 1$ for all large n . Then, under the null hypothesis in (??) we have

$$\frac{-2 \log \Lambda_n - n \log n}{n} \xrightarrow{d} - \sum_{j=v}^{r+v-1} \log Y_j, \quad (2.9)$$

where Y_j , $j \geq 1$ are independent random variables and the Y_j has a chi-square distribution with j degrees of freedom.

Remark 1. The classical likelihood method considers the case when both p and k are fixed integers. Assume that $q_1, q_2, \dots, q_k$ are fixed for all large n , then

$$-2\rho_n \log \Lambda_n \xrightarrow{d} \chi_f^2, \quad (2.10)$$

where

$$f = \frac{1}{2} \left(p^2 - \sum_{i=1}^k q_i^2 \right), \quad (2.11)$$

$$\rho_n = 1 - \frac{2 \left(p^3 - \sum_{i=1}^k q_i^3 \right) + 9 \left(p^2 - \sum_{i=1}^k q_i^2 \right)}{6n \left(p^2 - \sum_{i=1}^k q_i^2 \right)}; \quad (2.12)$$

See, e.g., Theorem 11.2.5 in Muirhead [?].

Remark 2. Under conditions that $q_{\max} \leq \delta p$ for some $\delta \in (0, 1)$ and $p \rightarrow \infty$ as $n \rightarrow \infty$, Qi et al. [?] established the following central limit theorem for $-2 \log \Lambda_n$

$$T_1 := \frac{-2 \log \Lambda_n - \bar{\mu}_n}{\tau_n} \xrightarrow{d} N(0, 1)$$

as $n \rightarrow \infty$, where

$$\bar{\mu}_n = n \sum_{i=1}^k \left(q_i - n + \frac{3}{2} \right) \log \left(1 - \frac{q_i}{n} \right) - n \left(p - n + \frac{3}{2} \right) \log \left(1 - \frac{p}{n} \right) + \frac{n}{3} \left(b(n, p) - \sum_{i=1}^k b(n, q_i) \right), \quad (2.13)$$

$$\tau_n^2 = 2n^2 \left(\sum_{i=1}^k \log \left(1 - \frac{q_i}{n} \right) - \log \left(1 - \frac{p}{n} \right) \right) + 2n^2 \left(b(n, p) - \sum_{i=1}^k b(n, q_i) \right).$$

Remark 3. Theorem ?? is still true if μ_n is replaced by $\bar{\mu}_n$ defined in (??) and σ_n^2 is replaced by $\bar{\sigma}_n^2$

$$\bar{\sigma}_n^2 = 2n^2 \left(\sum_{j=1}^p \frac{1}{n-j} - \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{n-j} \right). \quad (2.14)$$

In fact, we have

$$\lim_{n \rightarrow \infty} \frac{\mu_n - \bar{\mu}_n}{\sigma_n} = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\bar{\sigma}_n^2}{\sigma_n^2} = 1 \quad (2.15)$$

if $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$. The proof of (??) is given in Lemma ??. In practice, μ_n and σ_n^2 should be used since they give better approximation to the mean and variance of $-2 \log \Lambda_n$, and this selection can achieve a better accuracy for the normal approximation even when $n - q_{\max}$ is not very large.

Remark 4. The distribution of the random variable on the right-hand side of (??) is non-normal. This can be verified by using the moment-generating functions. The moment generating function of $\sum_{j=v}^{r+v-1} \log Y_j$ is equal to

$$2^{rt} \prod_{j=v}^{r+v-1} \frac{\Gamma(j/2 + t)}{\Gamma(j/2)},$$

which cannot be equal to $\exp(\mu t + \sigma^2 t^2 / 2)$, the moment generating function of a normal random variable with a mean μ and variance σ^2 . Otherwise, after taking the logarithm, it implies that the third derivative of

$\sum_{j=v}^{r+v-1} \log \frac{\Gamma(j/2+t)}{\Gamma(j/2)}$ is identically equal to zero. We can show this cannot be true by using some properties of the gamma function. The details are omitted here.

Now we are ready to establish the necessary and sufficient conditions under which the central limit theorem holds for $-2 \log \Lambda_n$.

Theorem 2.3. *There exist constants $a_n \in R$ and $b_n > 0$ such that under the null hypothesis in (??)*

$$\frac{-2 \log \Lambda_n - a_n}{b_n} \xrightarrow{d} N(0, 1) \quad (2.16)$$

if and only if $p_n \rightarrow \infty$ and $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$.

From (??), (??) and (??), we have only three different types of limiting distributions for $-2 \log \Lambda_n$, including chi-square distributions, normal distributions and distributions of linear combinations of logarithmic chi-square random variables. Theoretically, for any p and q_i 's one can use one of the three limiting distributions to approach the distribution of $-2 \log \Lambda_n$ when n is large. Since the convergence in (??), (??) or (??) does not provide clear cutoff values for p and $n - q_{\max}$, it may be difficult to select a limiting distribution in practice even if n is very large.

Qi et al. [?] proposed an adjusted log-likelihood ratio test statistic (ALRT) which can be approximated by a chi-square distribution when (??) or (??) holds. The ALRT is a linear function of $-2 \log \Lambda_n$ defined as

$$Z_n = (-2 \log \Lambda_n) \sqrt{\frac{2f_n}{\sigma_n^2}} + f_n - \mu_n \sqrt{\frac{2f_n}{\sigma_n^2}} \quad (2.17)$$

with f_n , μ_n and σ_n^2 being defined in (??), (??) and (??), respectively. We have the following result on chi-square approximation to the distribution of Z_n .

Theorem 2.4. *Let $p = p_n$ be a sequence of integers with $2 \leq p < n$. Assume $k = k_n$ is also a sequence of positive integers, and $q_1, \dots, q_k$ are k positive integers such that $p = \sum_{i=1}^k q_i$. Assume $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$. Then, under the null hypothesis in (??), we have*

$$\lim_{n \rightarrow \infty} \sup_{-\infty < x < \infty} |P(Z_n \leq x) - P(\chi_{f_n}^2 \leq x)| = 0. \quad (2.18)$$

We note that the chi-square approximation in (??) does not impose any restriction on dimension p and the number of degrees of freedom of the chi-square distribution changes with n .

3 Simulation study

3.1 Comparison of LRT related tests

In this subsection, we carry out a finite-sample simulation study to compare the performance of three different approaches to the likelihood ratio test statistic $-2 \log \Lambda_n$ under condition $n - q_{\max} \rightarrow \infty$ and the

null hypothesis in (??), including the classical chi-square approximation (??), the normal approximation (??), and the adjusted chi-square approximation in (??). The three approaches are denoted by “Chisq”, “CLT” and “ALRT”, respectively. We will also compare the performance of the normal approximation (??), the adjusted chi-square approximation in (??) and the non-normal approximation given in (??) (denoted by “LogChi”), and the comparison is made under conditions that both $r = p - q_{\max}$ and $v = n - p$ are fixed integers.

For all four approaches, we will demonstrate how well the proposed limiting distributions fit the histograms of the four test statistics. From (??) in Lemma ??, the moment-generating function of $\log W_n$ is distribution-free under the null hypothesis in (??). Since the four test statistics are functions of Λ_n and hence they are also functions of $\log W_n$ from (??), they are distribution-free as well under the null hypothesis in (??). Therefore, the underlying distribution in our study is set to be a multivariate normal distribution with independent standard normal components.

In our simulation study, we choose sample size $n = 101$. For each of selected combinations of p and q_i ’s under the regimes of the chi-square and normal limiting distributions, we repeat the sampling for 10,000 times and obtain 10,000 replicates for the four test statistics given in (??), (??), (??), and (??). We plot the histogram for each test statistic and its corresponding theoretical density function in one graph.

In the simulation study we consider the following four cases.

Case a. We set $k = 3$, $p = 10, 30, 60$ and 90 and use the ratio $q_1 : q_2 : q_3 = 3 : 1 : 1$. Figure ?? contains 12 plots in an array with four rows and three columns, and each row corresponds to one value of $p = 10, 30, 60$, and 90 .

Case b. We set $k = 2$, $q_1 = p - 1$ and $q_2 = 1$ for $p = 10, 30, 60$, and 90 , respectively. Figure ?? contains 12 plots in an array with four rows and three columns, and each row corresponds to one value of $p = 10, 30, 60$, and 90 .

Case c. We set $k = 2$, $p = 100$ and choose $(q_1, q_2) = (50, 50), (60, 40), (80, 20)$, and $(90, 10)$, respectively. Figure ?? contains 12 plots in an array with four rows and three columns, and each row corresponds to one combination of (q_1, q_2) with $q_{\max} = 50, 60, 80$, and 90 , respectively.

Case d. We set $n = 101$, $k = 2$, $r = p - q_{\max} = 1$ and $v = n - p = 1, 3, 5$ and 10 .

For the first three cases above, we select parameters to maintain a reasonably large value for $n - q_{\max}$ so that we compare the performance of classical chi-square approximation (??), the normal approximation (??), and the adjusted chi-square approximation in (??). Under Case **a**, the values of q_i ’s are proportional. Under Cases **b** and **c**, two extreme situations, either $q_{\max} = p - 1$ or $p = n - 1$, are considered.

Case **d** is used to compare the performance of the normal approximation (??), the adjusted chi-square approximation in (??) and the non-normal approximation in (??). Since we take $r = 1$, the limit on the right-hand side of (??) has only one term, that is, the limit is $-\log Y_v$, where Y_v is a random variable having

a chi-square distribution with v degrees of freedom.

The results under Cases **a** to **d** are given in Figures ?? to ??, respectively.

Now we summarize our findings from Figures ?? to ??.

1. The classical chi-square approximation (Chisq) works very well for small p , but it becomes worse with the increase of p and finally departs from the histograms of the test statistic.
2. When p is small such as $p = 10$, the normal approximation (CLT) shows lack of fit to the histograms and it is getting better with the increase of p .
3. The adjusted likelihood ratio method (ALRT) works very well for all cases, that is, for small p , the ALRT behaves like the classical chi-square approximation, while for large p , it performs very well too like the normal approximation.
4. In Figure ??, we select $p = n - 1$. Both the normal approximation and the adjusted likelihood method show a little bit departure from the histograms since p is too close to n and sample size $n = 101$ is not a large sample size. In this case, both the normal approximation and the adjusted likelihood method can improve when n is getting larger.
5. From Figure ??, when v is small, the non-normal approximation works much better than the normal approximation and the adjusted chi-square approximation. When the value of v increases from 5 to 10, both the normal approximation and the adjusted chi-square approximation improve significantly. This implies that one can use the normal approximation or the adjusted chi-square approximation when $r + v = n - p_{\max}$ is not too small. We note that the exact distribution of the limit on the right-hand side of (??) is not easy to obtain if $r = p - q_{\max} > 1$.

