On the Phase Transition of Wilk's Phenomenon

BY YINQIU HE
Department of Statistics, University of Michigan, MI 48109, U.S.A.
yqhe@umich.edu

BO MENG, ZHENGHAO ZENG
University of Science and Technology of China, Anhui, 230026, P.R. China
mb0529@mail.ustc.edu.cn, zzh98052@mail.ustc.edu.cn

AND GONGJUN XU
Department of Statistics, University of Michigan, MI 48109, U.S.A.
gongjun@umich.edu

SUMMARY

Wilk's theorem, which offers universal chi-squared approximations for likelihood ratio tests, is widely used in many scientific hypothesis testing problems. For modern datasets with increasing dimension, researchers have found that the conventional Wilk's phenomenon of the likelihood ratio test statistic often fails. Although new approximations have been proposed in high dimensional settings, there still lacks a clear statistical guideline regarding how to choose between the conventional and newly proposed approximations, especially for moderate-dimensional data. To address this issue, we develop the necessary and sufficient phase transition conditions for Wilk's phenomenon under popular tests on multivariate mean and covariance structures. Moreover, we provide an in-depth analysis of the accuracy of chi-squared approximations by deriving their asymptotic biases. These results may provide helpful insights into the use of chi-squared approximations in scientific practices.

Some key words: Wilk's phenomenon, phase transition

1. INTRODUCTION

The likelihood ratio test is a standard testing method for many hypothesis testing problems due to its nice statistical properties (Anderson, 2003; Muirhead, 2009). Under the low-dimensional setting with a fixed number of parameters p and large sample size n, classic theorems offer general asymptotic results for various likelihood ratio test statistics. One of the most celebrated and fundamental results is Wilks' theorem, which states that, under the null hypothesis, twice the negative log-likelihood ratio asymptotically approaches a χ^2_f distribution, where f is the difference of the degrees of freedom between the null and alternative hypotheses. The popularly used Bartlett correction provides a general rescaling strategy that further improves the finite sample accuracy of the chi-squared approximations (Cordeiro and Cribari-Neto, 2014; Barndorff-Nielsen and Hall, 1988). Similar Wilk's phenomenon and Bartlett correction were also studied for empirical likelihood (Owen, 1990; DiCiccio et al., 1991; Chen and Cui, 2006).

Despite the extensive literature on the Wilk's-type phenomenon of likelihood ratio tests under finite dimensions, it is of emerging interest to study the large n, diverging p asymptotic regions in a wide variety of modern applications. To understand how large the dimension p can be to ensure the validity of the classical Wilk's phenomenon, various works establish sufficient conditions on the growth rate of p as n increases. For instance, Portnoy (1988) showed that the chi-squared approximation of the likelihood ratio

test statistic for a simple hypothesis in canonical exponential families holds if $p/n^{2/3} \to 0$. Moreover, Hjort et al. (2009), Chen et al. (2009), and Tang and Leng (2010) studied the empirical likelihood ratio statistic when $p \to \infty$. Particularly, Chen et al. (2009) argued that $p/n^{1/2} \to 0$ is likely to be the best rate for the chi-squared approximation of general empirical likelihood ratio test, and showed that for the least-squares empirical likelihood, a simplified version of the empirical likelihood, the chi-squared approximation holds if $p/n^{2/3} \to 0$. The effect of data dimension was also studied in other inference problems; see, for example, Portnoy (1985), He and Shao (2000), and Wang (2011).

When the dimension p further increases, researchers have found that the chi-squared approximations based on Wilk's theorem often become inaccurate, resulting in the failure of the corresponding likelihood ratio tests. To address this issue, various corrections and alternative approximations for the likelihood ratio tests have been proposed. For example, when p is asymptotically proportional to n, namely, $p/n \to y \in (0, 1)$ as $n \to \infty$, Bai et al. (2009), Jiang and Yang (2013), and Jiang and Qi (2015) proposed normal approximations for the corrected likelihood ratio tests on testing mean vectors and covariance matrices. Zheng (2012), Bai et al. (2013), and He et al. (2020) proposed normal approximations for corrected likelihood ratio tests in multivariate linear regression models. Furthermore, Sur and Candès (2019), Sur et al. (2019), and Candès and Sur (2020) studied the phase transition of the maximum likelihood estimator for the logistic regression, and proposed a rescaled chi-squared approximation for the likelihood ratio test.

Despite the proposed distributional theory of the likelihood ratio tests for low- or high-dimensional data, there still lacks a quantitative guideline on which approximation should be chosen to use in practice, especially for moderate-dimensional data. For instance, when analyzing a dataset with the number of parameters $p \leq 5$ and sample size $n = 100$, the chi-squared approximation may be considered as reliable. However, when studying a data set with moderate dimension, e.g., p is between 6 to 20 and sample size $n = 100$, it may be unclear to practitioners whether they can still apply the classical chi-squared approximations or they should turn to other high-dimensional asymptotic results. To address this practical issue, it is of interest to investigate the phase transition boundary where the chi-squared approximation starts to fail as p increases, and also characterize the approximation accuracy. Theoretically, this needs a deep understanding of the limiting behavior of the likelihood ratio test statistics from low to high dimensions.

In this work, we focus on several standard likelihood ratio tests on multivariate mean and covariance structures that are widely used in biomedical and social sciences (Pituch and Stevens, 2015; Cleff, 2019). For each considered likelihood ratio test, we derive its phase transition boundary of Wilk's phenomenon and also provide an in-depth analysis of the accuracy of the chi-squared approximation. First, in terms of the phase transition boundary, we establish the *necessary* and *sufficient* condition for Wilk's theorem to hold when p increases with n. Specifically, we show that the chi-squared approximations hold if and only if $p/n^d \to 0$, where the value of d depends on the testing problem and whether the Bartlett correction is used. Interestingly, the proposed phase transition boundaries resonate with the abovementioned literature (e.g., Portnoy, 1988; Chen et al., 2009), which mostly focused on sufficient conditions without the Bartlett correction. Second, we provide a detailed characterization of the asymptotic bias of each chi-squared approximation. Specifically, we consider two local asymptotic regimes, depending on whether Wilk's theorem holds or not. Under the asymptotic regime when Wilk's theorem holds, the derived asymptotic bias sharply characterizes the convergence rate of the distribution of the likelihood ratio test statistic to the limiting chi-squared distribution, and thus provides a useful measure on the accuracy of the chi-squared approximation. When Wilk's theorem fails, the derived asymptotic bias describes the unignorable discrepancy between the chi-squared approximation and the true distribution of the likelihood ratio test statistic. As illustrated in the simulation studies, our theoretical results of the phase transition boundaries and the asymptotic biases may provide a helpful guideline on the use of the chi-squared approximations in practice.

2. RESULTS OF ONE-SAMPLE TESTS

In this section, we present the theoretical results under three one-sample testing problems. We also obtain similar results for other multiple-sample testing problems, which are introduced in § 4, and please

see their details in the Supplementary Material. Under one-sample problems, suppose $x_1, \ldots, x_n \in R^p$ are independent and identically distributed random vectors with distribution $\mathcal{N}_p(\mu, \Sigma)$, which denotes a p-variate multivariate normal distribution with mean vector μ and covariance matrix Σ. We define $\overline{x} = n^{-1}\sum_{i=1}^n x_i$ and $A = \sum_{i=1}^n (x_i - \overline{x})(x_i - \overline{x})^{\mathrm{T}}$, and denote the determinant and the trace of A by $|A|$ and $\mathrm{tr}(A)$, respectively. We next introduce the considered testing problems and the corresponding likelihood ratio tests (Anderson, 2003; Muirhead, 2009).

(I) *Testing Specified Value for the Mean Vector.* This test examines whether the population mean vector μ is equal to a specified vector $\mu_0 \in R^p$, that is, $H_0 : \mu = \mu_0$ against $H_a : H_0$ is not true. Through the transformation $x_i - \mu_0$, we consider, without loss of generality, $\mu_0 = (0, \ldots, 0)^{\mathrm{T}}$. Then, the likelihood ratio test statistic is $\Lambda_n = |A|^{n/2}(A + n\bar{x}\bar{x}^{\mathrm{T}})^{-n/2}$. When p is fixed and $n \to \infty$, under the null hypothesis, the classical chi-squared approximation without correction is $-2\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $\xrightarrow{d}$ represents the convergence in distribution and $f = p$, and the chi-squared approximation with the Bartlett correction is $-2\rho\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - (1 + p/2)/n$.

(II) *Testing the Sphericity of the Covariance Matrix.* This test examines whether the covariance matrix Σ is proportional to an identity matrix; that is, $H_0 : \Sigma = \lambda \mathrm{I}_p$ against $H_a : H_0$ is not true, where $\lambda > 0$ is an unspecified constant and I_p denotes the $p \times p$ identity matrix. The likelihood ratio test statistic is $\Lambda_n = |A|^{(n-1)/2}\{\mathrm{tr}(A)/p\}^{-p(n-1)/2}$. When p is fixed and $n \to \infty$, under the null hypothesis, the chi-squared approximation is $-2\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $f = (p-1)(p+2)/2$, and the chi-squared approximation with the Bartlett correction is $-2\rho\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - \{6(n-1)p\}^{-1}(2p^2 + p + 2)$.

(III) *Joint Testing Specified Values for the Mean Vector and Covariance Matrix.* Consider a specified vector $\mu_0 \in R^p$ and a specified positive-definite matrix $\Sigma_0 \in R^{p\times p}$. We study the test $H_0 : \mu = \mu_0$ and $\Sigma = \Sigma_0$ against $H_a : H_0$ is not true. By applying the transformation $\Sigma_0^{-1/2}(x_i - \mu_0)$, we assume, without loss of generality, that $\mu_0 = 0$ and $\Sigma_0 = \mathrm{I}_p$. Then, the likelihood ratio test statistic is $\Lambda_n = (e/n)^{np/2}|A|^{n/2}\exp\{-\mathrm{tr}(A)/2 - n\overline{x}^{\mathrm{T}}\overline{x}/2\}$. When p is fixed and $n \to \infty$, under the null hypothesis, the chi-squared approximation is $-2\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $f = p(p+3)/2$, and the chi-squared approximation with the Bartlett correction is $-2\rho\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - \{6n(p+3)\}^{-1}(2p^2 + 9p + 11)$.

For the likelihood ratio tests of the above three testing problems, Theorem 2.1 gives the phase transition boundaries of the chi-squared approximations without and with the Bartlett correction.

THEOREM 2.1. *Assume $n > p + 1$ for all $n \geq 3$ and $n - p \to \infty$ as $n \to \infty$. Under H_0, for the chi-squared approximations without and with the Bartlett correction of each likelihood ratio test in (I)–(III), we have the following necessary and sufficient conditions:*

(i) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\log\Lambda_n > \chi_f^2(\alpha)\} - \alpha| \to 0$ *if and only if* $p/n^{d_1} \to 0$;

(ii) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\rho\log\Lambda_n > \chi_f^2(\alpha)\} - \alpha| \to 0$ *if and only if* $p/n^{d_2} \to 0$,

where the values of d_1 and d_2 under the three testing problems are listed in the table below.

	(I) Mean	*(II) Covariance*	*(III) Joint*
(i) without correction d_1:	$2/3$	$1/2$	$1/2$
(ii) with correction d_2:	$4/5$	$2/3$	$2/3$

In Theorem 2.1, $n > p + 1$ is assumed to ensure the existence of the likelihood ratio tests. We next discuss the obtained phase transition boundaries of the classical chi-squared approximations without correction. When only testing mean parameters, Theorem 2.1 suggests that the chi-squared approximation holds if and only if $p/n^{2/3} \to 0$. This asymptotic regime is similarly assumed in Portnoy (1988), which considered testing p natural parameters in exponential families. However, Portnoy (1988) only showed the sufficiency of $p/n^{2/3} \to 0$ for the chi-squared approximation to be applied, and did not establish the necessary and sufficient result, which is essential for understanding the phase transition behaviors. In addition, when the likelihood ratio tests involve covariance matrices as in (II) and (III), Theorem 2.1 shows that the chi-squared approximation holds if and only if $p/n^{1/2} \to 0$, which is consistent with the discus-

sion in Chen et al. (2009). Particularly, under certain regularity conditions, Chen et al. (2009) established that the chi-squared approximation of the empirical likelihood ratio test holds if $p/n^{1/2} \to 0$. The authors further argued that $p/n^{1/2} \to 0$ is likely to be the best rate for p, because it is the necessary and sufficient condition for the convergence of the sample covariance matrix to the true covariance matrix Σ under the trace norm when the eigenvalues of Σ are bounded. The analysis provides an intuitive explanation for the phase transition boundaries obtained above, and our necessary and sufficient result would serve as another support for their conjecture, despite the different problem settings in Chen et al. (2009) and here.

Additionally, for the chi-squared approximations with the Bartlett correction, Theorem 2.1 also explicitly characterizes their phase transition boundaries, which generally achieve a larger asymptotic region than those without correction. When p is fixed, the Bartlett correction serves as a rescaling strategy that can improve the convergence rate of the likelihood ratio statistic from $O(n^{-1})$ to $O(n^{-2})$; however, when p grows with sample size n, the classical result cannot apply directly. Alternatively, the results in Theorem 2.1 provide a precise illustration of how the Bartlett correction improves the chi-squared approximations in terms of the phase transition boundaries.

The phase transition boundaries in Theorem 2.1 give the necessary and sufficient conditions on the asymptotic regimes of (n, p) in Wilk's phenomenon. When applying the likelihood ratio test in practice, it is desired to have a better understanding of the accuracy of the chi-squared approximation, especially near its phase transition boundary. The following Theorem 2.2 characterizes the accuracy of each chi-squared approximation for tests (I)–(III) when Wilk's phenomenon holds. Specifically, we consider the asymptotic regime where (n, p) satisfies the corresponding necessary and sufficient condition in Theorem 2.1, i.e., $p/n^{d_1} \to 0$ and $p/n^{d_2} \to 0$ for the chi-squared approximations without and with the Bartlett correction, respectively.

THEOREM 2.2. *For each likelihood ratio test (I)–(III), let d_i, $i = 1, 2$ take the corresponding values in Theorem* 2.1. *Let z_α denote the upper α-level quantile of the standard normal distribution. Consider $p \to \infty$ as $n \to \infty$. Then under H_0, given $\alpha \in (0, 1)$,*

(i) when $p/n^{d_1} \to 0$, the chi-squared approximation satisfies

$$\mathrm{pr}\{-2\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha = \frac{\vartheta_1(n,p)}{\sqrt{\pi}} \exp\Big(-\frac{z_\alpha^2}{2}\Big) + o\Big(\frac{p^{1/d_1}}{n}\Big); \tag{1}$$

(ii) when $p/n^{d_2} \to 0$, the chi-squared approximation with the Bartlett correction satisfies

$$\mathrm{pr}\{-2\rho\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha = \frac{\vartheta_2(n,p)}{\sqrt{\pi}} \exp\Big(-\frac{z_\alpha^2}{2}\Big) + o\Big(\frac{p^{2/d_2}}{n^2}\Big). \tag{2}$$

The values of $\vartheta_1(n, p)$ and $\vartheta_2(n, p)$ under three testing problems (I)–(III) are listed below.

(I) $\vartheta_1(n,p) = \dfrac{p^2+2p}{4n\sqrt{f}}$, $\vartheta_2(n,p) = \dfrac{p(p^2-4)}{24(\rho n)^2\sqrt{f}}$;

(II) $\vartheta_1(n,p) = \dfrac{p(2p^2+3p-1)-4/p}{24(n-1)\sqrt{f}}$, $\vartheta_2(n,p) = \dfrac{(p-2)(p-1)(p+2)(2p^3+6p^2+3p+2)}{144p^2\rho^2(n-1)^2\sqrt{f}}$;

(III) $\vartheta_1(n,p) = \dfrac{p\left(2p^2+9p+11\right)}{24n\sqrt{f}}$, $\vartheta_2(n,p) = \dfrac{p(2p^4+18p^3+49p^2+36p-13)}{144(p+3)(\rho n)^2\sqrt{f}}$.

In Theorem 2.2, the forms of $\vartheta_1(n, p)$ and $\vartheta_2(n, p)$ are derived from a nontrivial calculation of certain complicated infinite series (see Eq. (B.20) and (B.28) in the Supplementary Material). We can see that for each test, $\vartheta_1(n, p)$ and $\vartheta_2(n, p)$ are of orders of $p^{1/d_1}n^{-1}$ and $p^{2/d_2}n^{-2}$, respectively. It follows that $\vartheta_1(n, p)\exp(-z_\alpha^2/2)/\sqrt{\pi}$ in (1) is the leading term for the chi-squared approximation bias $\mathrm{pr}\{-2\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha$, and therefore can be used to measure the accuracy of the chi-squared approximation. Similar conclusion also holds for $\vartheta_2(n, p)\exp(-z_\alpha^2/2)/\sqrt{\pi}$ in (2) when using the chi-squared approximation with the Bartlett correction. We demonstrate the usefulness of (1) and (2) in practice by simulation studies in § 3.

In the above discussion, we focus on the local asymptotic regime when Wilk's phenomenon holds, and the derived bias describes the accuracy of the chi-squared approximation. When p further increases beyond this local asymptotic regime, the chi-squared approximation starts to fail, and the approximation bias becomes asymptotically unignorable. The following Theorem 2.3 characterizes such unignorable biases of the chi-squared approximations. Particularly, we consider the local asymptotic regime $p/n \to 0$, which includes the case when Wilk's theorem fails, that is, $p/n^{d_1} \not\to 0$ for the chi-squared approximation, and $p/n^{d_2} \not\to 0$ for the chi-squared approximation with the Bartlett correction.

THEOREM 2.3. *Assume $p \to \infty$ and $p/n \to 0$ as $n \to \infty$. For each likelihood ratio test (I)–(III), under H_0, there exists a small constant $\delta \in (0,1)$ such that for any $\alpha \in (0,1)$,*

(i) the chi-squared approximation satisfies

$$\mathrm{pr}\big\{ -2\log\Lambda_n > \chi_f^2(\alpha)\big\} - \alpha = \bar{\Phi}\left\{ \frac{\chi_f^2(\alpha) + 2\mu_n}{2n\sigma_n} \right\} - \alpha + O\left\{ \left(\frac{p}{n}\right)^{\frac{1-\delta}{2}} + f^{-\frac{1-\delta}{6}} \right\}, \qquad (3)$$

where $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$, and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution;

(ii) the chi-squared approximation with the Bartlett correction satisfies

$$\mathrm{pr}\big\{ -2\rho\log\Lambda_n > \chi_f^2(\alpha)\big\} - \alpha = \bar{\Phi}\left\{ \frac{\chi_f^2(\alpha) + 2\rho\mu_n}{2\rho n\sigma_n} \right\} - \alpha + O\left\{ \left(\frac{p}{n}\right)^{\frac{1-\delta}{2}} + f^{-\frac{1-\delta}{6}} \right\}. \qquad (4)$$

The values of μ_n and σ_n under each problem are listed below, where $L_{x,p} = \log(1 - p/x)$ for $x > p$.

(I) $\mu_n = \dfrac{n}{2}\Big\{\Big(n - p - \dfrac{3}{2}\Big)(L_{n,p} - L_{n-1,p}) + L_{n,p} + pL_{n,1}\Big\}$, $\sigma_n^2 = \dfrac{1}{2}(L_{n,p} - L_{n-1,p})$;

(II) $\mu_n = -\dfrac{n-1}{2}\big\{(n - p - 3/2)L_{n-1,p} + p\big\}$, $\sigma_n^2 = -\dfrac{1}{2}\left(\dfrac{p}{n-1} + L_{n-1,p}\right)\dfrac{(n-1)^2}{n^2}$;

(III) $\mu_n = -\dfrac{n}{2}\big\{(n - p - 3/2)L_{n-1,p} + p\big\} - \dfrac{p}{2}$, $\sigma_n^2 = -\dfrac{1}{2}\left(\dfrac{p}{n-1} + L_{n-1,p}\right)$.

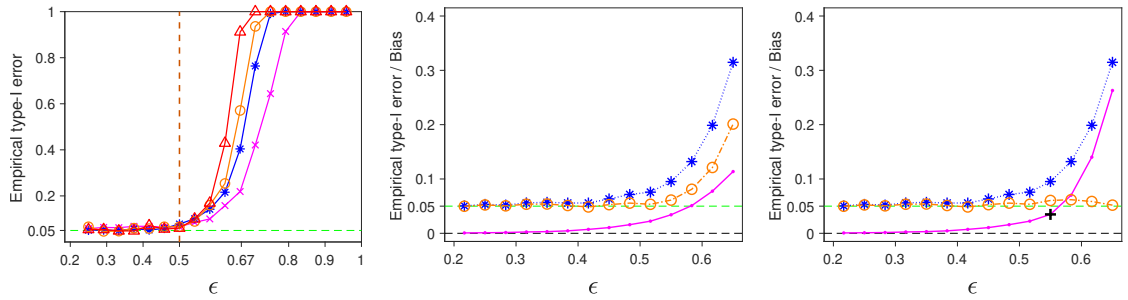
Theorem 2.3 is derived by quantifying the difference between the characteristic functions of $\log\Lambda_n$ and a normal distribution (see Lemma B.3.2 in the Supplementary Material). The local asymptotic regime $p/n \to 0$ is assumed mainly for the technical simplicity of evaluating the asymptotic expansions of the characteristic functions. Under the conditions of Theorem 2.3, $\bar{\Phi}[\{\chi_f^2(\alpha) + 2\mu_n\}/(2n\sigma_n)] - \alpha$ in (3) can be approximated by $\bar{\Phi}\{z_\alpha + (f + 2\mu_n)/(2n\sigma_n)\} - \bar{\Phi}(z_\alpha)$, where $(f + 2\mu_n)/(2n\sigma_n)$ is of the order of pn^{-d_1} (see Remark B.3.2 in the Supplementary Material). Consequently, when the chi-squared approximation fails, i.e., $pn^{-d_1} \not\to 0$, we know that $\bar{\Phi}[\{\chi_f^2(\alpha) + 2\mu_n\}/(2n\sigma_n)] - \alpha$ in (3) characterizes the corresponding unignorable bias of the chi-squared approximation. Similarly, we can show that $\bar{\Phi}[\{\chi_f^2(\alpha) + 2\rho\mu_n\}/(2\rho n\sigma_n)] - \alpha$ can be approximated by $\bar{\Phi}\{z_\alpha + (f + 2\rho\mu_n)/(2\rho n\sigma_n)\} - \bar{\Phi}(z_\alpha)$, where $(f + 2\rho\mu_n)/(2\rho n\sigma_n)$ is of the order of $p^{2/d_2}n^{-2}$. Therefore, when the chi-squared approximation with the Bartlett correction fails, i.e., $pn^{-d_2} \not\to 0$, we know that (4) characterizes the corresponding unignorable approximation bias.

Remark 2.0.1. Although the above discussions consider $p/n^{d_1} \not\to 0$ and $p/n^{d_2} \not\to 0$, (3) and (4) in Theorem 2.3 also hold under the asymptotic regimes $p/n^{d_1} \to 0$ and $p/n^{d_2} \to 0$ examined in Theorem 2.2. However, since Theorems 2.2 and 2.3 focus on different asymptotic regimes and are proved using different techniques, we can show that when $p/n^{d_1} \to 0$ and $p/n^{d_2} \to 0$, (3) and (4) have an additional remainder term $O\{(p/n)^{(1-\delta)/2} + f^{-(1-\delta)/6}\}$ compared to (1) and (2), respectively; see Remark B.3.2 in the Supplementary Material. Therefore, under the asymptotic regimes of Theorem 2.2, (1) and (2) provide a sharper characterization of the accuracy of the chi-squared approximations than (3) and (4), respectively.

3. SIMULATIONS

We conduct simulation studies to evaluate the finite-sample performance of the theoretical results. Particularly, under the null hypothesis of the one-sample tests, we generate data with $\mu = (0, \ldots, 0)^{\mathrm{T}}$ and $\Sigma = \mathrm{I}_p$ and use $\alpha = 0.05$. We next consider problem (III), jointly testing mean and covariance, as an illustration example, and present the results of the chi-squared approximation without the Bartlett correction. For test (III) with the Bartlett correction and problems (I)–(II), testing mean and covariance separately, the simulation results are similar and thus presented in § A.3 of the Supplementary Material.

First, to examine the phase transition boundary in Theorem 2.1, we take $p = \lfloor n^{\epsilon} \rfloor$, where $n \in \{100, 500, 1000, 5000\}$, $\epsilon \in \{6/24, \ldots, 23/24\}$, and $\lfloor \cdot \rfloor$ denotes the floor function. We plot the empirical type-I error versus ϵ in Part (a) of Fig. 1, which is based on 1,000 simulation replications. We can see that for all considered sample sizes, the empirical type-I errors start to inflate around $\epsilon = 1/2$, matching the phase transition boundary $d_1 = 1/2$ of test (III) in Theorem 2.1. Similar results are obtained for other tests as shown in the Supplementary Material.



(a) (b) (c)

Fig. 1: Chi-squared approximation without the Bartlett correction for test (III): (a) Empirical type-I error for $n = 100$ (cross), 500 (asterisk), 1000 (square), and 5000 (triangle); the theoretical phase transition boundary $\epsilon = 1/2$ (vertical dashed line). (b) Empirical type-I error for $n = 500$ (asterisk); asymptotic bias $\vartheta_1(n, p) \exp(-z_\alpha^2/2)/\sqrt{\pi}$ in (1) (dot); the difference between the empirical type-I error and the asymptotic bias in (1) (circle). (c) Empirical type-I error for $n = 500$ (asterisk); the maximum bias over the bias in (1) and the bias $\bar{\Phi}[\{\chi_f^2(\alpha) + 2\mu_n\}/(2n\sigma_n)] - \alpha$ in (3) (dot); the location where the bias in (3) starts to dominate the bias in (1) (plus sign); the difference between the empirical type-I error and the maximum bias (circle).

Second, we numerically evaluate the asymptotic biases in Theorems 2.2 and 2.3 with $p = \lfloor n^{\epsilon} \rfloor$, where $n \in \{100, 500\}$ and $\epsilon \in (0, 1)$. Parts (b) and (c) in Fig. 1 present the results with $n = 500$, while the results with $n = 100$ are similar and thus reported in the Supplementary Material. Part (b) shows that the asymptotic bias in (1) can be an informative indicator of the failure of Wilk's theorem. Particularly, as ϵ increases, the asymptotic bias in (1) increases accordingly. At the ϵ values where the empirical type-I error begins to inflate (e.g. $\epsilon \in [0.4, 0.5]$), the difference between the empirical type-I error and the asymptotic bias is still close to 0.05 as shown in the circle line, suggesting that (1) can approximate the bias well. When ϵ further increases beyond the phase transition boundary (e.g. $\epsilon > 0.5$), the asymptotic bias keeps increasing, and its large value indicates the failure of the chi-squared approximation, even though it now underestimates the approximation bias in this regime. To better characterize the approximation bias when ϵ is beyond the phase transition boundary, we can combine the results in Theorem 2.3 together with those in Theorem 2.2. Specifically, Part (c) shows that taking the maximum over the two asymptotic biases in (1) and (3) gives a good evaluation of the approximation bias for a full range of ϵ, below or above the phase transition boundary. We also find that using (3) itself does not evaluate the approximation bias well for small ϵ (results are not presented). Based on our theoretical and numerical results, when applying Wilk's theorem, we would recommend practitioners to compare the asymptotic bias, either (1) or the maximum over (1) and (3), with a small threshold value that they may specify beforehand, e.g., 0.01-0.02. If the

asymptotic bias is larger than the threshold, the chi-squared approximation should not be directly used, and other methods would be needed.

4. Results of Other Tests

In addition to three one-sample tests in § 2, we also obtain similar theoretical and numerical results for other four popular testing problems in the Supplementary Material. Particularly, we consider three multiple sample tests: (IV) Testing the equality of several mean vectors; (V) Testing the equality of several covariance matrices; (VI) Jointly testing the equality of several mean vectors and covariance matrices. We also study (VII) Testing independence between multiple vectors. Similarly to the results in § 2, for each likelihood ratio test, we establish not only the phase transition boundary of Wilk's theorem, but also the approximation biases under the two asymptotic regimes, where Wilk's theorem holds or not, respectively. Please see the details in § A of the Supplementary Material.

5. Discussion

This study derives the phase transition boundary and characterizes the approximation bias of Wilk's theorem in seven standard likelihood ratio tests. It is interesting to see that the phase transition boundary generally depends on the problem setting and whether the Bartlett correction is used or not, which emphasizes the necessity of statistically-principled guidelines. The approximation bias of Wilk's theorem was also recently studied by Anastasiou and Reinert (2018), which derived an explicit bound of the chi-squared approximation bias for a general family of regular likelihood ratio test statistics. However, as noted in that paper, their bounds are generally not optimized. It is thus of interest to further study the necessary and sufficient conditions for Wilk's phenomenon and the approximation accuracy in such a general setting. Beyond the regular parametric inference problems, Wilk's-type phenomenon has also been studied in geometrically irregular parametric models (Drton and Williams, 2011; Chen et al., 2018), and extended to nonparametric models and statistical learning theory (e.g., Fan et al., 2000, 2001; Fan and Zhang, 2004; Boucheron and Massart, 2011). Understanding the phase transition behavior of Wilk's phenomenon for the likelihood ratio tests would shed light on studying the general Wilk's phenomenon under these complicated statistical models. Besides the likelihood ratio tests, similar phase transition phenomena can also occur for other popular test statistics. For instance, Xu et al. (2019) recently studied the approximation theory for Pearson's chi-squared statistics when the number of cells is large, and demonstrated a similar phase transition phenomenon that the asymptotic distribution of the test statistic can be either a chi-squared or a normal distribution. It is interesting to further investigate the phase transition boundaries of these tests.

Acknowledgement

The authors are grateful to the editor, Professor Paul Fearnhead, an associate editor and three referees for their valuable comments and suggestions. This research was partially supported by the U.S. National Science Foundation.

Supplementary material

The supplementary material available at *Biometrika* online includes theoretical results for the other four testing problems in Section 4, additional simulation studies, and the proofs of the theorems.

References

Abramowitz, M. and I. A. Stegun (1970). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (9th ed.), Volume 55. US Government printing office.

Anastasiou, A. and G. Reinert (2018). Bounds for the asymptotic distribution of the likelihood ratio. *arXiv preprint arXiv:1806.03666*.

Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis*. Wiley Series in Probability and Statistics. Wiley.