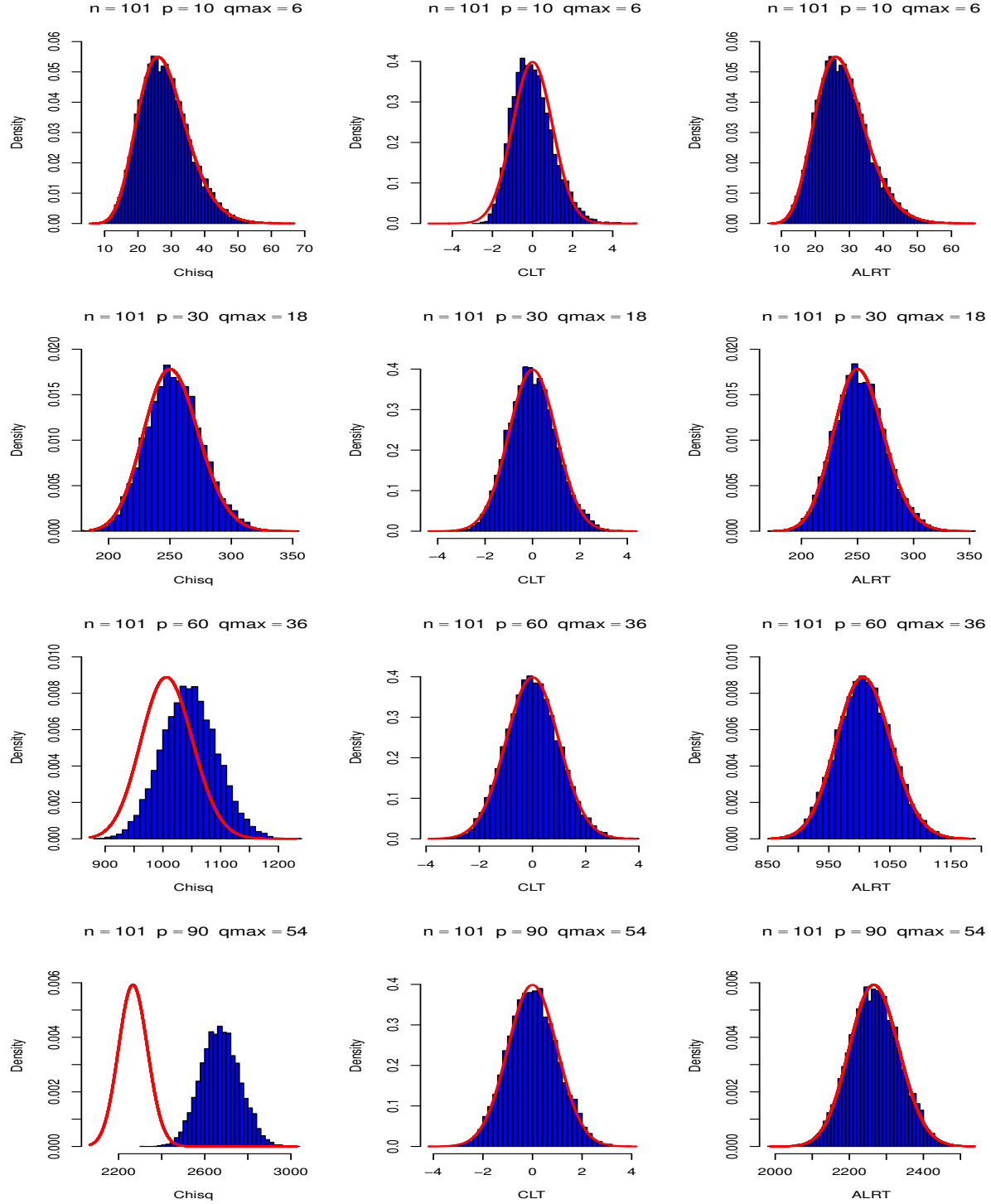


Figure 1: Plots of histograms and theoretical density curves (smooth curves in graphs) for three test statistics based on the classical chi-square approximation (??), the normal approximation (??), and the adjusted chi-square approximation in (??) with selection of $k = 3$ and ratio $q_1 : q_2 : q_3 = 3 : 1 : 1$.

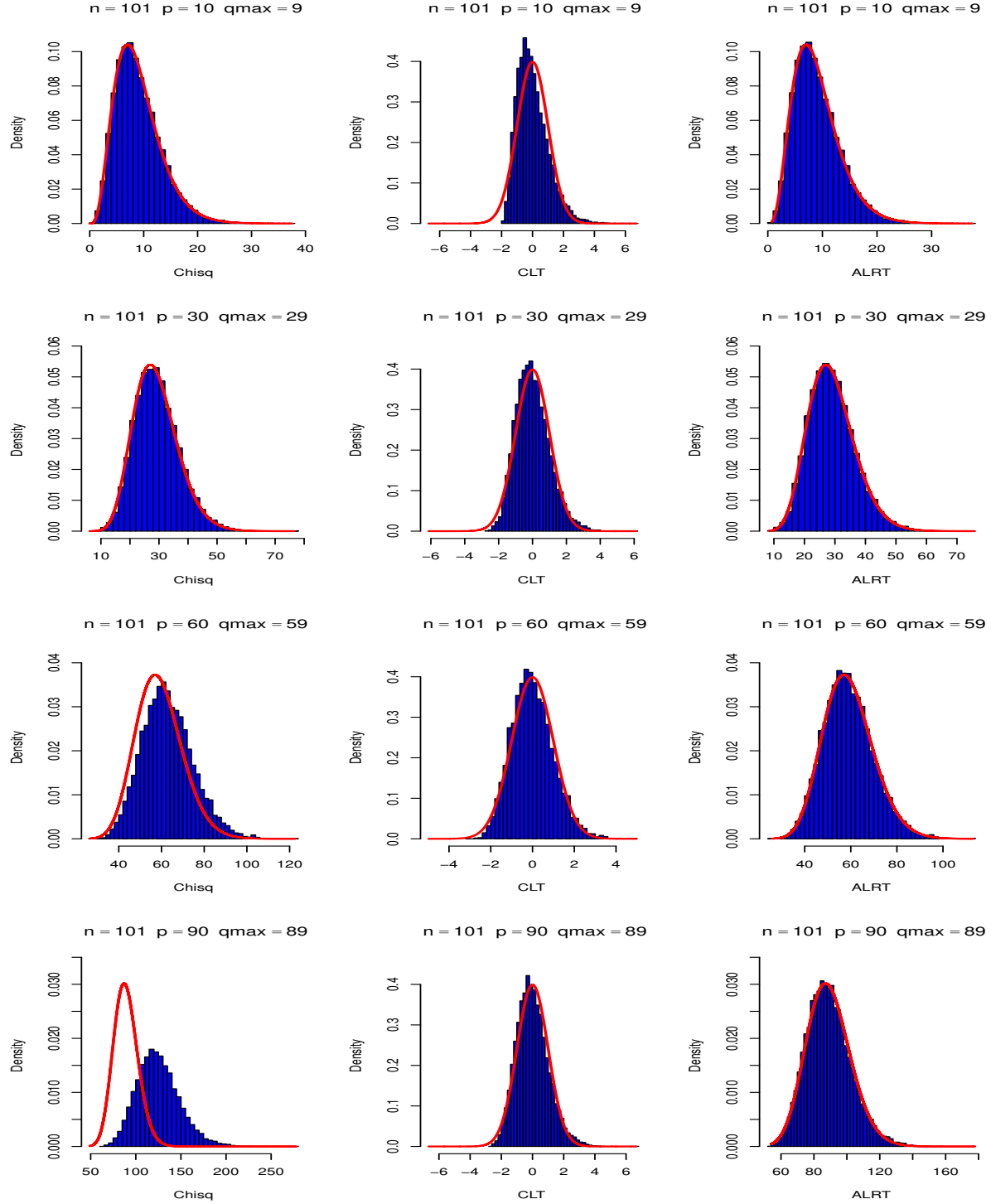


Figure 2: Plots of histograms and theoretical density curves (smooth curves in graphs) for three test statistics based on the classical chi-square approximation (??), the normal approximation (??), and the adjusted chi-square approximation in (??) with selection of $k = 2$, $(q_1, q_2) = (p - 1, 1)$ and $p = 10, 30, 60$ and 90 .