Bai, Z., 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.

Bai, Z., D. Jiang, J.-f. Yao, and S. Zheng (2013). Testing linear hypotheses in high-dimensional regressions. *Statistics 47*(6), 1207–1223.

Barndorff-Nielsen, O. and P. Hall (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika 75*(2), 374–378.

Boucheron, S. and P. Massart (2011). A high-dimensional Wilks phenomenon. *Probability theory and related fields 150*(3-4), 405–433.

Candès, E. J. and P. Sur (2020). The phase transition for the existence of the maximum likelihood estimate in high-dimensional logistic regression. *The Annals of Statistics*.

Chen, S. X. and H. Cui (2006). On bartlett correction of empirical likelihood in the presence of nuisance parameters. *Biometrika 93*(1), 215–220.

Chen, S. X., L. Peng, and Y.-L. Qin (2009). Effects of data dimension on empirical likelihood. *Biometrika 96*(3), 711–722.

Chen, Y., J. Huang, Y. Ning, K.-Y. Liang, and B. G. Lindsay (2018). A conditional composite likelihood ratio test with boundary constraints. *Biometrika 105*(1), 225–232.

Cleff, T. (2019). *Applied Statistics and Multivariate Data Analysis for Business and Economics: A Modern Approach Using SPSS, Stata, and Excel*. Springer.

Cordeiro, G. M. and F. Cribari-Neto (2014). *An introduction to Bartlett correction and bias reduction*. Springer.

DiCiccio, T., P. Hall, and J. Romano (1991). Empirical likelihood is Bartlett-correctable. *the Annals of Statistics 19*(2), 1053–1061.

Drton, M. and B. Williams (2011). Quantifying the failure of bootstrap likelihood ratio tests. *Biometrika 98*(4), 919–934.

Fan, J., H.-N. Hung, and W.-H. Wong (2000). Geometric understanding of likelihood ratio statistics. *Journal of the American Statistical Association 95*(451), 836–841.

Fan, J., C. Zhang, and J. Zhang (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of statistics 29*(1), 153–193.

Fan, J. and W. Zhang (2004). Generalised likelihood ratio tests for spectral density. *Biometrika 91*(1), 195–209.

He, X. and Q.-M. Shao (2000). On parameters of increasing dimensions. *Journal of Multivariate Analysis 73*(1), 120–135.

He, Y., T. Jiang, J. Wen, and G. Xu (2020). Likelihood ratio test in multivariate linear regression: from low to high dimension. *Statistica Sinica*.

Hjort, N. L., I. W. McKeague, and I. Van Keilegom (2009). Extending the scope of empirical likelihood. *The Annals of Statistics 37*(3), 1079–1111.

Jiang, T. and Y. Qi (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scandinavian Journal of Statistics 42*(4), 988–1009.

Jiang, T. 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.

Luke, Y. L. (1969). *Special functions and their approximations*, Volume 2. Academic press.

Muirhead, R. J. (2009). *Aspects of multivariate statistical theory*, Volume 197. John Wiley & Sons.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 90–120.

Pituch, K. A. and J. P. Stevens (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBMs SPSS*. Routledge.

Portnoy, S. (1985). Asymptotic behavior of M estimators of p regression parameters when p^2/n is large; II. Normal approximation. *The Annals of Statistics*, 1403–1417.

Portnoy, S. (1988). Asymptotic behavior of likelihood methods for exponential families when the number of parameters tends to infinity. *The Annals of Statistics*, 356–366.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1992). *Numerical recipes in Fortran 77: the art of scientific computing*. Cambridge university press.

Sur, P. and E. J. Candès (2019). A modern maximum-likelihood theory for high-dimensional logistic regression. *Proceedings of the National Academy of Sciences 116*(29), 14516–14525.

Sur, P., Y. Chen, and E. J. Candès (2019). The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square. *Probability Theory and Related Fields 175*(1-2), 487–558.

Tang, C. Y. and C. Leng (2010). Penalized high-dimensional empirical likelihood. *Biometrika 97*(4), 905–920.

Ushakov, N. G. (2011). *Selected topics in characteristic functions*. Walter de Gruyter.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, Volume 3. Cambridge university press.

Wang, L. (2011). GEE analysis of clustered binary data with diverging number of covariates. *The Annals of Statistics 39*(1), 389–417.

Whittaker, E. T. and G. N. Watson (1996). *A course of modern analysis*. Cambridge university press.
Xu, M., D. Zhang, and W. B. Wu (2019). Pearson's chi-squared statistics: approximation theory and beyond. *Biometrika 106*(3), 716–723.
Zheng, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Annales de l'IHP Probabilités et statistiques 48*(2), 444–476.
Zwillinger, D. (2002). *CRC standard mathematical tables and formulae* (31st ed.). CRC press.

Supplementary Material for "On the Phase Transition of Wilk's Phenomenon"

In this supplementary material, we present additional results in § A. Particularly, the theoretical results for tests (IV)–(VI) and test (VII) are given in § A.1 and § A.2, respectively. All the simulations for tests (I)–(VII) are provided in § A.3. We next present the proofs for the testing problem (III) as an illustration example in § B, where the corresponding results in Theorems 2.1–2.3 are proved in §§ B.1–B.3, respectively. The proofs for other tests are similar and given in § C. The technical lemmas are proved in § D.

A. ADDITIONAL RESULTS

A.1. *Multiple-Sample Tests*

This subsection presents the theoretical results of three multiple-sample tests (IV)–(VI). Under the multiple-sample problems, let k denote the number of samples, which is assumed to be fixed compared to the sample size. In each sample $i = 1, \ldots, k$, the observations $x_{i1}, \cdots, x_{in_i}$ are independent and identically distributed $\mathcal{N}_p(\mu_i, \Sigma_i)$ random vectors. In this subsection, we define $\overline{x}_i = n_i^{-1} \sum_{j=1}^{n_i} x_{ij}$ and $A_i = \sum_{j=1}^{n_i} (x_{ij} - \overline{x}_i)(x_{ij} - \overline{x}_i)^{\mathrm{T}}$ for $i = 1, \ldots, k$, and let $A = A_1 + \ldots + A_k$ and $n = n_1 + \ldots + n_k$. We next briefly review the likelihood ratio tests for the problems (IV)–(VI).

(IV) *Testing the Equality of Several Mean Vectors.* Consider $H_0 : \mu_1 = \ldots = \mu_k$ agains $H_a : H_0$ is not true, where the covariances of the k samples are assumed to be the same. Define $B = \sum_{i=1}^{k} n_i(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})^{\mathrm{T}}$ and $\overline{x} = n^{-1} \sum_{i=1}^{k} n_i \overline{x}_i$. Then, the likelihood ratio test statistic is $\Lambda_n = |A|^{n/2}|A + B|^{-n/2}$. When p is fixed and $n \to \infty$, the chi-squared approximation is $-2 \log \Lambda_n \xrightarrow{d} \chi_f^2$, where $f = (k-1)p$, and the chi-squared approximation with the Bartlett correction is $-2\rho \log \Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - \{1 + (k+p)/2\}/n$.

(V) *Testing the Equality of Several Covariance Matrices.* Consider $H_0 : \Sigma_1 = \ldots = \Sigma_k$ against $H_a : H_0$ is not true. For this test, $\Lambda_n = |A|^{-(n-k)/2}(n-k)^{(n-k)p/2} \times \prod_{i=1}^{k}(n_i - 1)^{-(n_i-1)p/2}|A_i|^{(n_i-1)/2}$ is the modified likelihood ratio test statistic with the unbiasedness property. When p is fixed and $\min_{1 \le i \le k} n_i \to \infty$, the chi-squared approximation is $-2 \log \Lambda_n \xrightarrow{d} \chi_f^2$, where $f = p(p+1)(k-1)/2$, and the chi-squared approximation with the Bartlett correction is $-2\rho \log \Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - \{6(p+1)(k-1)\}^{-1}(2p^2 + 3p - 1)\{\sum_{i=1}^{k}(n_i - 1)^{-1} - (n-k)^{-1}\}$.

(VI) *Joint Testing the Equality of Mean Vectors and Covariance Matrices.* Consider $H_0 : \mu_1 = \ldots = \mu_k,\ \Sigma_1 = \ldots = \Sigma_k$ against $H_a : H_0$ is not true. The likelihood ratio test statistic is $\Lambda_n = n^{pn/2}|A + B|^{-n/2} \times \prod_{i=1}^{k} n_i^{-pn_i/2}|A_i|^{n_i/2}$. When p is fixed and $\min_{1 \le i \le k} n_i \to \infty$, the chi-squared approximation is $-2 \log \Lambda_n \xrightarrow{d} \chi_f^2$, where $f = p(k-1)(p+3)/2$, and the chi-squared approximation with the Bartlett correction is $-2\rho \log \Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - \{6(k-1)(p+3)\}^{-1}(2p^2 + 9p + 11)(\sum_{i=1}^{k} n_i^{-1} - n^{-1})$.

For the likelihood ratio tests (IV)–(VI), Theorem A.1 gives the phase transition boundaries of the chi-squared approximations without and with the Bartlett correction.

THEOREM A.1. *Assume $n_i > p + 1$ for $i = 1, \ldots, k$, and there exists a constant $\delta \in (0, 1)$ such that $\delta < n_i/n_j < \delta^{-1}$ for any $1 \le i, j \le k$. Under H_0, for the chi-squared approximations without and with the Bartlett correction, we have the following necessary and sufficient conditions:*

(i) $\sup_{\alpha \in (0,1)} |\mathrm{pr}\{-2 \log \Lambda_n > \chi_f^2(\alpha)\} - \alpha| \to 0$ *if and only if* $p/n^{d_1} \to 0$;

(ii) when $p = o(n)$, $\sup_{\alpha \in (0,1)} |\mathrm{pr}\{-2\rho \log \Lambda_n > \chi_f^2(\alpha)\} - \alpha| \to 0$ *if and only if* $p/n^{d_2} \to 0$,

where the values of d_1 and d_2 under the three testing problems are listed in the table below.

	(IV) Mean	*(V) Covariance*	*(VI) Joint*
(i) without correction d_1:	$2/3$	$1/2$	$1/2$
(ii) with correction d_2:	$4/5$	$2/3$	$2/3$

In Theorem A.1, the boundedness of n_i/n_j suggests that the sizes of all the samples are comparable. The additional regularity condition $p = o(n)$ in (ii) specifies a local asymptotic region, which is of practical interest, and simulation studies suggest that the conclusion can hold more generally without this condition. With a fixed k, the phase transition boundaries in Theorem A.1 are parallel to those in Theorem 2.1, and the analyses after Theorem 2.1 apply to Theorem A.1 similarly. Particularly, examining covariances or not will yield different phase transition boundaries in the three problems. When k also increases with n, the phase transition boundaries would involve k, p, and n, as illustrated in the following proposition.

PROPOSITION A.1. *Consider $n > p + k$, $n - k \to \infty$, and $n - p \to \infty$. For Λ_n in problem (IV), under H_0, as $n \to \infty$,*

(i) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha| \to 0$ *if and only if* $\sqrt{pk}(p+k)/n \to 0$;
(ii) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\rho\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha| \to 0$ *if and only if* $\sqrt{pk}(p^2+k^2)/n^2 \to 0$.

Proposition A.1 suggests that the total number of samples k and the dimension of each observation p play symmetric roles in the phase transition boundary of problem (IV). When k is fixed, Proposition A.1 is consistent with Theorem A.1. To further illustrate the cases with increasing k, we consider $p = \lfloor n^\epsilon \rfloor$ and $k = \lfloor n^\eta \rfloor$, where $0 < \epsilon, \eta < 1$ and $\lfloor\cdot\rfloor$ denotes the floor of a number. Then the two phase transition boundaries in Proposition A.1 become (i) $\max\{\epsilon,\eta\} + (\epsilon+\eta)/2 < 1$ and (ii) $\max\{\epsilon,\eta\} + (\epsilon+\eta)/4 < 1$, respectively. Specifically, for (i), when ϵ is close to 0, the largest value of η is around $2/3$, and vice versa; when $\epsilon = \eta$, suggesting p and k are of the same order, the largest value of ϵ is $1/2$. For (ii), when ϵ is close to 0, the largest value of η is around $4/5$, and vice versa; when $\epsilon = \eta$, the largest value of ϵ becomes $2/3$.

In addition to the phase transition boundaries above, the following Theorem A.2, similarly to Theorem 2.2, further characterizes the accuracy of each chi-squared approximation for tests (IV)–(VI) when Wilk's theorem holds. Specifically, we consider $p/n^{d_1} \to 0$ and $p/n^{d_2} \to 0$ for the chi-squared approximations without and with the Bartlett correction, respectively.

THEOREM A.2. *Assume that there exists a constant $\delta \in (0,1)$ such that $\delta < n_i/n_j < \delta^{-1}$ for any $1 \le i, j \le k$, and $p \to \infty$ as $n \to \infty$. For each likelihood ratio test (IV)–(VI), let d_i, $i = 1, 2$ take the corresponding values in Theorem A.1. Then under H_0, for any $\alpha \in (0,1)$,*

(i) when $p/n^{d_1} \to 0$, (1) *in Theorem* 2.2 *holds with the value of $\vartheta_1(n,p)$ listed below;*
(ii) when $p/n^{d_2} \to 0$, (2) *in Theorem* 2.2 *holds with the values of $\vartheta_2(n,p)$ listed below.*

Let $D_{n,r} = \sum_{i=1}^k n_i^{-r} - n^{-r}$ and $\tilde{D}_{n,r} = \sum_{i=1}^k (n_i-1)^{-r} - (n-k)^{-r}$.

(IV) Mean:
$$\vartheta_1(n,p) = \frac{p(k-1)(p+2+k)}{4n\sqrt{f}},$$
$$\vartheta_2(n,p) = \frac{(k-1)p(p^2+k^2-2k-4)}{24n^2\rho^2\sqrt{f}};$$

(V) Covariance:
$$\vartheta_1(n,p) = \frac{\tilde{D}_{n,1}p(2p^2+3p-1)}{24\sqrt{f}},$$
$$\vartheta_2(n,p) = \frac{p(p+1)}{24\rho^2\sqrt{f}}\Big\{(p-1)(p+2)\tilde{D}_{n,2} - 6(k-1)(1-\rho)^2\Big\};$$

(VI) Joint:
$$\vartheta_1(n,p) = \frac{D_{n,1}p\left(2p^2+9p+11\right)}{24\sqrt{f}},$$
$$\vartheta_2(n,p) = \frac{p(p+3)}{24\rho^2\sqrt{f}}\Big\{(p+1)(p+2)D_{n,2} - 6(k-1)(1-\rho)^2\Big\}.$$

Theorem A.2 shows that for multiple-sample tests (IV)–(VI), (1) and (2) in Theorem 2.2 still hold. However, the values of $\vartheta_1(n,p)$ and $\vartheta_2(n,p)$ depend on the testing problems, and are different from those in Theorem 2.2. Similarly to Theorem 2.2, in each test (IV)–(VI), we also know that $\vartheta_1(n,p)$ and $\vartheta_2(n,p)$ are of the orders of $p^{1/d_1}n^{-1}$ and $p^{2/d_2}n^{-2}$, respectively. Then $\vartheta_1(n,p)\exp(-z_\alpha^2/2)/\sqrt{\pi}$ in (1) and $\vartheta_2(n,p)\exp(-z_\alpha^2/2)/\sqrt{\pi}$ in (2) are the leading terms of the biases of the chi-squared approximations without and with the Bartlett correction, respectively. We can similarly use the derived asymptotic biases to measure the approximation accuracy, and please see the simulation studies for multiple-sample tests (IV)–(VI) in § A.3.

Theorem A.2 focuses on the local asymptotic regime of (n, p) when Wilk's theorem holds. When p further increases such that Wilk's theorem fails, the biases of the chi-squared approximations become unignorable. The following Theorem A.3 characterizes such unignorable biases of the chi-squared approximations in testing problems (IV)–(VI). Similarly to Theorem 2.3, we consider a general local asymptotic regime $p/n \to 0$, which includes the case when Wilk's theorem fails, i.e., $p/n^{d_1} \not\to 0$ and $p/n^{d_2} \not\to 0$ for the chi-squared approximations without and with the Bartlett correction, respectively.

THEOREM A.3. *Assume that there exists a constant $\delta \in (0,1)$ such that $\delta < n_i/n_j < \delta^{-1}$ for any $1 \le i, j \le k$. Moreover, assume $p \to \infty$ and $p/n_i \to 0$ as $n_i \to \infty$. For each likelihood ratio test (I)–(III), under H_0, for any $\alpha \in (0,1)$,* (3) *and* (4) *in Theorem* 2.3 *hold under three testing problems (IV)–(VI) with μ_n and σ_n listed below.*

(IV) Mean:
$$\mu_n = \frac{n}{2}\big\{(n-p-k-1/2)(L_{n-1,p} - L_{n-k,p}) + (k-1)L_{n-1,p} + pL_{n-1,k-1}\big\},$$
$$\sigma_n^2 = \frac{1}{2}\big(L_{n-1,p} - L_{n-k,p}\big);$$

(V) Covariance:
$$\mu_n = \frac{1}{2}\sum_{i=1}^{k}(n_i-1)\Big\{(n-p-k-1/2)L_{n-k,p} - (n_i-p-3/2)\,L_{n_i-1,p}\Big\},$$
$$\sigma_n^2 = \frac{(n-k)^2}{2n^2}\left\{L_{n-k,p} - \sum_{i=1}^{k}\left(\frac{n_i-1}{n-k}\right)^2 L_{n_i-1,p}\right\};$$

(VI) Joint:
$$\mu_n = \frac{1}{2}\left[-kp + n\Big(n-p-\frac{3}{2}\Big)L_{n,p} - \sum_{i=1}^{k}\left\{\frac{p}{2n_i} + n_i\Big(n_i-p-\frac{3}{2}\Big)L_{n_i-1,p}\right\}\right],$$
$$\sigma_n^2 = \frac{1}{2}\left(L_{n,p} - \sum_{i=1}^{k}\frac{n_i^2}{n^2}\times L_{n_i-1,p}\right).$$

Theorem A.3 shows that (3) and (4) still hold for multiple-sample tests (IV)–(VI), where the values of μ_n and σ_n^2 depend on the specific testing problem. Similarly to Theorem 2.3, the analysis in Remark B.3.2 also applies here, and we know that when $pn^{-d_1} \not\to 0$, (3) characterizes the unignorable biases for the chi-squared approximation, and when $pn^{-d_2} \not\to 0$, (4) characterizes the unignorable biases for the chi-squared approximation with the Bartlett correction. Moreover, the analysis in Remark 2.0.1 also applies similarly to the multiple-sample tests (IV)–(VI), and thus is not repeated here.

A.2. *Testing Independence between Multiple Vectors*

This subsection studies testing the independence between k sets of multivariate normal variables. Suppose $x_1, \ldots, x_n \in R^p$ are independent and identically distributed $\mathcal{N}_p(\mu, \Sigma)$ random vectors, and we partition x_i and Σ as $x_i = (\xi_{i1}^{\mathrm{T}}, \ldots, \xi_{ik}^{\mathrm{T}})^{\mathrm{T}}$ and $\Sigma = (\Sigma_{jl})_{1\le j,l\le k}$, respectively, where $\xi_{i,j}$ is of size $p_j \times 1$, Σ_{jl} is a $p_j \times p_l$ sub-matrix of Σ, and $\sum_{j=1}^{k} p_j = p$. In this subsection, we define $\overline{x} = n^{-1}\sum_{i=1}^{n} x_i$, $\bar{\xi}_j = n^{-1}\sum_{i=1}^{n}\xi_{ij}$, $A = \sum_{i=1}^{n}(x_i - \overline{x})(x_i - \overline{x})^{\mathrm{T}}$, and $A_{jj} = \sum_{i=1}^{n}(\xi_{ij} - \bar{\xi}_j)(\xi_{ij} - \bar{\xi}_j)^{\mathrm{T}}$.

(VII) *Testing Independence of Subvectors of Multivariate Normal Distribution.* For the multivariate normal distribution, testing the independence between k sets of vectors $\xi_{i,1}, \ldots, \xi_{i,k}$ is equivalent to testing $H_0 : \Sigma_{jl} = 0$, for $1 \le j < l \le k$, against $H_a : H_0$ is not true. The likelihood ratio statistic is $\Lambda_n = |A|^{n/2}\prod_{j=1}^{k}|A_{jj}|^{-n/2}$. When $p_1, \ldots, p_k$ are fixed, the chi-squared approximation is $-2\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $f = (p^2 - \sum_{i=1}^{k} p_i^2)/2$; the chi-squared approximation with the Bartlett correction is $-2\rho\log\Lambda_n \xrightarrow{d} \chi_f^2$, where $\rho = 1 - (3/2n)^{-1} - \{3n(p^2 - \sum_{i=1}^{k} p_i^2)\}^{-1}(p^3 - \sum_{i=1}^{k} p_i^3)$.

Theorem A.4 below gives the phase transition boundaries of the chi-squared approximations without and with the Bartlett correction for test (VII).

THEOREM A.4. *Assume $n > p+1$ and there exists $\delta \in (0,1)$ such that $\delta < p_i/p_j < \delta^{-1}$ for $1 \le i, j \le k$. For Λ_n in problem (VII), under H_0, as $n \to \infty$,*

(i) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha| \to 0$ *if and only if* $p/n^{1/2} \to 0$;
(ii) when $p = o(n)$, $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\rho\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha| \to 0$ *if and only if* $p/n^{2/3} \to 0$.

The phase transition boundaries in Theorem A.4 are consistent with those in Theorems 2.1 and A.1 for testing problems (II), (III), (V), and (VI). This is reasonable because testing independence between multivariate normal vectors examines the structures of covariance matrices. In Theorem A.4, the boundedness of p_i/p_j suggests that the dimensions of the multiple vectors are comparable. The following Proposition A.2 relaxes this constraint for $k = 2$, a case closely related to the canonical correlation analysis.

PROPOSITION A.2. *Consider* $n > p_1 + p_2$ *and* $n - \max\{p_1, p_2\} \to \infty$. *For* Λ_n *in problem (VII), under* H_0, *as* $n \to \infty$,

(i) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha| \to 0$ *if and only if* $\sqrt{p_1p_2}(p_1 + p_2)/n \to 0$;
(ii) $\sup_{\alpha\in(0,1)} |\mathrm{pr}\{-2\rho\log\Lambda_n > \chi^2_f(\alpha)\} - \alpha| \to 0$ *if and only if* $\sqrt{p_1p_2}(p_1^2 + p_2^2)/n^2 \to 0$.

Proposition A.2 shows that the effects of p_1 and p_2 on the phase transition boundaries are symmetric. To further illustrate, consider $p_1 = \lfloor n^\epsilon \rfloor$ and $p_2 = \lfloor n^\eta \rfloor$, where $0 < \epsilon, \eta < 1$. Then the two phase transition boundaries in Proposition A.2 become (i) $\max\{\epsilon, \eta\} + (\epsilon + \eta)/2 < 1$ and (ii) $\max\{\epsilon, \eta\} + (\epsilon + \eta)/4 < 1$, respectively. When $\epsilon = \eta$, i.e., p_1 and p_2 are of the same order, the largest value of ϵ and η achievable is (i) $1/2$ and (ii) $2/3$ respectively, which are consistent with Theorem A.4. When η is close to 0, the largest value of ϵ is (i) $2/3$ and (ii) $4/5$ respectively. Therefore when one set of the vectors is of finite dimension, the chi-squared approximations without and with the Bartlett correction can be applied when $p/n^{2/3} \to 0$ and $p/n^{4/5} \to 0$, respectively. This demonstrates an interesting phenomenon that for the phase transition boundary, the growth rate of p changes as the ratio of p_1 and p_2 varies.

Similarly to Theorems 2.2 and A.2, the following Theorem A.5 further characterizes the accuracy of the chi-squared approximation under the asymptotic regime where p satisfies the corresponding necessary and sufficient conditions in Theorem A.4.

THEOREM A.5. *Assume that there exists* $\delta \in (0, 1)$ *such that* $\delta < p_i/p_j < \delta^{-1}$ *for* $1 \le i, j \le k$, *and* $p \to \infty$ *as* $n \to \infty$. *Let* $d_1 = 1/2$ *and* $d_2 = 2/3$ *as in Theorem* A.4. *For* Λ_n *in problem (VII), under* H_0, *for any* $\alpha \in (0, 1)$,

(i) when $p/n^{d_1} \to 0$, (1) *in Theorem* 2.2 *holds with the value of* $\vartheta_1(n, p)$ *below;*
(ii) when $p/n^{d_2} \to 0$, (2) *in Theorem* 2.2 *holds with the value of* $\vartheta_2(n, p)$ *below.*

Let $D_{p,r} = p^r - \sum_{j=1}^k p_j^r$. *Then*

$$\vartheta_1(n,p) = \frac{2D_{p,3} + 9D_{p,2}}{24n\sqrt{f}}, \qquad \vartheta_2(n,p) = \frac{1}{(\rho n)^2\sqrt{f}}\left(\frac{1}{24}D_{p,4} - \frac{5D_{p,2}}{48} - \frac{D_{p,3}^2}{36D_{p,2}}\right).$$

Similar to Theorems 2.2 and A.2, Theorem A.5 focuses on the local asymptotic regime when Wilk's theorem holds, and we know from a similar analysis that (1) and (2) provide useful information on the accuracy of the chi-squared approximations. Please see the simulations for test (VII) in § A.3. When p further increases such that Wilk's theorem fails, the following Theorem A.6 characterizes the unignorable chi-squared approximation biases for test (VII) similarly as in Theorems 2.3 and A.3.

THEOREM A.6. *Assume that there exists* $\delta \in (0, 1)$ *such that* $\delta < p_i/p_j < \delta^{-1}$ *for* $1 \le i, j \le k$, *and* $p \to \infty$ *and* $p/n \to 0$ *as* $n \to \infty$. *For* Λ_n *in problem (VII), under* H_0, *as* $n \to \infty$, *for any* $\alpha \in (0, 1)$, (3) *and* (4) *in Theorem* 2.3 *hold with* μ_n *and* σ_n *listed below.*

$$\mu_n = \frac{n}{2}\bigg[- \left(n - p - \frac{3}{2}\right) L_{n-1,p} + \sum_{j=1}^k \left\{ \left(n - p_j - \frac{3}{2}\right) L_{n-1,p_j} \right\} \bigg],$$

$$\sigma_n^2 = \frac{1}{2}\bigg(- L_{n-1,p} + \sum_{j=1}^k L_{n-1,p_j} \bigg).$$

Note that Theorem A.6 is analogous to Theorems 2.3 and A.3, and therefore similar analyses and conclusions as in Remarks 2.0.1 and B.3.2 also hold for test (VII), which are not repeated here.

A.3. *Additional Simulations*

We next introduce the simulation settings of each test and afterwards analyze the numerical results.

A.3.1. *One-Sample Tests (I)–(III).* Similarly to Section 3, under the null hypothesis of each one-sample test (I)–(III), we set $\mu = (0, \ldots, 0)^{\mathrm{T}}$ and $\Sigma = \mathrm{I}_p$.

(1) On the phase transition boundaries. We take $p = \lfloor n^{\epsilon} \rfloor$, where $n \in \{100, 500, 1000, 5000\}$ and $\epsilon \in \{6/24, \ldots, 23/24\}$. We next plot the empirical type-I error rates (over 1000 replications) versus ϵ for each chi-squared approximation in Fig. 2. We still include the results in § 3 for easy presentation of the figure.

(2) On the asymptotic biases. To evaluate the asymptotic biases in Theorems 2.2 and 2.3, we take $p = \lfloor n^{\epsilon} \rfloor$, where $n \in \{100, 500\}$ and $\epsilon \in (0, 1)$. The results of $n = 100$ and 500 (over 3000 replications) are given in Fig. 4 and Fig. 5, respectively. In each setting, the range of ϵ is chosen such that the largest empirical type-I error is below 0.5.

To facilitate the presentation of figures and the discussions below, we define

$$\varpi_1 = \vartheta_1(n,p)\exp(-z_\alpha^2/2)/\sqrt{\pi}, \quad \varpi_3 = \bar{\Phi}\big[\{\chi_f^2(\alpha) + 2\mu_n\}/(2n\sigma_n)\big] - \alpha,$$
$$\varpi_2 = \vartheta_2(n,p)\exp(-z_\alpha^2/2)/\sqrt{\pi}, \quad \varpi_4 = \bar{\Phi}\big[\{\chi_f^2(\alpha) + 2\rho\mu_n\}/(2\rho n\sigma_n)\big] - \alpha.$$

Then $\varpi_1, \varpi_2, \varpi_3$, and ϖ_4 denote the asymptotic biases in (1)–(4), respectively. For each test in Fig. 4 and Fig. 5, we plot ϖ_1 and ϖ_2 in the subfigures in the columns (a) and (c), respectively. Similarly to § 3, to better characterize each approximation bias when ϵ is beyond the corresponding phase transition boundary, we combine the results in Theorem 2.2 and those in Theorem 2.3. Specifically, in the column (b) of Fig. 4 and Fig. 5, we plot $M_c(\varpi_1, \varpi_3) \equiv \varpi_1 1\{\varpi_1 < c\} + \max\{\varpi_1, \varpi_3\}1\{\varpi_1 \geq c\}$, where $1\{\cdot\}$ denotes an indicator function, and c denotes a small positive threshold, and we choose $c = 0.002$ in the simulations. This definition of $M_c(\varpi_1, \varpi_3)$ suggests that ϖ_1 is used when the approximation bias is smaller than c, and $\max\{\varpi_1, \varpi_3\}$ is used when the approximation bias becomes larger. Similarly, we define $M_c(\varpi_2, \varpi_4) \equiv \varpi_2 1\{\varpi_2 < c\} + \max\{\varpi_2, \varpi_4\}1\{\varpi_2 \geq c\}$, and plot it in the column (d) of Fig. 4 and Fig. 5.