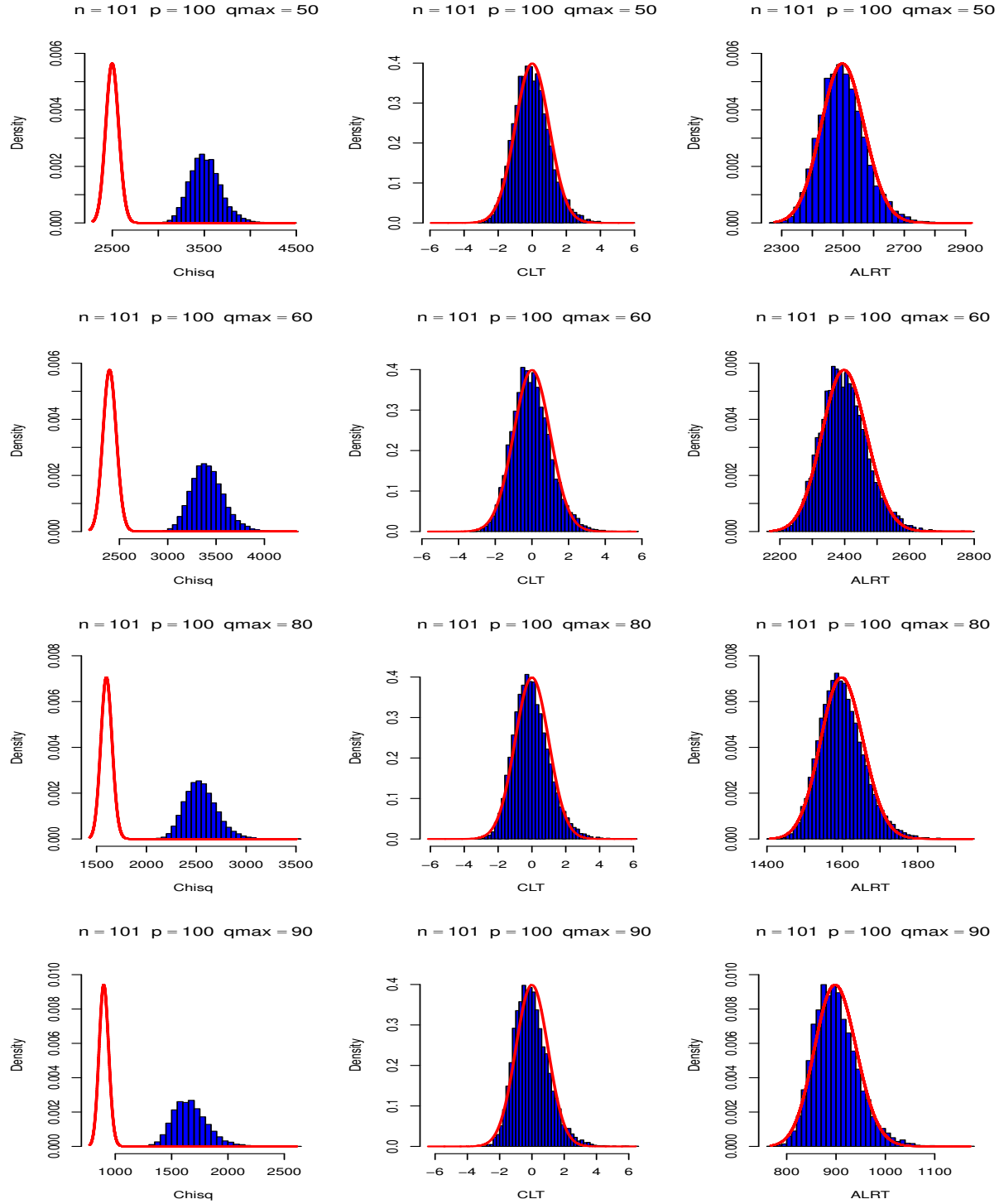


Figure 3: Plots of histograms and theoretical density curves (smooth curves in graphs) for three test statistics based on the classical chi-square approximation (??), the normal approximation (??), and the adjusted chi-square approximation in (??) with selection of $k = 2$, $p = 100$ and $(q_1, q_2) = (50, 50)$, $(60, 40)$, $(80, 20)$, and $(90, 10)$.

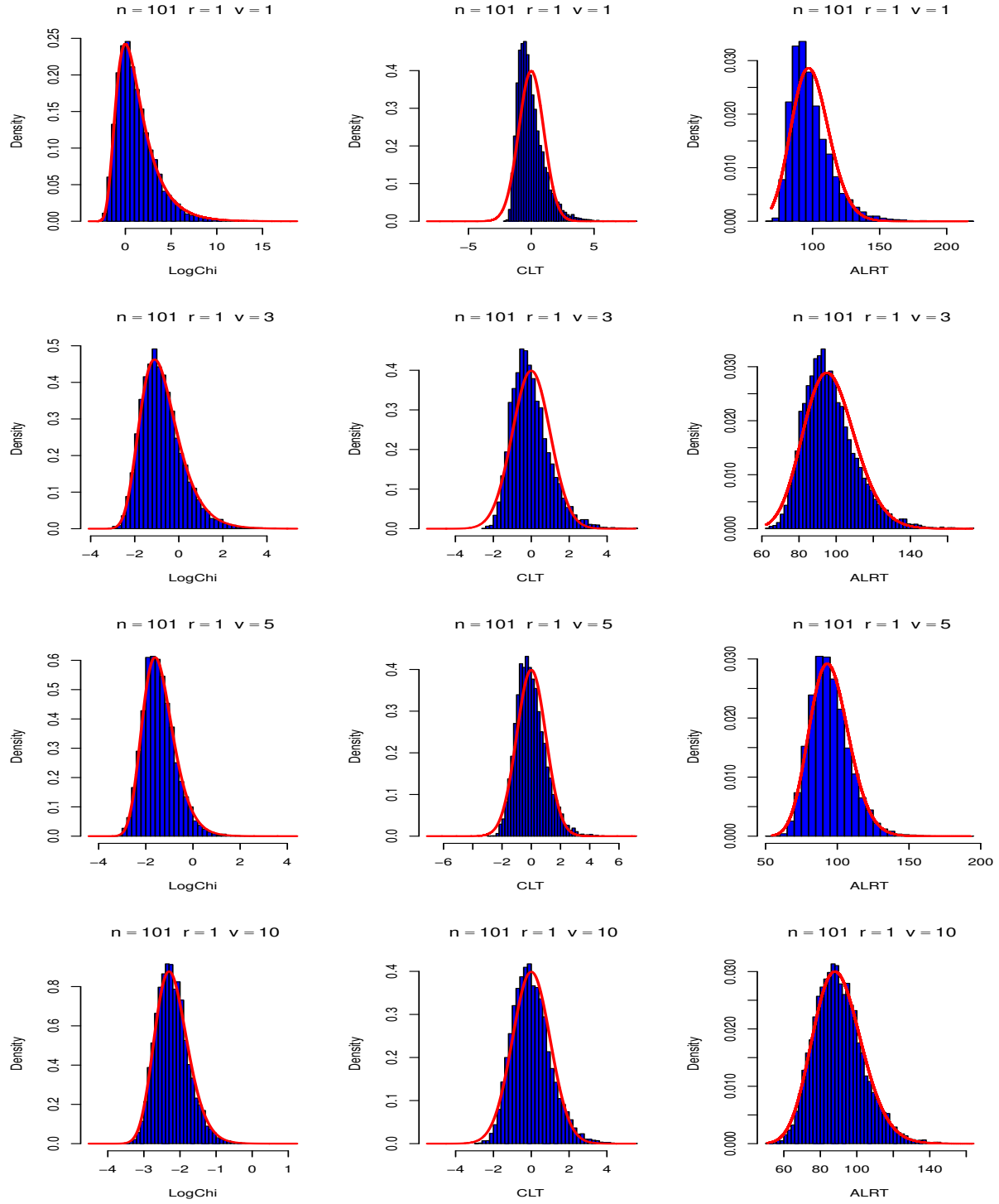


Figure 4: Plots of histograms and theoretical density curves (smooth curves in graphs) for three test statistics based on non-normal approximation (??), the normal approximation (??), and the adjusted chi-square approximation in (??) with selection of $n = 101$, $k = 2$, $r = p - q_{\max} = 1$, and $v = n - p = 1, 3, 5$ and 10. For the plots in the first column above, LogChi denotes the non-normal approximation in (??).

3.2 Comparison of adjusted log-LRT and other methods

In this subsection, we compare our adjusted log-likelihood ratio test statistic, i.e., Z_n in (??) with other three test statistics, including two trace criterion test statistics by Jiang et al. [?] and Li et al. [?] and Schott type statistics by Bao et al. [?]. We notice that Jiang et al. [?] and Bao et al. [?] investigate their test statistics for test (??) for any fixed $k \geq 2$ while Li et al. [?] consider test (??) for $k = 2$ only. Given that Li et al. [?]'s test statistics are suitable for $k = 2$ only, we set $k = 2$ for the comparison in this section.

In Jiang et al. [?], a large-dimensional trace criterion test statistic is defined as

$$L_n = \mathbf{tr}(A_{21}A_{11}^{-1}A_{12}A_{22}^{-1}),$$

where $\mathbf{tr}(A)$ denotes the trace of matrix A . Under conditions $r_{n1} := q_2/q_1 \rightarrow r_1 \in (0, \infty)$, $r_{n2} := q_2/(n - 1 - q_2) \rightarrow r_2 \in (0, \infty)$, and $q_2 < n$, it is shown that

$$T_2 := \frac{L_n - a_n}{\sqrt{b_n}} \xrightarrow{d} N(0, 1) \text{ as } n \rightarrow \infty \quad (3.1)$$

under the null hypothesis in (??), where

$$b_n = \frac{2h_n^2 r_{n1}^2 r_{n2}^2}{(r_{n1} + r_{n2})^2}, \quad a_n = \frac{q_2 r_{n2}}{r_{n1} + r_{n2}}, \quad h_n = \sqrt{r_{n1} + r_{n2} - r_{n1} r_{n2}}.$$

Some calculation shows that $a_n = q_1 q_2 / (n - 1)$ and $b_n = 2q_1 q_2 (n - 1 - q_1)(n - 1 - q_2) / (n - 1)^4$.

When $k = 2$, the test statistic in Bao et al. [?] is equal to

$$\mathbf{tr}(A_{22}^{-1/2} A_{21} A_{11}^{-1} A_{12} A_{22}^{-1/2}) = \mathbf{tr}(A_{21} A_{11}^{-1} A_{12} A_{22}^{-1}),$$

which is the same as L_n . Meanwhile, Theorem 3.1 in [?] also implies (??).

To introduce Li et al. [?]'s test statistic, set for $i, j = 1, 2$

$$\gamma_{ij} = \frac{1}{(n-2)(n+1)} \left(\mathbf{tr}(A_{ij} A_{ji}) - \frac{1}{n-1} \mathbf{tr}(A_{ii}) \mathbf{tr}(A_{jj}) \right).$$

The trace criterion test statistic by [?] is defined as γ_{12} . Under the null hypothesis in (??) for $k = 2$, it is shown in Li et al. [?] that

$$T_3 := \sqrt{\frac{(n-2)(n+1)}{2}} \frac{\gamma_{12}}{\sqrt{\gamma_{11}\gamma_{22}}} \xrightarrow{d} N(0, 1) \text{ as } n \rightarrow \infty \quad (3.2)$$

given that $p = q_1 + q_2 \rightarrow \infty$ as $n \rightarrow \infty$ and

$$0 < \lim_{n \rightarrow \infty} \frac{1}{p} \mathbf{tr}(\Sigma^i) < \infty \text{ for } i = 1, 2, 4.$$

We notice that the test statistic T_3 works only for the case $k = 2$, but condition $p = q_1 + q_2 \rightarrow \infty$ is less restrictive than those required for other statistics we just discussed.

For any fixed size $\alpha \in (0, 1)$, test T_2 (or T_3) rejects the null hypothesis in (??) if $T_2 > z_\alpha$ (or $T_3 > z_\alpha$), where z_α denotes the α -level critical value of the standard normal distribution.

To compare our adjusted log-likelihood ratio test statistic Z_n and test statistics T_2 and T_3 , we assume $q_1 > q_2 \geq 1$ and $p = q_1 + q_2 < n$. Our samples are generated from the populations similar to those in Jiang et al. [?] and Qi et al. [?]. Let $\mathbf{z} = (z_1, \dots, z_p)'$ be a random vector whose components are independent normal random variables with mean 0 and variance 1.