Remark A.3.1. For each chi-squared approximation, $\max\{\varpi_1, \varpi_3\}$ already characterizes the bias well most of the time. We use $M_c(\varpi_1, \varpi_3)$ instead of $\max\{\varpi_1, \varpi_3\}$ because ϖ_3 can mistakenly indicate a large bias under small ϵ, especially when n is small. Compared to $\max\{\varpi_1, \varpi_3\}$, $M_c(\varpi_1, \varpi_3)$ does not use ϖ_3 when ϖ_1 indicates that the bias is still small. As long as c is sufficiently small but not too close to zero, $M_c(\varpi_1, \varpi_3)$ will not take the wrong value given by ϖ_3, and thus gives a good evaluation of the approximation bias under a wide range of ϵ values. Despite the difference between $M_c(\varpi_1, \varpi_3)$ and $\max\{\varpi_1, \varpi_3\}$, we note that $M_c(\varpi_1, \varpi_3)$ is equal to $\max\{\varpi_1, \varpi_3\}$ under most cases. For instance, in all our simulations with $n = 500$ and $c = 0.002$, $M_c(\varpi_1, \varpi_3) = \max\{\varpi_1, \varpi_3\}$. Thus in § 3, we did not highlight this difference. When the Bartlett correction is used, we know that similar analysis applies to $\max\{\varpi_2, \varpi_4\}$ and $M_c(\varpi_2, \varpi_4)$.

A.3.2. *Multiple-Sample Tests (IV)–(VI).* Consider $k = 3$, $n_1 = n_2 = n_3$, and $n = n_1 + n_2 + n_3$. Under the null hypothesis of each multiple-sample test (IV)–(VI), we set $\mu_i = (0, \ldots, 0)^{\mathrm{T}}$, and $\Sigma_i = \mathrm{I}_p$ for $i = 1, 2, 3$.

(1) On the phase transition boundaries. Let $p = \lfloor n^{\epsilon} \rfloor$, where $n = n_1 + n_2 + n_3$ and $n_i \in \{100, 500, 1000, 5000\}$ for $i = 1, 2, 3$. We then plot the empirical type-I error rates (over 1000 replications) versus ϵ for each chi-squared approximation in Fig. 3.

(2) On the asymptotic biases. To evaluate the asymptotic biases in Theorems A.2 and A.3, we take $p = \lfloor n^{\epsilon} \rfloor$, where $n = n_1 + n_2 + n_3$, $n_i \in \{100, 500\}$ for $i = 1, 2, 3$, and $\epsilon \in (0, 1)$. The results of $n_i = 100$ and 500 (over 3000 replications) are given in Fig. 6 and Fig. 7, respectively. Similarly to Fig. 4 and Fig.

5, in each row of Fig. 6 and Fig. 7, the lines with dot markers in the four columns (a)–(d) give ϖ_1, $M_c(\varpi_1, \varpi_3)$, ϖ_2, and $M_c(\varpi_2, \varpi_4)$, respectively.

A.3.3. *Testing Independence between Multiple Tests (VII).* Consider $k = 3$. Under the null hypothesis of test (VII), we set $\mu = (0, \ldots, 0)^{\mathrm{T}}$ and $\Sigma = \mathrm{I}_p$.

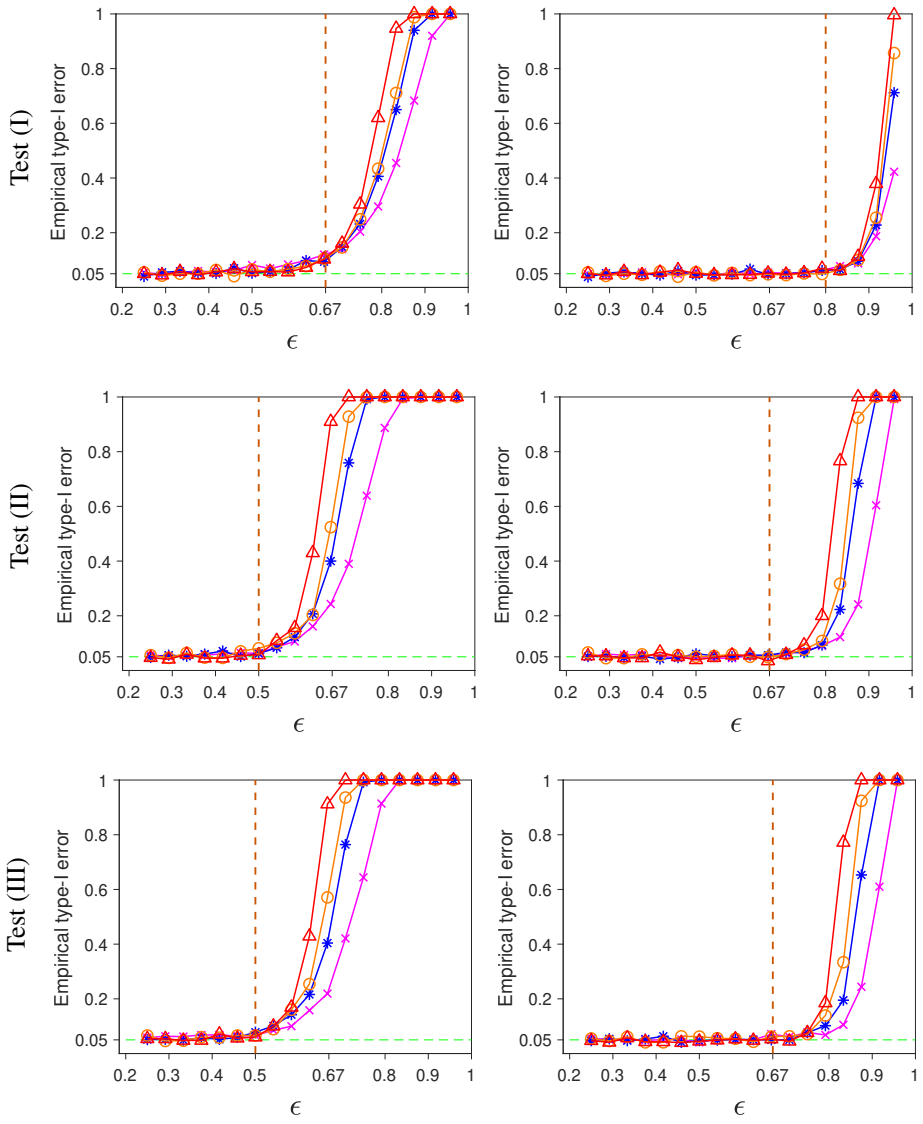
(1) On the phase transition boundaries. Let $p = \lfloor n^{\epsilon} \rfloor$, where $\epsilon \in \{6/24, 7/24, \ldots, 23/24\}$ and $n \in \{100, 500, 1000, 5000\}$. Under each (n, p), we set $p_1 = p_2 = \lfloor p/3 \rfloor$ and $p_3 = p - p_1 - p_2$, and then plot the empirical type I error (over 1000 replications) versus ϵ in Fig. 3.

(2) On the asymptotic biases. To evaluate the asymptotic biases in Theorems A.5 and A.6, we set $p = \lfloor n^{\epsilon} \rfloor$, where $n \in \{100, 500\}$ and $\epsilon \in (0, 1)$. Under each (n, p), we take $p_1 = p_2 = \lfloor p/3 \rfloor$ and $p_3 = p - p_1 - p_2$. The results of $n = 100$ and 500 (over 3000 replications) are given in Fig. 8 and Fig. 9, respectively. Similarly to Figures 4–7, in Fig. 8 and Fig. 9, the lines with dot markers in the four columns (a)–(d) give ϖ_1, $M_c(\varpi_1, \varpi_3)$, ϖ_2, and $M_c(\varpi_2, \varpi_4)$, respectively.

We next analyze the simulation results. First, as shown in Figures 2 and 3, the theoretical phase transition boundary, denoted by a vertical line, is observed to be consistent with where each chi-squared approximation starts to fail. For instance, the two plots in the first row of Fig. 2 show that for test (I), the type-I error rates of the chi-squared approximations without and with the Bartlett correction begin to inflate when ϵ is around $2/3$ and $4/5$, respectively. These are consistent with $d_1 = 2/3$ and $d_2 = 4/5$ for test (I) in Theorem 2.1. Similarly for other tests, we can see that the numerical results are also consistent with the corresponding conclusions in Theorems 2.1, A.1, and A.4.

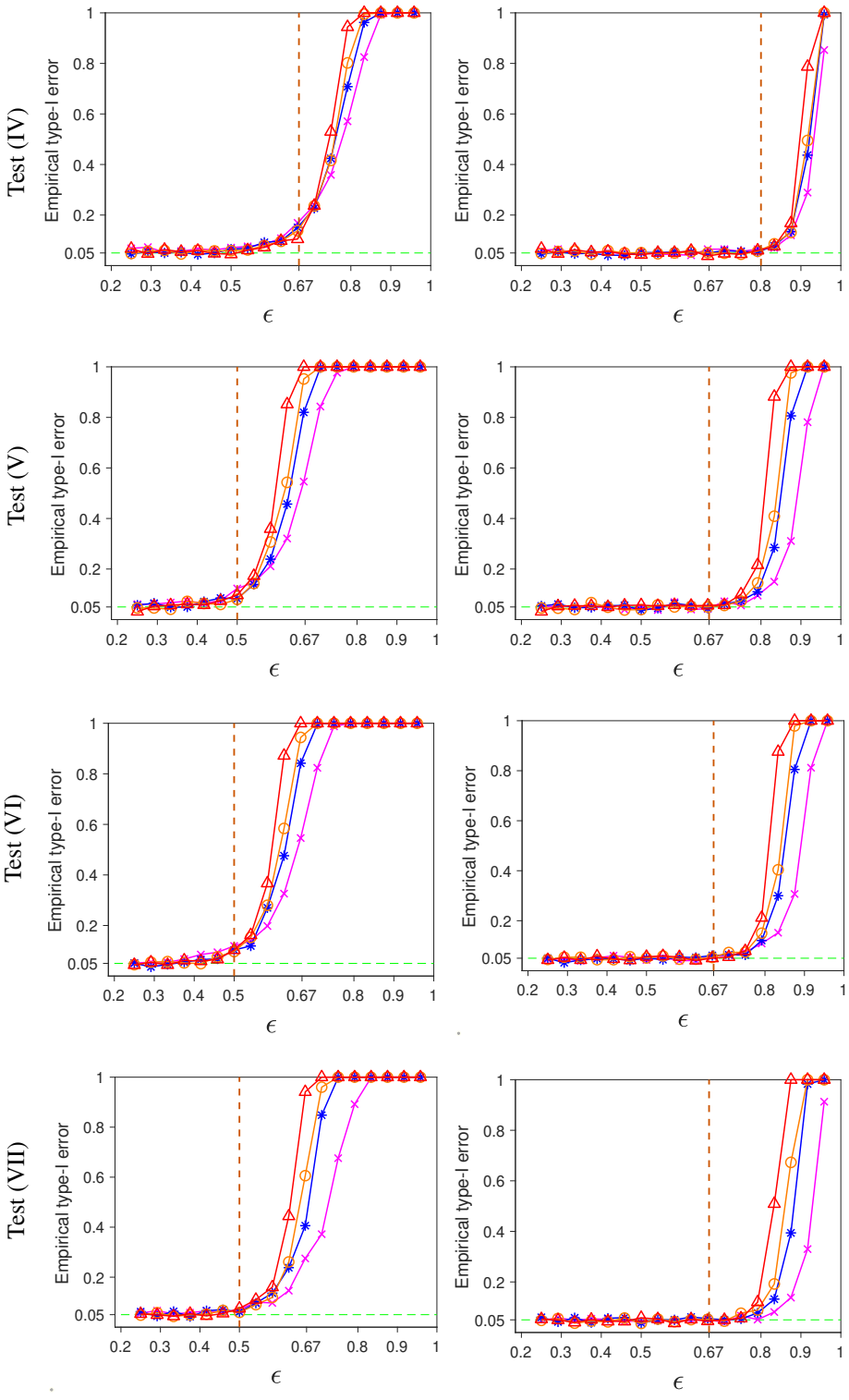
Second, similarly to § 3, the results in Figures 4–9 show that the derived theoretical asymptotic biases provide good evaluations of the corresponding chi-squared approximation biases. From the subfigures in the column (a) of Figures 4–9, we can see that as ϵ increases, the empirical type-I error inflates, and ϖ_1 also increases accordingly. At the ϵ values where the type-I error begins to inflate, the difference between the empirical type-I error and ϖ_1 is close to 0.05, as shown by the circle line, which suggests that ϖ_1 approximates the chi-squared approximation bias $\mathrm{pr}\{-2 \log \Lambda_n > \chi^2_f(\alpha)\} - \alpha$ well in this regime. When ϵ further increases beyond the corresponding phase transition boundary, the asymptotic bias ϖ_1 keeps increasing, and its large value indicates the failure of the chi-squared approximation, even though now ϖ_1 underestimates the approximation bias in this regime. To better characterize the approximation bias when ϵ is beyond the phase transition boundary, we combine ϖ_1 and ϖ_3 by plotting $M_c(\varpi_1, \varpi_3)$ in the column (b) of Figures 4–9. The results suggest that utilizing the two asymptotic biases in (1) and in (3) together can give a good evaluation of the approximation bias under a wide range of ϵ values, either below or above the phase transition boundary. Moreover, in each subfigure in the column (b), we also highlight the location with x-axis ϵ^* where $M_c(\varpi_1, \varpi_3)$ starts to be larger than ϖ_1 (the plus sign). When $\epsilon < \epsilon^*$, $M_c(\varpi_1, \varpi_3) = \varpi_1$, indicating that ϖ_1 approximates the bias better than ϖ_3 does in this regime, while ϖ_3 performs better than ϖ_1 when $\epsilon \geq \epsilon^*$. Similarly, for the chi-squared approximation with the Bartlett correction, similar conclusions can be obtained by the results in the columns (c) and (d) of Figures 4–9.

The simulations under the finite sample suggest that the derived asymptotic biases can be used as practical guidelines for the considered likelihood ratio tests. Specifically, when using the chi-squared approximation in each test, similarly to our recommendation in § 3, the practitioners can compare the asymptotic bias, either ϖ_1 or $M_c(\varpi_1, \varpi_3)$, with a small threshold value that they may specify in advance, e.g., 0.01–0.02. If the asymptotic bias is larger than the threshold, the chi-squared approximation should not be directly used, and other methods would be needed. In addition, when using the chi-squared approximation with the Bartlett correction in each test, we can compare the asymptotic bias, either ϖ_2 or $M_c(\varpi_2, \varpi_4)$ with the pre-specified threshold value. Similarly, if the asymptotic bias is larger than the threshold, the chi-squared approximation with the Bartlett correction should not be directly applied, and other methods would be needed.