Model 1. $\mathbf{x} = (x_1, \dots, x_p)'$, where $x_i = (1 + c)z_i$ for $i = 1, \dots, p_1$, $x_{p_1+j} = z_{p_1+j} + cz_j$ for $j = 1, \dots, p_2$, and c is a constant;

Model 2. $\mathbf{x} = (x_1, \dots, x_p)'$, where $x_i = (1 + c)z_i$ for $i = 1, \dots, p_1$, $x_{p_1+j} = z_{p_1+j} + cz_j$ for $j = 1, \dots, p_2 - 1$, $x_p = p^{-1/4}z_p$, and c is a constant.

With different selections of (q_1, q_2, n, c) , we draw 10000 random samples of size n from each model above (Model 1 and Model 2) and then we calculate the empirical sizes of the tests (when $c = 0$) or the powers of the tests (when $c \neq 0$). The nominal level of size α (type I error) is set to be 0.05 in the simulation.

Tables ?? and ?? present results for the numerical comparisons on the three test statistics. From the two tables, the adjusted log-likelihood ratio test statistic Z_n is constantly accurate in terms of type I error, and T_3 has larger type I errors than the nominal level 0.05 for small values of p . In terms of empirical powers, Z_n and T_2 are comparable in most cases while T_3 is better than both Z_n and T_2 under Model 1. Under Model 2, Z_n has a slightly larger power than T_2 in most cases and both are significantly better than T_3 .

4 Some lemmas

The multivariate gamma function, denoted by $\Gamma_p(x)$, is defined as

$$\Gamma_p(x) := \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma(x - \frac{1}{2}(i-1)) \quad (4.1)$$

with $x > \frac{p-1}{2}$ from Muirhead [?].

For any positive integers n and p with $1 < p < n$, let $q_1, \dots, q_k$ be k positive integers such that $p = \sum_{i=1}^k q_i$, where $k \geq 2$ is an integer which may depend on n . Denote $q_{\max} = \max_{1 \leq i \leq k} q_i$.

For any function g defined over $(0, \infty)$, set

$$\Psi_{g,n,p}(x) = \Delta_{g,n,p}(x) - \sum_{i=1}^k \Delta_{g,n,q_i}(x) = \sum_{j=1}^p g(\frac{n-j}{2} + x) - \sum_{i=1}^k \sum_{j=1}^{q_i} g(\frac{n-j}{2} + x) \quad (4.2)$$

for $x > -\frac{n-p}{2}$, where Δ_{g,n,q_i} is defined in (??). For brevity, we omit $q_1, \dots, q_k$ in the definition of $\Psi_{g,n,p}(x)$.

Let g be a differentiable function. Both $\Delta_{g,n,q}$ and $\Psi_{g,n,p}$ are linear functionals in g with following property

$$\frac{d}{dx} \Delta_{g,n,q}(x) = \Delta_{g',n,q}(x), \quad \frac{d}{dx} \Psi_{g,n,p}(x) = \Psi_{g',n,p}(x). \quad (4.3)$$

Table 1: Comparisons on Size and Power under Model 1. The sizes and powers are estimated based on 10,000 simulations under Model 1, and the nominal type I errors for all tests are set to be 0.05.

(q_1, q_2)	n	Size ($c = 0$)			Power ($c = 0.1$)			Power ($c = 0.2$)		
		Z_n	T_2	T_3	Z_n	T_2	T_3	Z_n	T_2	T_3
(6, 4)	20	0.0507	0.0514	0.0645	0.0605	0.0583	0.0806	0.0909	0.0923	0.1394
	50	0.0507	0.0593	0.0641	0.0857	0.0999	0.1128	0.2400	0.2654	0.3044
	100	0.0527	0.0615	0.0645	0.1350	0.1570	0.1690	0.5643	0.6028	0.6262
(8, 2)	20	0.0498	0.0501	0.0609	0.0532	0.0547	0.0705	0.0717	0.0704	0.1029
	50	0.0487	0.0581	0.0641	0.0699	0.0838	0.0927	0.1499	0.1719	0.1983
	100	0.0530	0.0665	0.0646	0.0957	0.1154	0.1251	0.3086	0.3461	0.3709
(18, 12)	50	0.0499	0.0542	0.0545	0.1897	0.1983	0.2255	0.8617	0.8732	0.9147
	100	0.0506	0.0528	0.0554	0.1233	0.1289	0.1482	0.5492	0.5702	0.6757
	150	0.0499	0.0542	0.0545	0.1897	0.1983	0.2255	0.8617	0.8732	0.9147
(24, 6)	50	0.0499	0.0472	0.0562	0.0647	0.0614	0.0777	0.1133	0.1115	0.1811
	100	0.0478	0.0519	0.0540	0.0854	0.0881	0.1096	0.2990	0.3128	0.3855
	150	0.0547	0.0595	0.0588	0.1214	0.1334	0.1482	0.5393	0.5587	0.6210
(36, 24)	100	0.0511	0.0543	0.0519	0.1053	0.1112	0.1495	0.4243	0.4599	0.6915
	150	0.0493	0.0494	0.0527	0.1619	0.1629	0.2224	0.8144	0.8361	0.9328
	200	0.0538	0.0554	0.0562	0.2428	0.2498	0.3138	0.9715	0.9756	0.9915
(48, 12)	100	0.0519	0.0523	0.0509	0.0771	0.0795	0.1039	0.2235	0.2325	0.3836
	150	0.0503	0.0496	0.0568	0.1115	0.1137	0.1368	0.4657	0.4793	0.6399
	200	0.0473	0.0503	0.0513	0.1471	0.1493	0.1817	0.7222	0.7334	0.8261
(60, 40)	150	0.0499	0.0495	0.0502	0.1248	0.1362	0.2202	0.6453	0.7046	0.9407
	200	0.0524	0.0552	0.0555	0.2036	0.2140	0.3101	0.9358	0.9497	0.9940
	300	0.0532	0.0547	0.0526	0.4046	0.4153	0.5247	1.0000	1.0000	1.0000
(80, 20)	150	0.0501	0.0514	0.0553	0.0907	0.0959	0.1343	0.3269	0.3543	0.6420
	200	0.0521	0.0529	0.0527	0.1287	0.1325	0.1795	0.6093	0.6287	0.8362
	300	0.0507	0.0519	0.0502	0.2162	0.2203	0.2791	0.9453	0.9500	0.9844

Now we set $\beta_{nr}(x) = \Psi_{g,n,p}(x)$ when $g(x) = 1/x^r$ for $r \geq 1$, that is, we define

$$\beta_{nr}(x) = \sum_{j=1}^p \frac{1}{\left(\frac{n-j}{2} + x\right)^r} - \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{\left(\frac{n-j}{2} + x\right)^r}, \quad x > -\frac{n-p}{2}. \quad (4.4)$$

Note that $\beta_{nr}(x) \geq 0$ since $g(x) = 1/x^r$, $x > 0$ is a decreasing function over $(0, \infty)$.

Table 2: Comparisons on Size and Power under Model 2. The sizes and powers are estimated based on 10,000 simulations under Model 2, and the nominal type I errors for all tests are set to be 0.05.

(q_1, q_2)	n	Size ($c = 0$)			Power ($c = 0.1$)			Power ($c = 0.2$)		
		Z_n	T_2	T_3	Z_n	T_2	T_3	Z_n	T_2	T_3
(6, 4)	20	0.0507	0.0514	0.0658	0.0797	0.0798	0.0834	0.1965	0.1991	0.1525
	50	0.0507	0.0593	0.0658	0.1916	0.2083	0.1175	0.7428	0.7472	0.3532
	100	0.0527	0.0615	0.0647	0.4467	0.4774	0.1880	0.9938	0.9945	0.7142
(8, 2)	20	0.0498	0.0501	0.0626	0.0787	0.0763	0.0749	0.1855	0.1805	0.1214
	50	0.0487	0.0581	0.0695	0.1966	0.2178	0.1057	0.7311	0.7443	0.2567
	100	0.0530	0.0665	0.0654	0.4261	0.4626	0.1456	0.9901	0.9921	0.5237
(18, 12)	50	0.0494	0.0492	0.0579	0.1365	0.1330	0.1000	0.5468	0.4722	0.3095
	100	0.0506	0.0528	0.0564	0.4118	0.3876	0.1582	0.9961	0.9841	0.7051
	150	0.0499	0.0542	0.0571	0.7148	0.6869	0.2317	1.0000	1.0000	0.9343
(24, 6)	50	0.0499	0.0472	0.0570	0.1378	0.1279	0.0852	0.5213	0.4103	0.1983
	100	0.0478	0.0519	0.0536	0.4234	0.3949	0.1132	0.9937	0.9718	0.4257
	150	0.0547	0.0595	0.0601	0.7320	0.6968	0.1587	1.0000	1.0000	0.6789
(36, 24)	100	0.0511	0.0543	0.0517	0.2865	0.2672	0.1492	0.9389	0.8449	0.7075
	150	0.0493	0.0494	0.0535	0.6008	0.5303	0.2269	0.9999	0.9979	0.9392
	200	0.0538	0.0554	0.0576	0.8552	0.7843	0.3194	1.0000	1.0000	0.9926
(48, 12)	100	0.0519	0.0523	0.0511	0.2822	0.2407	0.1048	0.9086	0.7314	0.4058
	150	0.0503	0.0496	0.0553	0.6079	0.5098	0.1447	0.9998	0.9906	0.6671
	200	0.0473	0.0503	0.0519	0.8579	0.7741	0.1913	1.0000	0.9999	0.8484
(60, 40)	150	0.0499	0.0495	0.0493	0.3966	0.3446	0.2209	0.9915	0.9530	0.9460
	200	0.0524	0.0552	0.0560	0.7092	0.5988	0.3147	1.0000	0.9993	0.9962
	300	0.0532	0.0547	0.0535	0.9768	0.9346	0.5361	1.0000	1.0000	1.0000
(80, 20)	150	0.0501	0.0514	0.0564	0.3863	0.3046	0.1403	0.9766	0.8280	0.6582
	200	0.0521	0.0529	0.0528	0.7005	0.5473	0.1865	1.0000	0.9924	0.8517
	300	0.0507	0.0519	0.0510	0.9791	0.9093	0.2885	1.0000	1.0000	0.9886

Define

$$s(x) = (x - \frac{1}{2}) \log(x) - x, \quad x > 0$$

and set

$$h(x) = \log \Gamma(x) - s(x), \quad x > 0. \tag{4.5}$$

Then we can verify that

$$s'(x) = \log x - \frac{1}{2x}, \quad s''(x) = \frac{1}{x} + \frac{1}{2x^2}, \quad s'''(x) = -\frac{1}{x^2} - \frac{1}{x^3}, \quad (4.6)$$

and for some constant $C > 0$

$$|h''(x)| \leq \frac{C}{x^3}, \quad x \geq \frac{1}{4}; \quad (4.7)$$

See Lemma 4.4 in Guo and Qi [?].

For the digamma function ψ defined in (??), it follows from Formula 6.3.18 in Abramowitz and Stegun [?] that

$$\psi(x) = \log x - \frac{1}{2x} - \frac{1}{12x^2} + O\left(\frac{1}{x^4}\right) \quad (4.8)$$

as $x \rightarrow \infty$.

From now on, we adopt the following notation in our lemmas and our proofs. For any two sequences, $\{a_n\}$ and $\{b_n\}$ with $b_n > 0$, notation $a_n = o(b_n)$ implies $\lim_{n \rightarrow \infty} a_n/b_n = 0$, and notation $a_n = O(b_n)$ means a_n/b_n is uniformly bounded.

We first introduce the formula for the h -th moment of W_n .

Lemma 4.1. (Theorem 11.2.3 in Muirhead [?]) Let $p = \sum_{i=1}^k q_k$ and W_n be Wilk's likelihood ratio statistics defined in (??). Then, under the null hypothesis in (??), we have

$$\mathbf{E}(W_n^h) = \frac{\Gamma_p(\frac{n-1}{2} + h)}{\Gamma_p(\frac{n-1}{2})} \prod_{i=1}^k \frac{\Gamma_{q_i}(\frac{n-1}{2})}{\Gamma_{q_i}(\frac{n-1}{2} + h)} \quad (4.9)$$

for any $h > \frac{p-n}{2}$, where $\Gamma_p(x)$ is defined in (??).

Next, we introduce a distributional representation for W_n .

Lemma 4.2. (Theorem 11.2.4 in Muirhead [?]) Let $p = \sum_{i=1}^k q_k$ and W_n be Wilk's likelihood ratio statistics defined in (??). Then, under the null hypothesis in (??), W_n has the same distribution as

$$\prod_{i=2}^k \prod_{j=1}^{q_i} V_{ij},$$

where $q_i^* = \sum_{j=1}^{i-1} q_j$ for $2 \leq i \leq k$, the V_{ij} 's are independent random variables and V_{ij} has a $\text{beta}(\frac{1}{2}(n - q_i^* - j), \frac{1}{2}q_i^*)$ distribution.

Lemma 4.3. (Lemma 5 in Qi et al. [?]) As $n \rightarrow \infty$,

$$\sum_{i=1}^q \left(\frac{1}{n-i} - \frac{1}{n-1} \right) = -\frac{q}{n} - \log\left(1 - \frac{q}{n}\right) + O\left(\frac{q}{n(n-q)}\right), \quad (4.10)$$

$$\sum_{i=1}^q \left(\log(n-i) - \log(n-1) \right) = \left(q - n + \frac{1}{2} \right) \log\left(1 - \frac{q}{n}\right) - \frac{(n-1)q}{n} + O\left(\frac{q}{n(n-q)}\right) \quad (4.11)$$

uniformly over $1 \leq q < n$.

Lemma 4.4. *With β_{nr} defined in (??), we have*

$$\frac{2p}{3n} \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) \leq \beta_{n1}(0) \leq \frac{4p}{n} \log \left(1 + \frac{2(p - q_{\max})}{n - p} \right), \quad (4.12)$$

$$\beta_{n2}(x) \leq \frac{32\beta_{n1}(0)}{n - p}, \quad (4.13)$$

$$\beta_{n2}(x) \leq \frac{8192p(p - q_{\max})}{n(n - p)(n - q_{\max})} \quad (4.14)$$

and

$$\beta_{n3}(x) \leq \frac{192\beta_{n1}(0)}{(n - p)^2} \quad (4.15)$$

for all $|x| \leq \frac{n-p}{4}$.

Proof. Without loss of generality, assume $q_1 \geq q_2 \geq \dots \geq q_k$. Therefore, $q_1 = q_{\max} = \max_{1 \leq i \leq k} q_i$. Set $q_i^* = \sum_{j=1}^{i-1} q_j$ for $2 \leq i \leq k+1$. Then we see that

$$\begin{aligned} \beta_{n1}(0) &= 2 \sum_{j=1}^p \frac{1}{n-j} - \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{n-j} \\ &= 2 \sum_{i=2}^k \sum_{j=1}^{q_i} \left(\frac{1}{n - q_i^* - j} - \frac{1}{n-j} \right) \\ &= 2 \sum_{i=2}^k \sum_{j=1}^{q_i} \frac{q_i^*}{(n - q_i^* - j)(n-j)}. \end{aligned} \quad (4.16)$$

Note that $q_i^* \geq q_1 = q_{\max}$, and $q_i^* + q_i \leq p$ for all $2 \leq i \leq k$. Moreover, $q_i \leq p/2 \leq n/2$ for all $2 \leq i \leq k$, and thus $n > n - j \geq n/2$ for $1 \leq j \leq q_i$, $2 \leq i \leq k$. We have

$$\begin{aligned} \beta_{n1}(0) &\leq \frac{4p}{n} \sum_{i=2}^k \sum_{j=1}^{q_i} \frac{1}{n - q_i^* - j} \\ &= \frac{4p}{n} \sum_{j=1}^{p-q_{\max}} \frac{1}{n - q_{\max} - j} \\ &\leq \frac{4p}{n} \int_{0.5}^{p-q_{\max}+0.5} \frac{1}{n - q_{\max} - x} dx \\ &= \frac{4p}{n} \log \left(\frac{n - q_{\max} - 0.5}{n - p - 0.5} \right) \\ &= \frac{4p}{n} \log \left(1 + \frac{2(p - q_{\max})}{n - p} \right), \end{aligned}$$

which gives the upper bound in (??).

To verify the lower bound in (??), define $m = \min\{i \geq 2 : q_i^* \geq p/3\}$. We see that $m = 2$ if $q_1 = q_{\max} \geq p/3$. If $q_1 < p/3$, then $q_k \leq \dots \leq q_2 \leq q_1 < p/3$, which implies that $k \geq 4$, $2 < m \leq k$, $q_{m-1}^* < p/3$, $p/3 \leq q_m^* < 2p/3$, and thus, $p - q_m^* \geq p/3 \geq (p - q_{\max})/3$. When $q_1 < p/3$, it is trivial that

$p - q_m^* \geq p/3 \geq (p - q_{\max})/3$. Then we conclude from (??) that

$$\begin{aligned}
\beta_{n1}(0) &\geq 2 \sum_{i=m}^k \sum_{j=1}^{q_i} \frac{q_i^*}{(n - q_i^* - j)(n - j)} \\
&\geq \frac{2p}{3n} \sum_{i=m}^k \sum_{j=1}^{q_i} \frac{1}{n - q_i^* - j} \\
&= \frac{2p}{3n} \sum_{j=1}^{p-q_m^*} \frac{1}{n - q_m^* - j} \\
&\geq \frac{2p}{3n} \int_0^{p-q_m^*} \frac{1}{n - q_m^* - x} dx \\
&= \frac{2p}{3n} \log \left(\frac{n - q_m^*}{n - p} \right) \\
&= \frac{2p}{3n} \log \left(1 + \frac{p - q_m^*}{n - p} \right) \\
&= \frac{2p}{3n} \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right),
\end{aligned}$$

proving (??).

We can also verify that

$$\beta_{n2}(x) = 4 \sum_{i=2}^k \sum_{j=1}^{q_i} \left(\frac{1}{(n - q_i^* - j + 2x)^2} - \frac{1}{(n - j + 2x)^2} \right) \quad (4.17)$$

and

$$\beta_{n3}(x) = 8 \sum_{i=2}^k \sum_{j=1}^{q_i} \left(\frac{1}{(n - q_i^* - j + 2x)^3} - \frac{1}{(n - j + 2x)^3} \right). \quad (4.18)$$

For any x with $|x| \leq (n - p)/4$, we have for any $1 \leq j \leq q_i$, $2 \leq i \leq k$

$$\begin{aligned}
\frac{1}{(n - q_i^* - j + 2x)^2} - \frac{1}{(n - j + 2x)^2} &= \frac{q_i^*(2n - q_i^* - 2j + 4x)}{(n - q_i^* - j + 2x)^2(n - j + 2x)^2} \\
&\leq \frac{2q_i^*}{(n - q_i^* - j + 2x)^2(n - j + 2x)} \\
&\leq \frac{16q_i^*}{(n - q_i^* - j)^2(n - j)} \\
&\leq \frac{16}{n - p} \frac{q_i^*}{(n - q_i^* - j)(n - j)},
\end{aligned} \quad (4.19)$$

which, together with (??), implies

$$\beta_{n2}(x) \leq \frac{64}{n - p} \sum_{i=2}^k \sum_{j=1}^{q_i} \frac{q_i^*}{(n - q_i^* - j)(n - j)} = \frac{32\beta_{n1}(0)}{n - p},$$

proving (??).

Likewise, for any x with $|x| \leq (n-p)/4$, we have for any $1 \leq j \leq q_i$, $2 \leq i \leq k$

$$\begin{aligned} \frac{1}{(n-q_i^*-j+2x)^3} - \frac{1}{(n-j+2x)^3} &\leq \frac{3q_i^*}{(n-q_i^*-j+2x)^3(n-j+2x)} \\ &\leq \frac{48q_i^*}{(n-q_i^*-j)^3(n-j)} \\ &\leq \frac{48}{(n-p)^2} \frac{q_i^*}{(n-q_i^*-j)(n-j)}, \end{aligned}$$

which coupled with (??) yields (??).

Now we will show (??). We see that $\max_{2 \leq i \leq k} q_i \leq n/2$ since $\max_{2 \leq i \leq k} q_i + q_1 \leq p < n$. This implies $n-j \geq n/2$ for $1 \leq j \leq q_i$, $2 \leq i \leq k$. Then it follows from (??) and (??) that

$$\beta_{n2}(x) \leq 64 \sum_{i=2}^k \sum_{j=1}^{q_i} \frac{16q_i^*}{(n-q_i^*-j)^2(n-j)} \leq \frac{2048p}{n} \sum_{i=2}^k \sum_{j=1}^{q_i} \frac{1}{(n-q_i^*-j)^2} = \frac{2048p}{n} \sum_{j=q_1+1}^p \frac{1}{(n-j)^2}.$$

Since the sum in the last expression is dominated by the integral

$$\int_{q_1+0.5}^{p+0.5} \frac{1}{(n-y)^2} dy \leq \frac{1}{n-p-0.5} - \frac{1}{n-q_1-0.5} = \frac{p-q_1}{(n-p-0.5)(n-q_1-0.5)} \leq \frac{4(p-q_{\max})}{(n-p)(n-q_{\max})},$$

we obtain (??). $\square$

Lemma 4.5. *There exists a universal constant D such that*

$$|\Psi_{h'',n,p}(x)| \leq \frac{Dp}{n(n-p)^2} \quad (4.20)$$

and

$$|\Psi_{s''',n,p}(x)| \leq \frac{32\beta_{n1}(0)}{n-p} \quad (4.21)$$

uniformly over $|x| \leq \frac{n-p}{4}$ for all large n .