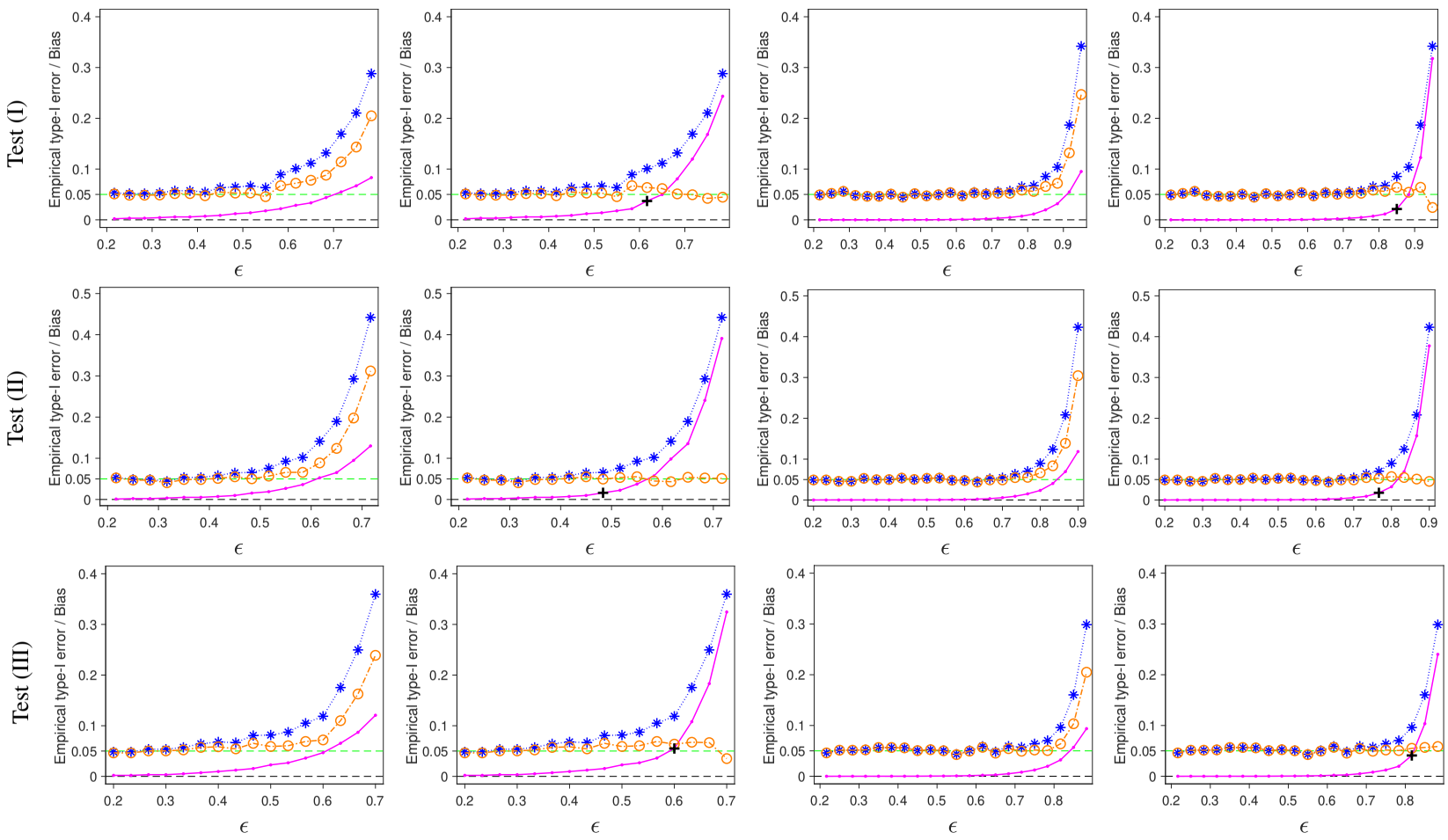
(i) Without the Bartlett correction (ii) With the Bartlett correction correction

Fig. 2: One-sample tests (I)–(III). Rows 1-3 give the results for tests (I)–(III), respectively. Columns (i) and (ii) correspond to the chi-squared approximations without and with the Bartlett correction, respectively. Within each subfigure: empirical type-I error versus ϵ with $n = 100$ (cross), 500 (asterisk), 1000 (square), and 5000 (triangle); theoretical phase transition boundary (vertical dashed line).



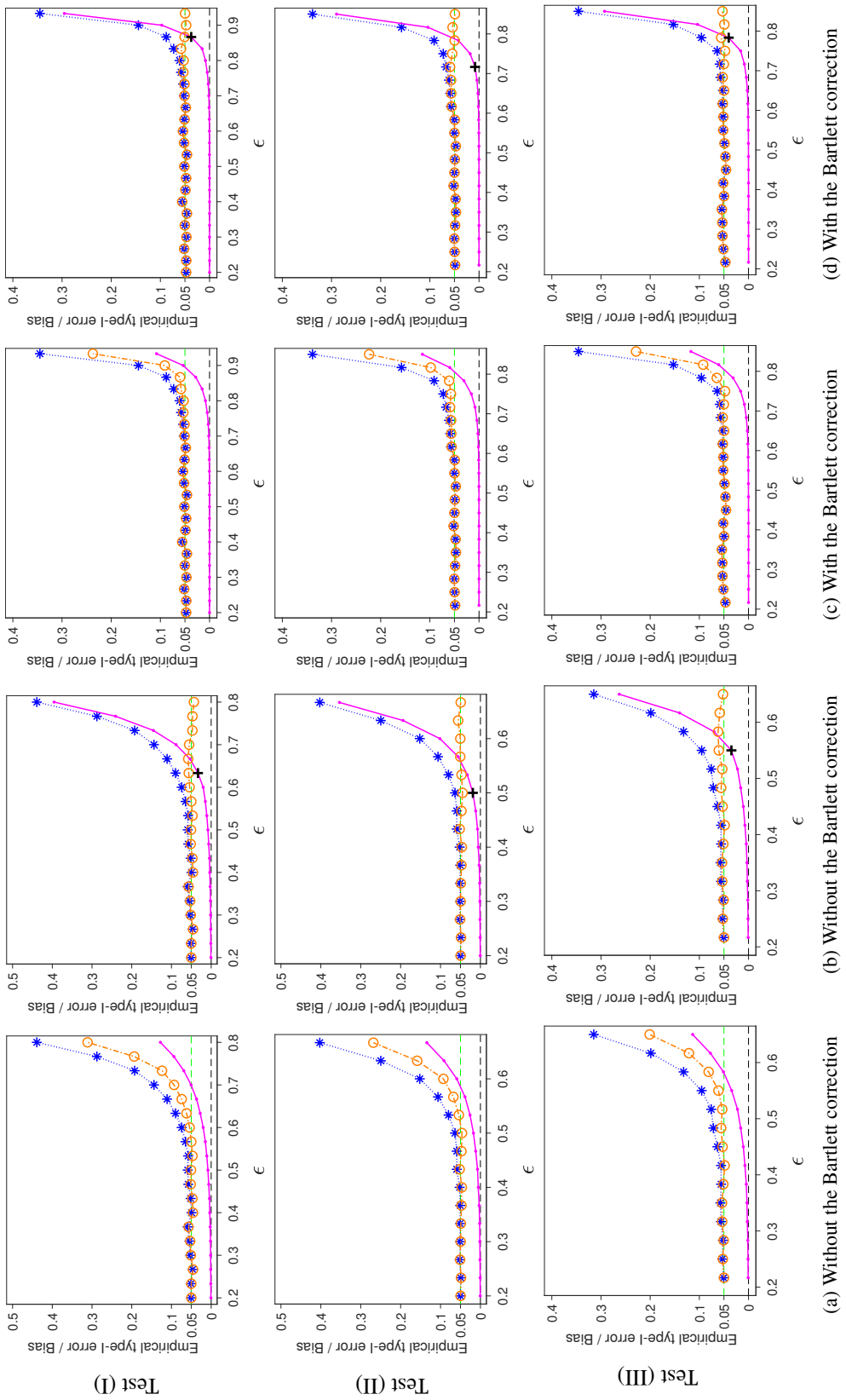
(i) Without the Bartlett correction (ii) With the Bartlett correction

Fig. 3: Multiple-sample tests (IV)–(VI) and the independence test (VII). Rows 1-4 give results for tests (IV)–(VII), respectively. Columns (i) and (ii) are for the chi-squared approximations without and with the Bartlett correction, respectively. Within each subfigure, please see the caption description in Fig. 2.



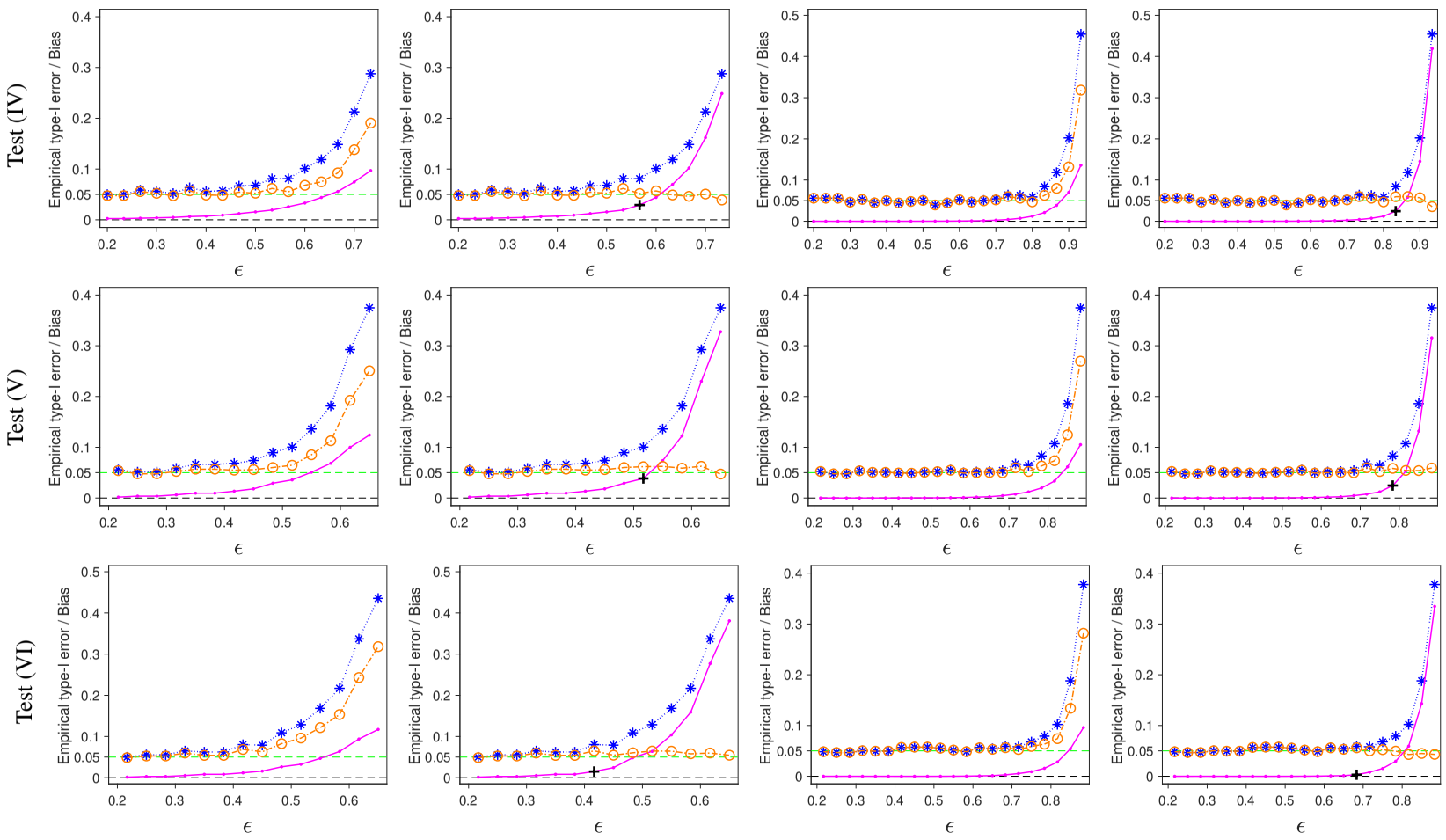
(a) Without the Bartlett correction (b) Without the Bartlett correction (c) With the Bartlett correction (d) With the Bartlett correction

Fig. 4: One-sample tests (I)–(III) when $n = 100$. Rows 1–3 present the results for tests (I)–(III), respectively. For four columns in each row: (a) Without the Bartlett correction: empirical type-I error versus ϵ (asterisk); ϖ_1, i.e., the asymptotic bias in (1) (dot); the difference between the empirical type-I error and ϖ_1 (circle). (b) Without the Bartlett correction: empirical type-I error versus ϵ (asterisk); $M_c(\varpi_1, \varpi_3)$ with $c = 0.002$ (dot); the location with x-axis ϵ^* satisfying $M_c(\varpi_1, \varpi_3) = \varpi_1$ when $\epsilon < \epsilon^*$ and $M_c(\varpi_1, \varpi_3) > \varpi_1$ when $\epsilon \geq \epsilon^*$ (plus sign); the difference between the empirical type-I error and $M_c(\varpi_1, \varpi_3)$ (circle). (c) With the Bartlett correction: empirical type-I error versus ϵ (asterisk); ϖ_2, i.e., the asymptotic bias in (2) (dot); the difference between the empirical type-I error and ϖ_2 (circle). (d) With the Bartlett correction: empirical type-I error versus ϵ (asterisk); $M_c(\varpi_2, \varpi_4)$ with $c = 0.002$ (dot); the location with x-axis ϵ^* satisfying $M_c(\varpi_2, \varpi_4) = \varpi_2$ when $\epsilon < \epsilon^*$ and $M_c(\varpi_2, \varpi_4) > \varpi_2$ when $\epsilon \geq \epsilon^*$ (plus sign); the difference between the empirical type-I error and $M_c(\varpi_2, \varpi_4)$ (circle).



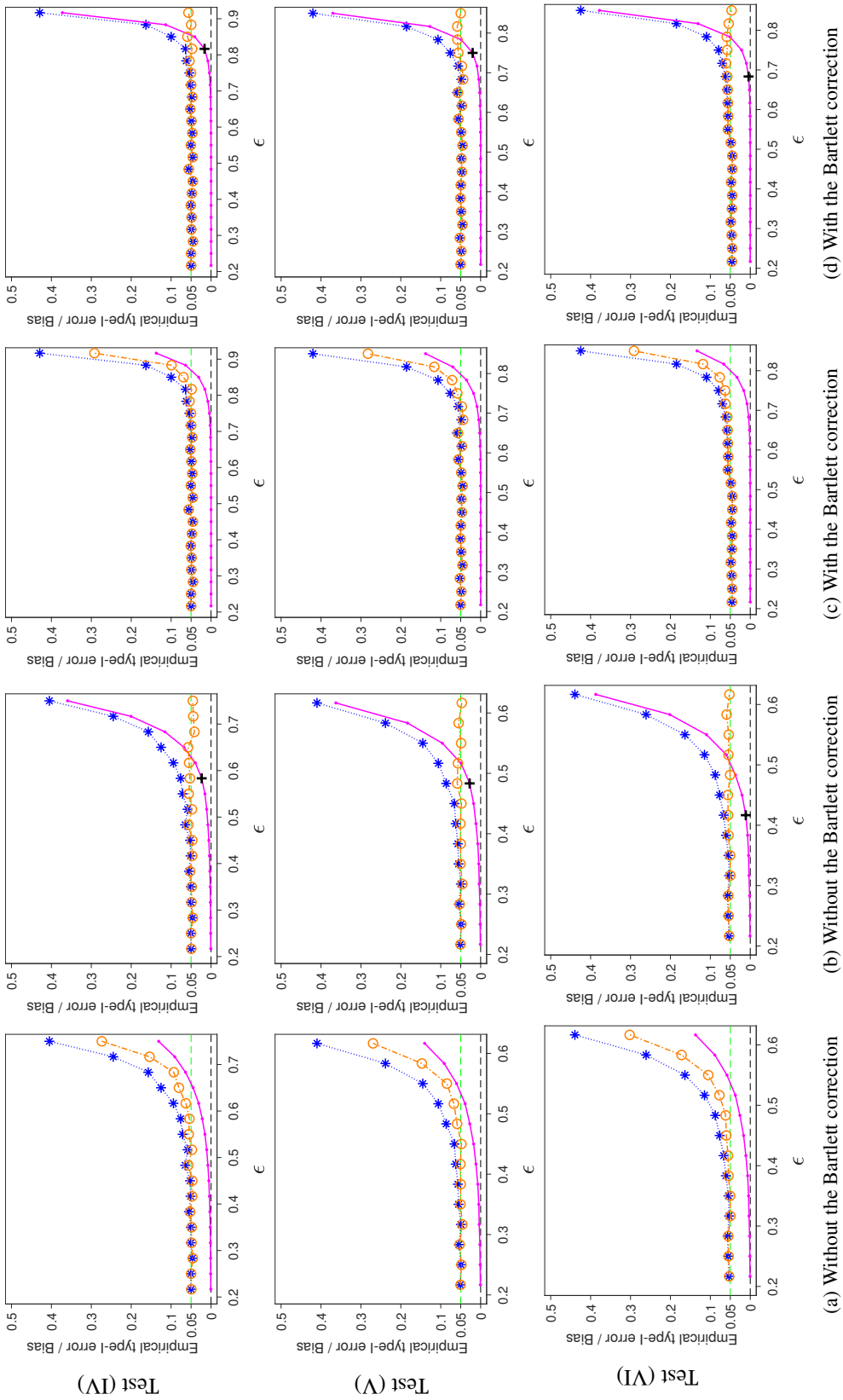
(a) Without the Bartlett correction (b) Without the Bartlett correction (c) With the Bartlett correction (d) With the Bartlett correction

Fig. 5: One-sample tests (I)–(III) when $n = 500$. Rows 1–3 present the results for tests (I)–(III), respectively. For four columns in each row, please see the caption description in Fig. 4.