Proof. For integers q and n with $1 \leq q < n$,

$$\sum_{j=1}^q \frac{1}{(n-j)^3} \leq \int_{0.5}^{q+0.5} \frac{1}{(n-x)^3} dx = \frac{1}{2} \left(\frac{1}{(n-q-0.5)^2} - \frac{1}{(n-0.5)^2} \right) \leq \frac{q}{(n-q-0.5)^2(n-0.5)} \leq \frac{8q}{n(n-q)^2}.$$

We apply the above inequality to $q = q_1, \dots, q_k$ and p . Then we obtain that

$$\sum_{j=1}^p \frac{1}{(n-j)^3} \leq \frac{8p}{n(n-q)^2} \quad (4.22)$$

and

$$\sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{(n-j)^3} \leq \sum_{i=1}^k \frac{8q_i}{n(n-q_i)^2} \leq \sum_{i=1}^k \frac{8q_i}{n(n-p)^2} = \frac{8p}{n(n-p)^2}, \quad (4.23)$$

where we have used the fact that $p = \sum_{i=1}^k q_i$.

In view of (??), we have

$$\begin{aligned}
|\Psi_{h'',n,p}(x)| &\leq C \sum_{j=1}^p \frac{1}{(\frac{n-j}{2} + x)^3} + C \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{(\frac{n-j}{2} + x)^3} \\
&= 8C \sum_{j=1}^p \frac{1}{(n-j+2x)^3} + 8C \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{(n-j+2x)^3} \\
&\leq 64C \sum_{j=1}^p \frac{1}{(n-j)^3} + 64C \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{(n-j)^3}
\end{aligned}$$

uniformly over $|x| \leq \frac{n-p}{4}$. In the last step above, we have used the inequality $n-j+2x \geq (n-j)/2$ for all $1 \leq j \leq p$ and $|x| \leq (n-p)/4$. By combining (??) and (??), we obtain (??) with $D = 1024C$.

(??) can be verified by using Lemma ?? and (??) and the fact that $|\Psi_{s''',n,p}(x)| \leq \beta_{n2}(x) + \beta_{n3}(x)$. This completes the proof of the lemma. $\square$

Lemma 4.6. Assume $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$. Then as $n \rightarrow \infty$

$$\frac{(n - q_{\max})(n - p)}{p - q_{\max}} \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) > \frac{1}{3} \log(1 + n - q_{\max}) \rightarrow \infty \quad (4.24)$$

and

$$(n - p)^2 \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) \rightarrow \infty. \quad (4.25)$$

In addition, if $p \rightarrow \infty$, then

$$p(n - p) \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) \rightarrow \infty. \quad (4.26)$$

Proof. Define $f(x) = \frac{\log(1+x)}{x}$, $x > 0$. We can verify that

$$f'(x) = -\frac{1}{x^2} \left(\log(1+x) + \frac{1}{1+x} - 1 \right) = -\frac{1}{x^2} \int_0^x \frac{t}{1+t^2} dt < 0,$$

that is, $f(x)$ is decreasing in $x > 0$. Now set $x_n = (p - q_{\max})/3(n - p)$. Then $x_n < n - q_{\max}$. We get

$$\begin{aligned}
\frac{(n - q_{\max})(n - p)}{p - q_{\max}} \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) &= \frac{1}{3} (n - q_{\max}) f(x_n) \\
&> \frac{1}{3} (n - q_{\max}) f(n - q_{\max}) \\
&= \frac{1}{3} \log(1 + n - q_{\max}) \\
&\rightarrow \infty,
\end{aligned}$$

proving (??). (??) follows from (??) since

$$\frac{(n - q_{\max})(n - p)}{p - q_{\max}} \Big/ (n - p)^2 = \frac{n - q_{\max}}{(p - q_{\max})(n - p)} = \frac{1}{p - q_{\max}} + \frac{1}{n - p} \leq 2.$$

One can easily verify that for any $y > 0$, the function $x \log(1 + y/x)$ is increasing in $x \geq 1$. Therefore, we have

$$p(n - p) \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) = \frac{p}{3} \times 3(n - p) \log \left(1 + \frac{p - q_{\max}}{3(n - p)} \right) \geq \frac{p}{3} \log(1 + p - q_{\max}) \geq \frac{p}{3} \log(2) \rightarrow \infty$$

since $p \rightarrow \infty$ as $n \rightarrow \infty$. This proves (??). $\square$

Lemma 4.7. *If $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$, then*

$$\frac{b(n, p) - \sum_{i=1}^k b(n, q_i)}{\beta_{n1}(0)} \rightarrow 0, \quad \frac{b(n, p) - \sum_{i=1}^k b(n, q_i)}{\sqrt{\beta_{n1}(0)}} \rightarrow 0, \quad (4.27)$$

$$\frac{p}{n(n-p)\sqrt{\beta_{n1}(0)}} \rightarrow 0, \quad \frac{p}{n(n-p)^2\sqrt{\beta_{n1}(0)}} \rightarrow 0 \quad (4.28)$$

and

$$\lim_{n \rightarrow \infty} \frac{\mu_n - \bar{\mu}_n}{\sigma_n} = 0, \quad \lim_{n \rightarrow \infty} \frac{n^2 \beta_{n1}(0)}{\sigma_n^2} = 1 \quad (4.29)$$

as $n \rightarrow \infty$, where $b(n, q)$ is defined in (??). Moreover, (??) implies (??).

Proof. We have $b(n, p) - \sum_{i=1}^k b(n, q_i) = \frac{1}{4}\beta_{n2}(0)$, where β_{n2} is defined in (??). From (??), (??) and (??) we get

$$\frac{b(n, p) - \sum_{i=1}^k b(n, q_i)}{\beta_{n1}(0)} = O\left(\left(\frac{(n - q_{\max})(n - p)}{p - q_{\max}} \log\left(1 + \frac{p - q_{\max}}{3(n - p)}\right)\right)^{-1}\right) \rightarrow 0.$$

Similarly, from (??), (??) and (??) we have

$$\begin{aligned} \frac{b(n, p) - \sum_{i=1}^k b(n, q_i)}{\sqrt{\beta_{n1}(0)}} &= O\left(\frac{\sqrt{p}(p - q_{\max})}{\sqrt{n}(n - q_{\max})(n - p)\sqrt{\log\left(1 + \frac{p - q_{\max}}{3(n - p)}\right)}}\right) \\ &= O\left(\frac{1}{\sqrt{(n - p)^2 \log\left(1 + \frac{p - q_{\max}}{3(n - p)}\right)}}\right) \\ &\rightarrow 0. \end{aligned}$$

This completes the proof of (??).

By using (??) and (??) we get

$$\frac{p}{n(n-p)\sqrt{\beta_{n1}(0)}} = O\left(\frac{\sqrt{p/n}}{\sqrt{(n-k)^2 \log\left(1 + \frac{p - q_{\max}}{3(n-p)}\right)}}\right) = O\left(\frac{1}{\sqrt{(n-k)^2 \log\left(1 + \frac{p - q_{\max}}{3(n-p)}\right)}}\right) \rightarrow 0,$$

and

$$\frac{p}{n(n-p)^2\sqrt{\beta_{n1}(0)}} \leq \frac{p}{n(n-p)\sqrt{\beta_{n1}(0)}} \rightarrow 0,$$

proving (??).

The second limit in (??) follows from (??) since $\sigma_n^2 = n^2 \beta_{n1}(0) + 2n^2(b(n, p) - \sum_{i=1}^k b(n, q_i))$.

Set $\ell(x) = \log x - \frac{1}{2x} - \frac{1}{12x^2}$ for $x > 0$.

It follows from (??) that

$$g(x) := \psi(x) - \ell(x) = O(x^{-4}) \quad \text{as } x \rightarrow \infty.$$

Since $g(x)$ is a smooth function in $x > 0$, it is easy to show that

$$|g(x)| \leq \frac{C}{x^3}, \quad x \geq \frac{1}{4}$$

for some constant C . Following the same lines in the proof of (??) we have

$$|\Psi_{g,n,p}(0)| \leq \frac{Dp}{n(n-p)^2}$$

for some $D > 0$, which together with (??) implies

$$\frac{n\Psi_{\psi,n,p}(0) - n\Psi_{\ell,n,p}(0)}{\sqrt{n^2\beta_{n1}(0)}} = \frac{\Psi_{\psi,n,p}(0) - \Psi_{\ell,n,p}(0)}{\sqrt{\beta_{n1}(0)}} = \frac{\Psi_{g,n,p}(0)}{\sqrt{\beta_{n1}(0)}} \rightarrow 0. \quad (4.30)$$

In virtue of (??),

$$\begin{aligned} & \sum_{j=1}^p \log(n-j) - \sum_{i=1}^k \sum_{j=1}^{q_i} \log(n-j) \\ &= \sum_{j=1}^p (\log(n-j) - \log(n-1)) - \sum_{i=1}^k \sum_{j=1}^{q_i} (\log(n-j) - \log(n-1)) \\ &= (p-n+\frac{1}{2})\log(1-\frac{p}{n}) - \sum_{i=1}^k (q_i-n+\frac{1}{2})\log(1-\frac{q_i}{n}) + O(\frac{p}{n(n-p)}). \end{aligned}$$

We have used the condition $p = \sum_{i=1}^k q_i$ to simplify the expression. We skip the details here. Similarly, by using (??), we obtain

$$\sum_{j=1}^p \frac{1}{n-j} - \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{n-j} = -\log(1-\frac{p}{n}) + \sum_{i=1}^k \log(1-\frac{q_i}{n}) + O(\frac{p}{n(n-p)}).$$

Therefore, we have from (??) that

$$\begin{aligned} & \Psi_{\ell,n,p}(0) \\ &= \sum_{j=1}^p \log(n-j) - \sum_{i=1}^k \sum_{j=1}^{q_i} \log(n-j) - (\sum_{j=1}^p \frac{1}{n-j} - \sum_{i=1}^k \sum_{j=1}^{q_i} \frac{1}{n-j}) - \frac{1}{3}(b(n,p) - \sum_{i=1}^k b(n,q_i)) \\ &= (p-n+\frac{3}{2})\log(1-\frac{p}{n}) - \sum_{i=1}^k (q_i-n+\frac{3}{2})\log(1-\frac{q_i}{n}) - \frac{1}{3}(b(n,p) - \sum_{i=1}^k b(n,q_i)) + O(\frac{p}{n(n-p)}) \\ &= -\frac{\bar{\mu}_n}{n} + o(\sqrt{\beta_{n1}(0)}), \end{aligned}$$

that is,

$$n\Psi_{\ell,n,p}(0) = -\bar{\mu}_n + o(\sqrt{n^2\beta_{n1}(0)}),$$

which coupled with (??), (??) and the fact that $\mu_n = -n\Psi_{\psi,n,p}(0)$ proves the first limit in (??).