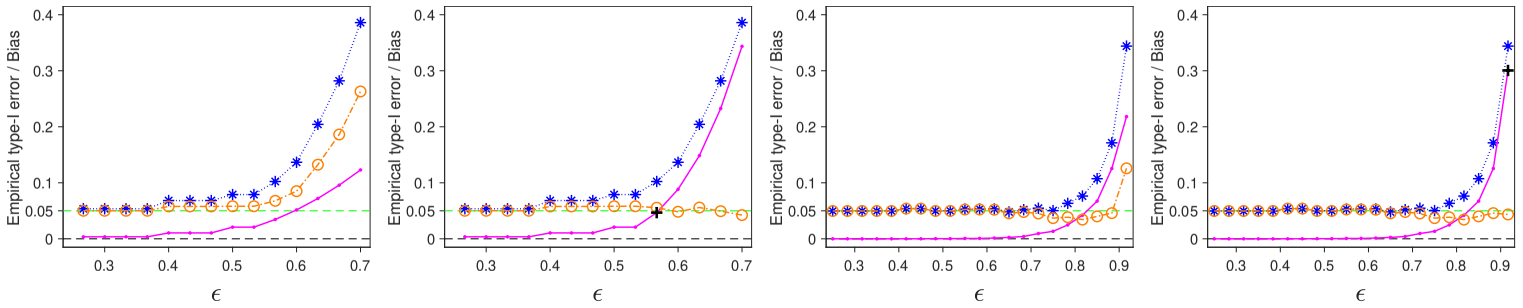
(a) Without the Bartlett correction (b) Without the Bartlett correction (c) With the Bartlett correction (d) With the Bartlett correction

Fig. 6: Multiple-sample tests (IV)–(VI) when $n = 100$. Rows 1–3 present the results for tests (IV)–(VI), respectively. For four columns in each row, please see the caption description in Fig. 4.



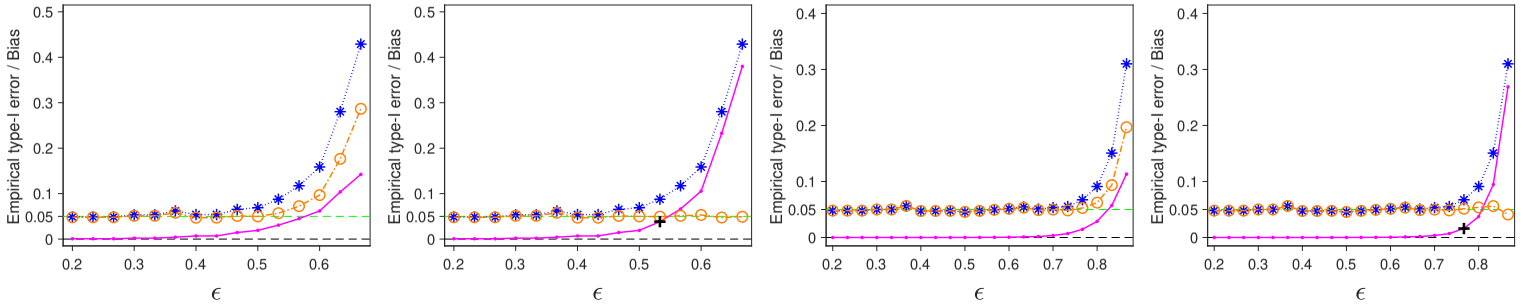
(a) Without the Bartlett correction (b) Without the Bartlett correction (c) With the Bartlett correction (d) With the Bartlett correction

Fig. 7: Multiple-sample tests (IV)–(VI) when $n = 500$. Rows 1–3 present the results for tests (IV)–(VI), respectively. For four columns in each row, please see the caption description in Fig. 4.



(a) Without the Bartlett correction (b) Without the Bartlett correction (c) With the Bartlett correction (d) With the Bartlett correction

Fig. 8: Independence test (VII) when $n = 100$: for columns (a)–(d), please see the caption description in Fig. 4.



(a) Without the Bartlett correction (b) Without the Bartlett correction (c) With the Bartlett correction (d) With the Bartlett correction

Fig. 9: Independence test (VII) when $n = 500$: for columns (a)–(d), please see the caption description in Fig. 4.

B. Proof Illustration with Problem (III)

In this section, we illustrate the proofs of Theorems 2.1–2.3 by focusing on the testing problem (III), which jointly tests the the one-sample mean vector and covariance matrix. Other testing problems (I)–(II) and (IV)–(VII) can be proved following a similar analysis, and are discussed in Section C. We define some notation to facilitate the proofs. For two sequences of numbers $\{a_n; n \geq 1\}$ and $\{b_n; n \geq 1\}$, $a_n = O(b_n)$ denotes $\limsup_{n\to\infty} |a_n/b_n| < \infty$; $a_n = o(b_n)$ denotes $\lim_{n\to\infty} a_n/b_n = 0$; $a_n = \Theta(b_n)$ represents that $a_n = O(b_n)$ and $b_n = O(a_n)$ hold simultaneously; $a_n \sim b_n$ denotes $\lim_{n\to\infty} |a_n/b_n| = 1$.

B.1. *Proof of Theorem* 2.1 *(III)*

When p is fixed, the chi-squared approximations hold by the classical multivariate analysis (Anderson, 2003; Muirhead, 2009). Therefore, without loss of generality, the proofs below focus on $p \to \infty$.

Deriving the necessary and sufficient conditions for the chi-squared approximations requires the correct understanding of the limiting behavior of $\log \Lambda_n$ under both low and high dimensions. Particularly, we examine the limiting distribution of the log likelihood ratio test statistic $\log \Lambda_n$ based on the moment generating function of $\log \Lambda_n$, that is, $\mathrm{E}\{\exp(t \log \Lambda_n)\}$. For Λ_n in question (III), by Theorem 8.5.3 and Corollary 8.5.4 in Muirhead (2009), we have that under H_0,

$$\mathrm{E}\{\exp(t \log \Lambda_n)\} = \mathrm{E}(\Lambda_n^t) = \left(\frac{2e}{n}\right)^{npt/2} (1+t)^{-np(1+t)/2} \times \frac{\Gamma_p[\{n(1+t)-1\}/2]}{\Gamma_p\{(n-1)/2\}}, \tag{B.1}$$

where $\Gamma_p(\cdot)$ is the multivariate Gamma function; see Definition 2.1.10 in Muirhead (2009).

When p is fixed, the moment generating function of $-2 \log \Lambda_n$ approximates that of a chi-squared variable χ_f^2, where $f = p(p+3)/2$; see, Sections 8.2.4 and 8.5 in Muirhead (2009). When $p \to \infty$, Jiang and Yang (2013) and Jiang and Qi (2015) derived an approximate expansion of the multivariate Gamma function, and their Theorem 5 utilized (B.1) to show that under the conditions of Theorem 2.1,

$$\mathrm{E}[\exp\{s(-2 \log \Lambda_n + 2\mu_n)/(2n\sigma_n)\}] \to \exp(s^2/2), \tag{B.2}$$

where $\exp(s^2/2)$ is the moment generating function of $\mathcal{N}(0,1)$, and

$$\mu_n = -\frac{1}{4}\left\{n(2n-2p-3)\log\left(1-\frac{p}{n-1}\right) + 2(n+1)p\right\}, \tag{B.3}$$

$$\sigma_n^2 = -\frac{1}{2}\left\{\frac{p}{n-1} + \log\left(1-\frac{p}{n-1}\right)\right\}. \tag{B.4}$$

We next prove (i) in Theorem 2.1 when $p \to \infty$ based on (B.2). Particularly, we write

$$\sup_{\alpha\in(0,1)} \left|\mathrm{pr}\{-2 \log \Lambda_n > \chi_f^2(\alpha)\} - \alpha\right| = \sup_{\alpha\in(0,1)} \left|\mathrm{pr}(T_n > q_{n,\alpha}) - \bar{\Phi}(q_{n,\alpha}) + \bar{\Phi}(q_{n,\alpha}) - \bar{\Phi}(z_\alpha)\right|, \tag{B.5}$$

where $T_n = (-2 \log \Lambda_n + 2\mu_n)/(2n\sigma_n)$, $q_{n,\alpha} = \{\chi_f^2(\alpha) + 2\mu_n\}/(2n\sigma_n)$, and $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$ with $\Phi(\cdot)$ being the cumulative distribution function of $\mathcal{N}(0,1)$. Since (B.2) suggests that T_n converges to $\mathcal{N}(0,1)$ in distribution, and the cumulative distribution function of $\mathcal{N}(0,1)$ is continuous, by Pólya-Cantelli Lemma (see, e.g., Lemma 2.11 in Van der Vaart (2000)), we have $\sup_{\alpha\in(0,1)} |\mathrm{pr}(T_n > q_{n,\alpha}) - \bar{\Phi}(q_{n,\alpha})| \to 0$. Consequently, (B.5) $\to 0$ if and only if $\sup_{\alpha\in(0,1)} |\bar{\Phi}(q_{n,\alpha}) - \bar{\Phi}(z_\alpha)| \to 0$, which is equivalent to $\sup_{\alpha\in(0,1)} |q_{n,\alpha} - z_\alpha| \to 0$, as $\bar{\Phi}(\cdot)$ is a continuous and strictly decreasing function with bounded derivative. Since χ_f^2 can be viewed as a summation over f independent χ_1^2 variables, and $f \to \infty$ as $p \to \infty$, we can apply BerryEsseen theorem to χ_f^2 variable, and obtain

$$\sup_{\alpha\in(0,1)} \left|\{\chi_f^2(\alpha) - f\}/\sqrt{2f} - z_\alpha\right| = O(f^{-1/2}). \tag{B.6}$$

Therefore, $\sup_{\alpha\in(0,1)}|q_{n,\alpha}-z_\alpha|\to 0$ is equivalent to

$$\sqrt{2f}\times(2n\sigma_n)^{-1}\to 1, \tag{B.7}$$

$$(O(1)+f+2\mu_n)\times(2n\sigma_n)^{-1}\to 0. \tag{B.8}$$

Following similar analysis, we know that under the conditions of Theorem 2.1 and $p\to\infty$, for the chi-squared approximation with the Bartlett correction, $\sup_{\alpha\in(0,1)}|\mathrm{pr}\{-2\rho\log\Lambda_n>\chi_f^2(\alpha)\}-\alpha|$ holds if and only if

$$\sqrt{2f}\times(2n\rho\sigma_n)^{-1}\to 1, \tag{B.9}$$

$$(O(1)+f+2\rho\mu_n)\times(2n\rho\sigma_n)^{-1}\to 0. \tag{B.10}$$

We next examine (B.7)–(B.8) and (B.9)–(B.10) for the chi-squared approximation without and with the Bartlett correction, respectively.

(III.i) The chi-squared approximation. We next discuss two cases $\lim_{n\to\infty}p/n=0$ and $\lim_{n\to\infty}p/n=C\in(0,1]$, respectively.

Case (III.i.1) $\lim_{n\to\infty}p/n=0$. Under this case, we prove that (B.7) holds. As $\sqrt{2f}\sim p$, it is equivalent to show that $p/(2n\sigma_n)\to 1$. By Taylor's expansion of σ_n^2 in (B.4), we have

$$2\sigma_n^2=-\frac{p}{n-1}-\log\left(1-\frac{p}{n-1}\right)=\frac{p^2}{2(n-1)^2}+o\left(\frac{p^2}{n^2}\right),$$

and therefore $\sqrt{2f}\times(2n\sigma_n)^{-1}\to 1$. We next show that (B.8) holds if and only if $p^2/n\to 0$. Given (B.7) and $\sqrt{2f}\sim p$, (B.8) is equivalent to $(f+2\mu_n)/p\to 0$. By $p/n=o(1)$ and Taylor's expansion of $\log(1-x)$, for μ_n in (B.3), we have

$$\begin{aligned}4\mu_n/p=&-2(n+1)+n(2n-2p-3)\left\{\frac{1}{n-1}+\frac{p}{2(n-1)^2}+\frac{p^2}{3(n-1)^3}+O\left(\frac{p^3}{n^4}\right)\right\}\\ =&-2(n+1)+(2n-2p-3)\left\{1+\frac{p}{2(n-1)}+\frac{p^2}{3(n-1)^2}\right\}+2+o(1)+O\left(\frac{p^3}{n^2}\right)\\ =&-2p-3+\frac{(2n-2p-3)p}{2(n-1)}+\frac{(2n-2p-3)p^2}{3(n-1)^2}+o(1)+O\left(\frac{p^3}{n^2}\right).\end{aligned} \tag{B.11}$$

As $2f/p=p+3$, we obtain

$$\begin{aligned}2\times(f+2\mu_n)/p=&-p+\frac{\{2(n-1)-2p-1\}p}{2(n-1)}+\frac{2p^2}{3(n-1)}+o(1)+O\left(\frac{p^3}{n^2}\right)\\ =&-\frac{p^2}{3(n-1)}+o(1)+O\left(\frac{p^3}{n^2}\right).\end{aligned} \tag{B.12}$$

Therefore when $p/n\to 0$, (B.8) holds if and only if $p^2/n\to 0$.

Case (III.i.2) $\lim_{n\to\infty}p/n=C\in(0,1]$. Under this case, we have

$$\sqrt{2f}\times(2n\sigma_n)^{-1}\sim p(2n\sigma_n)^{-1}\sim C(2\sigma_n)^{-1}. \tag{B.13}$$

If $C=1$, $\sigma_n^2\to\infty$ and thus (B.13) $\to 0$. If $C\in(0,1)$, we have $C(2\sigma_n)^{-1}\sim C[-2\{C+\log(1-C)\}]^{-1/2}<1$ when $0<C<1$. In summary, (B.7) does not hold, which suggests that the chi-squared approximation fails.