Finally, (??) follows from (??) since $\bar{\sigma}_n^2 = n^2\beta_{n1}(0)$ in view of (??). This completes the proof of the lemma. $\square$

5 Proofs of the main results

Proof of Theorem ??.

We will prove (??) under conditions $n - q_{\max} \rightarrow \infty$ and $p \rightarrow \infty$.

Set $V_n = -2 \log \Lambda_n$. Then it follows from (??) that $V_n = -n \log W_n$ and

$$\frac{-(V_n - \mu_n)}{\sigma_n} = \frac{n}{\sigma_n} \log W_n + \frac{\mu_n}{\sigma_n}.$$

In order to prove (??), it is sufficient to show the moment-generating function of $-(V_n - \mu_n)/\sigma_n$ converges to that of the standard normal, that is, for any z ,

$$\lim_{n \rightarrow \infty} \left(\log E(W_n^{nz/\sigma_n}) + \frac{\mu_n z}{\sigma_n} \right) = \frac{z^2}{2}. \quad (5.1)$$

In view of (??), (??) and (??), we have for $t > \frac{p-n}{2}$

$$\log \frac{\Gamma_q(\frac{n-1}{2} + t)}{\Gamma_q(\frac{n-1}{2})} = \Delta_{\log \Gamma, n, q}(t) - \Delta_{\log \Gamma, n, q}(0)$$

and

$$\begin{aligned} \log \left(\mathbf{E}(W_n^t) \right) &= \log \left(\frac{\Gamma_p(\frac{n-1}{2} + t)}{\Gamma_p(\frac{n-1}{2})} \prod_{i=1}^k \frac{\Gamma_{q_i}(\frac{n-1}{2} + t)}{\Gamma_{q_i}(\frac{n-1}{2})} \right) \\ &= \log \frac{\Gamma_p(\frac{n-1}{2} + t)}{\Gamma_p(\frac{n-1}{2})} - \sum_{i=1}^k \log \frac{\Gamma_{q_i}(\frac{n-1}{2} + t)}{\Gamma_{q_i}(\frac{n-1}{2})} \\ &= \Psi_{\log \Gamma, n, p}(t) - \Psi_{\log \Gamma, n, p}(0). \end{aligned} \quad (5.2)$$

Now we apply Taylor's theorem to expand $\Psi_{h, n, p}(t)$ with a second order remainder. For any t with $|t| \leq (n-p)/4$, there exists a c_t with $|c_t| \leq t \leq (n-p)/4$ such that

$$\Psi_{h, n, p}(t) - \Psi_{h, n, p}(0) = \Psi_{h', n, p}(0)t + \Psi_{h'', n, p}(c_t) \frac{t^2}{2}. \quad (5.3)$$

Note that we have used the property given in (??). Similarly, we expand $\Psi_{h, n, p}(t)$ to the third order

$$\Psi_{s, n, p}(t) - \Psi_{s, n, p}(0) = \Psi_{s', n, p}(0)t + \Psi_{s'', n, p}(0) \frac{t^2}{2} + \Psi_{s''', n, p}(\bar{c}_t) \frac{t^3}{6}, \quad (5.4)$$

where $|\bar{c}_t| \leq |t| \leq (n-p)/4$.

We will show

$$\lim_{n \rightarrow \infty} \frac{n^2}{\sigma_n^2(n-p)^2} = 0. \quad (5.5)$$

Note that

$$\sigma_n^2 \geq n^2 \beta_{n1}(0). \quad (5.6)$$

Then it follows from (??) that

$$\sigma_n^2 \geq \frac{2}{3} np \log \left(1 + \frac{p - q_{\max}}{3(n-p)} \right),$$

which coupled with Lemma ?? implies that

$$\frac{n^2}{\sigma_n^2(n-p)^2} \leq \frac{2n}{p(n-p)^2 \log(1 + \frac{p-q_{\max}}{3(n-p)})} = \frac{2}{p(n-p) \log(1 + \frac{p-q_{\max}}{3(n-p)})} + \frac{2}{(n-p)^2 \log(1 + \frac{p-q_{\max}}{3(n-p)})} \rightarrow 0,$$

proving (??).

Now we proceed to prove (??) for any fixed z . For fixed z , set $t = t_n = \frac{nz}{\sigma_n}$. Then it follows from (??) that $t_n = o(n-p)$, which implies $|t_n| \leq (n-p)/4$ for all large n . Therefore, we can apply (??), (??) and (??) with $t = t_n$. From Lemma ?? and (??) we have

$$|\Psi_{h'',n,p}(c_{t_n})|t_n^2 \leq \frac{Dp}{n(n-p)^2} o((n-p)^2) = o(1)$$

and

$$|\Psi_{s''',n,p}(c_{t_n})||t_n|^3 \leq \frac{32\beta_{n1}(0)}{n-p} t_n^2 o(n-p) = o(\frac{n^2\beta_{n1}(0)}{\sigma_n^2}) = o(1).$$

Then by combining (??), (??), (??) and (??) that

$$\begin{aligned} \log(\mathbf{E}(W_n^{nz/\sigma_n})) &= \Psi_{\log \Gamma, n, p}(t_n) - \Psi_{\log \Gamma, n, p}(0) \\ &= \Psi_{s, n, p}(t_n) + \Psi_{h, n, p}(t_n) - (\Psi_{s, n, p}(0) + \Psi_{h, n, p}(0)) \\ &= \Psi_{s, n, p}(t_n) - \Psi_{s, n, p}(0) + \Psi_{h, n, p}(t_n) - \Psi_{h, n, p}(0) \\ &= (\Psi_{s', n, p}(0) + \Psi_{h', n, p}(0))t_n + \Psi_{s'', n, p}(0)\frac{t_n^2}{2} + \Psi_{s''', n, p}(c_{t_n})\frac{t_n^3}{6} + \Psi_{h'', n, p}(c_{t_n})\frac{t_n^2}{2} \\ &= \Psi_{\psi, n, p}(0)t_n + \Psi_{s'', n, p}(0)\frac{t_n^2}{2} + o(1) \\ &= \frac{n\Psi_{\psi, n, p}(0)}{\sigma_n}z + \frac{n^2\Psi_{s'', n, p}(0)}{\sigma_n^2}\frac{z^2}{2} + o(1) \\ &= -\frac{\mu_n z}{\sigma_n} + \frac{z^2}{2} + o(1), \end{aligned}$$

where we have used the following facts

$$n\Psi_{\psi, n, p}(0) = -\mu_n, \quad n^2\Psi_{s'', n, p}(0) = \sigma_n^2.$$

This proves (??). □

Proof of Theorem ??. Without loss of generality, assume $q_1 = q_{\max}$. When $n - q_{\max} = r$ and $n - p = v$ are fixed integers, k is bounded and q_i , $2 \leq i \leq k$ are also bounded. We can employ the subsequence argument to prove (??), that is, for any subsequence of n , we will show that there exists its further subsequence along which (??) holds. Our criterion for selection of subsequences is first to choose a subsequence of the given subsequence of n , along which $k = k_n$ converges to a finite limit, which implies k_n is ultimately a constant, and then to select its further subsequence along which q_2 converges. We repeat the same procedure until we find a subsequence along which q_k converges. The last sub-sequence will be the one along which k and q_i 's are ultimately constant integers. The proof of (??) along such a subsequence is essentially the same as the

proof when $k = k_n$ is fixed and all q_i for $2 \leq i \leq k$ are also fixed integers. For brevity, we will prove (??) by assuming k and q_i for $2 \leq i \leq k$ are constants for all large n .

We first work on W_n defined in (??). From Lemma ??, $\log W_n - r \log n$ has the same distribution as

$$\sum_{i=2}^k \sum_{j=1}^{q_i} \log V_{ij} - r \log n = \sum_{i=2}^k \sum_{j=1}^{q_i} \log n V_{ij},$$

where $q_i^* = \sum_{j=1}^{i-1} q_j$ for $2 \leq i \leq k$, the V_{ij} 's are independent random variables and V_{ij} has a $\text{beta}(\frac{1}{2}(n - q_i^* - j), \frac{1}{2}q_i^*)$ distribution.

Set $\bar{q}_i = q_i^* - q_1 = q_i^* - (p - r) = q_i^* - n + v + r$ for $2 \leq i \leq k$. We have $\bar{q}_2 = 0$, and $\bar{q}_i = \sum_{j=2}^{i-1} q_j$ if $i > 2$. This implies $\bar{q}_i$'s are fixed integers for all large n . For any $1 \leq j \leq q_i$, $2 \leq i \leq k$, V_{ij} has a $\text{beta}(\frac{1}{2}(r + v - \bar{q}_i - j), \frac{1}{2}(n - r - v + \bar{q}_i))$ distribution.

It follows from Section 8.5 in Blitzstein and Hwang [?] that a $\text{beta}(a, b)$ random variable has the same distribution as $X(a)/(X(a) + Y(b))$, where $X(a)$ and $Y(b)$ are two independent random variables, $X(a)$ has a $\text{Gamma}(a, 2)$ distribution with density function $f(x) = x^{a-1}e^{-x/2}/2^a\Gamma(a)$, $x > 0$, and $Y(b)$ has a $\text{Gamma}(b, 2)$ distribution.

Now for each pair of (i, j) with $1 \leq j \leq q_i$ and $2 \leq i \leq k$, set $a = \frac{1}{2}(r + v - \bar{q}_i - j)$ and $b = \frac{1}{2}(n - r - v + \bar{q}_i)$. Then $a + b = \frac{1}{2}(n - j)$. Note that $X(a) = X(\frac{1}{2}(r + v - \bar{q}_i - j))$ and $Y(b) = Y(\frac{1}{2}(n - r - v + \bar{q}_i))$ are independent chi-square random variables with $r + v - \bar{q}_i - j$ and $n - r - v + \bar{q}_i$ degrees of freedom, respectively, and $X(\frac{1}{2}(r + v - \bar{q}_i - j)) + Y(\frac{1}{2}(n - r - v + \bar{q}_i))$ is also a chi-square random variable with $n - j$ degrees of freedom. By using the law of large numbers,

$$\frac{X(\frac{1}{2}(r + v - \bar{q}_i - j)) + Y(\frac{1}{2}(n - r - v + \bar{q}_i))}{n} = \frac{n - j}{n} \frac{X(\frac{1}{2}(r + v - \bar{q}_i - j)) + Y(\frac{1}{2}(n - r - v + \bar{q}_i))}{n - j}$$

converges in probability to 1 as $n \rightarrow \infty$. Since nV_{ij} has the same distribution as

$$X(\frac{1}{2}(r + v - \bar{q}_i - j)) \frac{n}{X(\frac{1}{2}(r + v - \bar{q}_i - j)) + Y(\frac{1}{2}(n - r - v + \bar{q}_i))}$$

which converges in distribution to a chi-square random variable with $r + v - \bar{q}_i - j$ degrees of freedom, that is

$$nV_{ij} \xrightarrow{d} Y_{r+v-\bar{q}_i-j},$$

we have

$$\sum_{i=2}^k \sum_{j=1}^{q_i} \log V_{ij} - r \log n = \sum_{i=2}^k \sum_{j=1}^{q_i} \log n V_{ij} \xrightarrow{d} \sum_{i=2}^k \sum_{j=1}^{q_i} \log Y_{r+v-\bar{q}_i-j} = \sum_{j=v}^{r+v-1} \log Y_j,$$

which is the limiting distribution of $\log W_n - r \log n$. We obtain (??) by noting that

$$\frac{-2 \log \Lambda_n + rn \log n}{n} = -(\log W_n - r \log n).$$

This completes the proof of Theorem ??. □

Proof of Theorem ??.

The sufficiency follows from Theorem ??, that is, under conditions $p \rightarrow \infty$ and $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$, the central limit theorem (??) holds with $a_n = \mu_n$ and $b_n = \sigma_n$.

Now assume (??) holds. We need to show $p \rightarrow \infty$ and $n - q_{\max} \rightarrow \infty$. If any one of the two conditions is not true, there must exist a subsequence of $\{n\}$, say $\{n'\}$, along which

- a.** p is fixed, k is fixed and all q_i 's are fixed, or
- b.** $p - q_{\max} = r$ and $n - p = v$ for some fixed integers $r \geq 1$ and $v \geq 1$.

Condition **b** holds when $n - q_{\max}$ is bounded because both $q - q_{\max}$ and $n - p$ are bounded.

The subsequence $\{n'\}$ along which condition **a** holds can be embedded in an entire sequence along which condition **a** holds. Since the limiting distribution of $-2 \log \Lambda_n$ is a chi-square distribution according to (??), its subsequential limit along $\{n'\}$ cannot be normal. For the same reason, under condition **b**, the subsequential limit is also non-normal from Theorem ??; See Remark 4. Under either condition **a** or condition **b**, it results in a contradiction to the central limit theorem in (??). This completes the proof of the necessity. $\square$

Proof of Theorem ??.

Similar to the proof of Theorem 2 in Qi et al. [?], we use the subsequence argument to prove the theorem. It suffices to prove (??) under each of the following two assumptions:

Case 1: $p_n = p$ and $k_n = k$ and all q_i 's are fixed integers for all large n ;

Case 2: $p_n \rightarrow \infty$ and $n - q_{\max} \rightarrow \infty$ as $n \rightarrow \infty$.

Under Case 1, $f_n = f$ is a constant for all large n , and (??) holds. Since ρ_n defined in (??) converges to one, $-2 \log \Lambda_n$ converges in distribution to a chi-square distribution with f degrees of freedom. Note that Z_n is defined in (??). To prove (??), it suffices to verify that

$$\lim_{n \rightarrow \infty} \frac{2f}{\sigma_n^2} = 1 \text{ and } \lim_{n \rightarrow \infty} (f - \mu_n \sqrt{\frac{2f}{\sigma_n^2}}) = 0. \quad (5.7)$$

Using the notation in the proof of Lemma ??, we have q_i^* 's are fixed integers. it follows from (??) that

$$\beta_{n1}(0) = \frac{2(1+o(1))}{n^2} \sum_{i=2}^k \sum_{j=1}^{q_i} q_i^* = \frac{2(1+o(1))}{n^2} \sum_{i=2}^k q_i q_i^* = \frac{1+o(1)}{n^2} (p^2 - \sum_{i=1}^k q_i^2) = \frac{2(1+o(1))f}{n^2},$$

which together with (??) implies $2f/\sigma_n^2 \rightarrow 1$ and proves the first limit in (??). To prove the second limit, it suffices to show $\lim_{n \rightarrow \infty} \mu_n = f$ or equivalently $\lim_{n \rightarrow \infty} \bar{\mu}_n = f$ by using (??).

From (??) and (??), $n(b(n, p) - \sum_{i=1}^k b(n, q_i)) = o(n\beta_{n1}(0)) = o(\frac{1}{n})$. Then by using Taylor's expansion

we have from (??) that

$$\begin{aligned}
\bar{\mu}_n &= n \sum_{i=1}^k (q_i - n + \frac{3}{2}) \log(1 - \frac{q_i}{n}) - n(p - n + \frac{3}{2}) \log(1 - \frac{p}{n}) + o(\frac{1}{n}) \\
&= n \left(\sum_{i=1}^k (q_i - n + \frac{3}{2}) \left(-\frac{q_i}{n} - \frac{1}{2} \left(\frac{q_i}{n} \right)^2 + O\left(\frac{1}{n^3}\right) \right) - (p - n + \frac{3}{2}) \left(\frac{p}{n} + \frac{1}{2} \left(\frac{p}{n} \right)^2 + O\left(\frac{1}{n^3}\right) \right) \right) + o(\frac{1}{n}) \\
&= \frac{1}{2} (p^2 - \sum_{i=1}^k q_i^2) + O(\frac{1}{n}) \\
&= f + o(1),
\end{aligned}$$

This proves the second limit in (??).

The proof under Case 2 is the same as that in Qi et al. [?], and it is outlined as follows. First, rewrite (??) as

$$\lim_{n \rightarrow \infty} \sup_{-\infty < x < \infty} |P(\frac{Z_n - f_n}{\sqrt{2f_n}} \leq x) - P(\frac{\chi_{f_n}^2 - f_n}{\sqrt{2f_n}} \leq x)| = 0. \quad (5.8)$$

We can show $f_n \rightarrow \infty$ under assumption $\lim_{n \rightarrow \infty} (n - q_{\max}) = \infty$. Since $\chi_{f_n}^2$ can be written as a sum of f_n independent chi-square random variables with one degree of freedom, we have from the central limit theorem that

$$\lim_{n \rightarrow \infty} \sup_{-\infty < x < \infty} |P(\frac{\chi_{f_n}^2 - f_n}{\sqrt{2f_n}} \leq x) - \Phi(x)| = 0.$$

To show (??), it suffices to prove that

$$\frac{Z_n - f_n}{\sqrt{2f_n}} \xrightarrow{d} N(0, 1),$$

which is a direct consequence of Theorem ?? since $\frac{Z_n - f_n}{\sqrt{2f_n}} = \frac{-2 \log \Lambda_n - \mu_n}{\sigma_n}$. This completes the proof of Theorem ??. $\square$

Acknowledgements. The authors would like to thank the two referees whose constructive suggestions have led to improvement in the readability of the paper. The research of Yongcheng Qi was supported in part by NSF Grant DMS-1916014.

References

- [1] M. Abramowitz and I.A. Stegun (1972). Handbook of Mathematical Functions. *Dover, New York*.
- [2] Z. Bai, D. Jiang, J.F. Yao, and S. Zheng (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *The Annals of Statistics* 37(6B), 3822–3840.
- [3] Z. Bao, J. Hu, G. Pan and W. Zhou (2017). Test of independence for high-dimensional random vectors based on freeness in block correlation matrices. *Electronic Journal of Statistics* 11(1), 1527–1548.
- [4] J.K. Blitzstein and J. Hwang (2014). Introduction to Probability. Chapman & Hall/CRC Texts in Statistical Science. CRC Press

- [5] T. Bodnar, H. Dette and N. Parolya (2019). Testing for independence of large dimensional vectors. *The Annals of Statistics* 47(5) 2977–3008.
- [6] S.X. Chen, L.X. Zhang and P.S. Zhong (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association* 105(490), 810–819.
- [7] H. Dette and N. Dörnemann (2020). Likelihood ratio tests for many groups in high dimensions. *Journal of Multivariate Analysis* 178, 104605.
- [8] N. Dörnemann (2022). Likelihood ratio tests under model misspecification in high dimensions. <https://arxiv.org/abs/2203.05423>.
- [9] W. Guo and Y. Qi (2021). Asymptotic distributions for likelihood ratio tests for the equality of covariance matrices. University of Minnesota Duluth, Department of Mathematics and Statistics. *Technical Report*. <http://arxiv.org/abs/2110.02384>.
- [10] D. Jiang, Z. Bai and S. Zheng (2013). Testing the independence of sets of large-dimensional variables. *Science China Mathematics* 56(1), 135–147.
- [11] D. Jiang, T. Jiang and F. Yang (2012). Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. *Journal of Statistical Planning and Inference* 142(8), 2241–2256.
- [12] T. Jiang and Y. Qi (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scandinavian Journal of Statistics* 42(4), 988–1009.
- [13] T. Jiang and F. Yang (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *The Annals of Statistics* 41(4), 2029–2074.
- [14] O. Ledoit and M. Wolf (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics* 30(4), 1081–1102.
- [15] W. Li, J. Chen and J. Yao (2017). Testing the independence of two random vectors where only one dimension is large. *Statistics* 51(1), 141–153.
- [16] R. J. Muirhead (1982). Aspects of multivariate statistical theory. John Wiley & Sons, New York, Wiley Series in Probability and Mathematical Statistics.
- [17] Y. Qi, F. Wang and L. Zhang (2019). Likelihood ratio test of independence of components for high-dimensional normal vectors. *Annals of the Institute of Statistical Mathematics* 71(4), 911–946.
- [18] J.R. Schott (2001). Some tests for the equality of covariance matrices. *Journal of Statistical Planning and Inference* 94(1), 25–36.
- [19] J.R. Schott (2005). Testing for complete independence in high dimensions. *Biometrika* 92(4), 951–956.
- [20] J.R. Schott (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Computational Statistics and Data Analysis* 51(12), 6535–6542.
- [21] M.S. Srivastava and N. Reid (2012). Testing the structure of the covariance matrix with fewer observations than the dimension. *Journal of Multivariate Analysis* 112(C), 156–171.