Finally, we consider a general sequence $p/n=p_n/n\in[0,1]$, where we write p as p_n to emphasize that p changes with n. Similarly, we also write f as f_n. Note that a sequence converges if and only if every subsequence converges. For the sequence $\{p_n/n\}$, by the BolzanoWeierstrass theorem, we can further take a subsequence $\{n_t\}$ such that $p_{n_t}/n_t\to C\in[0,1]$. If $C\in(0,1]$, the above analysis still applies, which shows that the chi-squared approximation fails. Alternatively, if all the subsequences of $\{p/n\}$

converge to 0, we know $p/n \to 0$. In summary, the above analysis shows that (B.7) and (B.8) hold if and only if $p^2/n \to 0$.

(III.ii) The chi-squared approximation with the Bartlett correction. Similarly to the analysis above, we discuss two cases $\lim_{n\to\infty} p/n = 0$ and $\lim_{n\to\infty} p/n = C \in (0, 1]$, respectively.

Case (III.ii.1) $\lim_{n\to\infty} p/n = 0$. Under this case, we know (B.9) holds since $\rho = 1 + O(p/n) \to 1$ and $p/(2n\sigma_n) \to 1$ as shown in Case (III.i.1) above. Given (B.9), deriving the condition for (B.10) is equivalent to examine when $p^{-1}(f + 2\rho\mu_n) \to 0$. Following the analysis of (B.12), we further obtain

$$2 \times (f + 2\mu_n)/p = (p + 3) - 2(n + 1) + n(2n - 2p - 3) \sum_{j=1}^{4} \frac{p^{j-1}}{j(n-1)^j} + O\left(\frac{p^4}{n^3}\right) \tag{B.14}$$

$$= -\frac{p^2}{3(n-1)} - \frac{p^3}{6(n-1)^2} + O\left(\frac{p^4}{n^3}\right) + o(1).$$

We write $\rho = 1 - \Delta_n$ where $\Delta_n = \{6n(p + 3)\}^{-1}(2p^2 + 9p + 11)$, which is $O(p/n)$. By (B.12), we have $4\mu_n/p = -p - 3 - p^2/\{3(n - 1)\} + o(1) + O(p^3 n^{-2})$. Together with (B.14), we have

$$2 \times (f + 2\rho\mu_n)/p = 2 \times (f + 2\mu_n)/p - 4\Delta_n \times \mu_n/p \tag{B.15}$$

$$= -\frac{p^2}{3(n-1)} - \frac{p^3}{6(n-1)^2} - \Delta_n \left\{-p - 3 - \frac{p^2}{3(n-1)}\right\} + O\left(\frac{p^4}{n^3}\right) + o(1)$$

$$= -\frac{p^2}{3(n-1)} - \frac{p^3}{6(n-1)^2} + \frac{2p^2(p+3)}{6n(p+3)} + \frac{2p^2 \times p^2}{6n(p+3) \times 3(n-1)} + O\left(\frac{p^4}{n^3}\right) + o(1)$$

$$= -\frac{p^3}{18n^2} + O\left(\frac{p^4}{n^3}\right) + o(1).$$

Therefore under this case (B.10) holds if and only if $p^3/n^2 \to 0$.

Case (III.ii.2): When $\lim_{n\to\infty} p/n = C \in (0, 1]$, we have $\rho \to 1 - C/3$ and

$$\sqrt{2f} \times (2n\rho\sigma_n)^{-1} \sim C \times (1 - C/3)^{-1}(2\sigma_n)^{-1}.$$

Similarly to the Case (III.i.2) above, if $C = 1$, (B.9) $\to 0$; if $C \in (0, 1)$, we have $C(1 - C/3)^{-1}(2\sigma_n)^{-1} \sim C(1 - C/3)^{-1}[-2\{C + \log(1 - C)\}]^{-1/2} < 1$ when $0 < C < 1$. In summary, (B.9) does not hold, which suggests the failure of the chi-squared approximation with the Bartlett correction.

For a general sequence $p/n = p_n/n \in [0, 1]$, the analysis of taking subsequences above can be applied similarly. In summary, we know that for the likelihood ratio test in problem (III), the chi-squared approximation with the Bartlett correction holds if and only if $p^3/n^2 \to 0$.

B.2. *Proof of Theorem* 2.2 *(III)*

Similarly to § B.1, in this subsection, we prove Theorem 2.2 for problem (III) as an illustration example, while the proofs of other problems are similar and the details are provided in § C.3. Particularly, we prove Theorem 2.2 for problem (III) by examining the characteristic function of $-2\eta \log \Lambda_n$, where $\eta = 1$ or $\eta = \rho$, and ρ is the corresponding Bartlett correction factor, given in § 2. The following Lemma B.2.1 gives an asymptotic expansion for the characteristic function $\mathrm{E}\{\exp(-2it\eta \log \Lambda_n)\}$, where the notation i is reserved to denote the solution of the equation $x^2 = -1$, i.e., the imaginary unit.

LEMMA B.2.1. *Under H_0 of the testing problem (III), when $\eta = 1$ or $\eta = \rho$ with the Bartlett correction factor ρ in § 2, the characteristic function of $-2\eta \log \Lambda_n$ satisfies that for a given integer L, when $p^{L+2}/n^L \to 0$,*

$$\mathrm{E}\{\exp(-2it\eta \log \Lambda_n)\} = (1 - 2it)^{-f/2} \exp\left[\sum_{l=1}^{L-1} \varsigma_l \{(1 - 2it)^{-l} - 1\} + O\left(\frac{p^{L+2}}{n^L}\right)\right],$$

where $f = p(p+3)/2$ is the corresponding degrees of freedom, and

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \sum_{j=1}^{p} \left\{ B_{l+1}\left(\frac{(1-\eta)n}{2} - \frac{j}{2}\right) - \left(\frac{(1-\eta)n}{2}\right)^{l+1} \right\} \left(\frac{\eta n}{2}\right)^{-l}. \tag{B.16}$$

For any integer $l \geq 1$, $B_l(\cdot)$ represents the Bernoulli polynomial of degree l; see, e.g., Eq. (25) in Section 8.2.4 of Muirhead (2009).

Proof. Section D.2.1 on Page 55. □

With Lemma B.2.1, we next prove (1) and (2) in Theorem 2.2 for the chi-squared approximations without and with the Bartlett correction, respectively.

(i) The chi-squared approximation. When $\rho = 1$, as $B_{l+1}(\cdot)$ is a polynomial of order $l+1$, we have $\varsigma_l = O(p^{l+2}n^{-l})$ for $l \geq 2$, and we can check that $\varsigma_1 = \Theta(p^3 n^{-1})$; see (B.23). Thus when $p^2/n \to 0$, $\varsigma_l \to 0$ for $l \geq 2$. Let $\Psi(t) = \mathrm{E}\{\exp(-2it \log \Lambda_n)\}$. Then by Lemma B.2.1,

$$\Psi(t) = (1-2it)^{-f/2} \left\{ \exp\left[\sum_{l=1}^{2} \varsigma_l \big\{ (1-2it)^{-l} - 1 \big\} + O\big(p^5 n^{-3}\big) \right] \right\}. \tag{B.17}$$

By Taylor' expansion, we can write $\exp[\varsigma_l\{(1-2it)^{-l}-1\}] = 1 + V_l(t)$, where

$$V_l(t) = \sum_{v=1}^{\infty} \frac{\varsigma_l^v}{v!} \sum_{w=0}^{v} \binom{v}{w} (1-2it)^{-lw} (-1)^{v-w}. \tag{B.18}$$

Then by (B.17) and $p^2/n \to 0$, we have $\Psi(t) = \tilde{\Psi}(t)\{1 + O(p^5/n^3)\}$, where

$$\begin{aligned}
\tilde{\Psi}(t) &= (1-2it)^{-f/2} \big\{ 1 + V_1(t) \big\} \big\{ 1 + V_2(t) \big\} \\
&= (1-2it)^{-f/2} + \sum_{v=1}^{\infty} \frac{\varsigma_1^v}{v!} \sum_{w=0}^{v} \binom{v}{w} (1-2it)^{-f/2-w} (-1)^{v-w} \\
&\quad + \sum_{v=1}^{\infty} \frac{\varsigma_2^v}{v!} \sum_{w=0}^{v} \binom{v}{w} (1-2it)^{-f/2-2w} (-1)^{v-w} \\
&\quad + \sum_{\substack{v_1 \geq 1;\ 0 \leq w_1 \leq v_1 \\ v_2 \geq 1;\ 0 \leq w_2 \leq v_2}} \frac{\varsigma_1^{v_1} \varsigma_2^{v_2}}{v_1! v_2!} \binom{v_1}{w_1} \binom{v_2}{w_2} (1-2it)^{-f-w_1-2w_2} (-1)^{v_1-w_1+v_2-w_2}.
\end{aligned} \tag{B.19}$$

Note that $(1-2it)^{-f/2}$ is the characteristic function of χ_f^2 distribution. Following similar analysis to Section 8.5 in Anderson (2003), we use the inversion property of the characteristic function, and then by (B.19), we obtain that

$$\begin{aligned}
&\Pr(-2\log\Lambda_n \leq x) \\
&= \left\{ \Pr(\chi_f^2 \leq x) + \sum_{v=1}^{\infty} \frac{\varsigma_1^v}{v!} \sum_{w=0}^{v} \binom{v}{w} \Pr(\chi_{f+2w}^2 \leq x)(-1)^{v-w} \right. \\
&\quad + \sum_{v=1}^{\infty} \frac{\varsigma_2^v}{v!} \sum_{w=0}^{v} \binom{v}{w} \Pr(\chi_{f+4w}^2 \leq x)(-1)^{v-w} \\
&\quad \left. + \sum_{\substack{v_1 \geq 1;\ 0 \leq w_1 \leq v_1 \\ v_2 \geq 1;\ 0 \leq w_2 \leq v_2}} \frac{\varsigma_1^{v_1} \varsigma_2^{v_2}}{v_1! v_2!} \binom{v_1}{w_1} \binom{v_2}{w_2} \Pr(\chi_{2f+2w_1+4w_2}^2 \leq x)(-1)^{v_1-w_1+v_2-w_2} \right\} \left\{ 1 + O\left(\frac{p^5}{n^3}\right) \right\}.
\end{aligned} \tag{B.20}$$

(From (B.19) to (B.20), Fubini's theorem is implicitly used to exchange the order of the infinite sum and the integration of characteristic functions.) We next utilize the following Propositions B.1 and B.2 to evaluate (B.20).

PROPOSITION B.1. *Given an integer $h \in \{1,2,3,4\}$, when $x = \chi^2_f(\alpha)$, there exists a constant C such that as $f \to \infty$,*

$$\sum_{w=0}^{v} \binom{v}{w} \Pr(\chi^2_{f+2hw} \le x)(-1)^{v-w} = O(v! C^v f^{-v/2}) \tag{B.21}$$

uniformly over $v \ge 1$.

Proof. Please see Section D.2.4 on Page 58. □

PROPOSITION B.2. *For $(h_1,h_2) = (1,2)$ or $(h_1,h_2) = (2,3)$, when $x = \chi^2_f(\alpha)$, there exists a constant C such that as $f \to \infty$,*

$$\begin{aligned}&\sum_{w_1=0}^{v_1}\sum_{w_2=0}^{v_2} \binom{v_1}{w_1}\binom{v_2}{w_2} \Pr(\chi^2_{2f+2h_1w_1+2h_2w_2} \le x)(-1)^{v_1-w_1+v_2-w_2} \\ &= O\{v_1!v_2!C^{v_1+v_2} f^{-(v_1+v_2)/2}\}\end{aligned}$$

uniformly over $v_1, v_2 \ge 1$.

Proof. Please see Section D.2.5 on Page 63. □

*Remark B.*2.1. In Propositions B.1 and B.2, C denotes a universal constant and its value can change. This is similarly used in the following proofs. In addition, for a series $\{b_{v,f}\}$ that depends on positive integers v and f, we say $b_{v,f} = O(v!C^v f^{-v/2})$ as $f \to \infty$ and uniformly over $v \ge 1$, if there exists a constant C such that $\sup_{v\ge1} \limsup_{f\to\infty} |b_{v,f}/(v!C^v f^{-v/2})| < \infty$.

When $x = \chi^2_f(\alpha)$ and $f \to \infty$, we apply Proposition B.1 with $h = 1$ and $h = 2$, and Proposition B.2 with $(h_1,h_2) = (1,2)$ to (B.20). Then as $\varsigma_1 = \Theta(p^3 n^{-1})$, $\varsigma_2 = O(p^4 n^{-2})$, and $f = \Theta(p^2)$, when $p \to \infty$ and $p^2/n \to 0$, we obtain

$$\Pr(-2\log\Lambda_n \le x) = \Pr(\chi^2_f \le x) + \varsigma_1 \left\{\Pr(\chi^2_{f+2} \le x) - \Pr(\chi^2_f \le x)\right\} + o(p^2/n). \tag{B.22}$$

We next compute ς_1. Particularly, for the chi-squared approximation, $\rho = 1$, and then by (B.16),

$$\varsigma_1 = \frac{1}{2}\sum_{j=1}^{p} B_2\Big(-\frac{j}{2}\Big)\Big(\frac{n}{2}\Big)^{-1} = \frac{1}{24n} p\left(2p^2 + 9p + 11\right), \tag{B.23}$$

where we use $B_2(z) = z^2 - z + 1/6$; see, e.g., Eq. (26) in Section 8.2.4 of Muirhead (2009). To finish the proof of (1), we use the following lemma.

LEMMA B.2.2. *When $x = \chi^2_f(\alpha)$ and $f \to \infty$, for $h \in \{1,2,3,4\}$,*

$$\begin{aligned}\Pr(\chi^2_{f+2h} \le x) - \Pr(\chi^2_f \le x) =& -\sum_{k=1}^{h} \left\{\Gamma\left(\frac{f}{2} + h - k + 1\right)\right\}^{-1} \left(\frac{x}{2}\right)^{\frac{f}{2}+h-k} e^{-x/2} &\text{(B.24)}\\ =& -\frac{h}{\sqrt{f\pi}} \exp\left(-\frac{z_\alpha^2}{2}\right)\left\{1 + O(f^{-1/2})\right\}. &\text{(B.25)}\end{aligned}$$

Proof. Please see Section D.2.3 on Page 57. □

As $p \to \infty$, $f \to \infty$. Then by (B.22) and (B.23), and applying Lemma B.2.2 with $h = 1$, (1) is proved, where $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) The chi-squared approximation with the Bartlett correction. Similarly to the proof in Part (i) above, we prove (2) by examining the expansion of the characteristic function in Lemma B.2.1. In particular, for the chi-squared approximation with the Bartlett correction, we note that the Bartlett correction factor ρ is chosen such that $\varsigma_1 = 0$ (see Section 8.5.3 in Muirhead (2009)). This can be checked by plugging $\rho = 1 -$

$\{6n(p+3)\}^{-1}(2p^2+9p+11)$ into (B.16) to calculate ς_1. In addition, by $B_3(z)=z^3-3z^2/2+z/2$ (see, e.g., Eq. (26) in Section 8.2.4 of Muirhead (2009)), we calculate

$$\varsigma_2=\frac{p(2p^4+18p^3+49p^2+36p-13)}{288(p+3)(\rho n)^2}, \tag{B.26}$$

and therefore $\varsigma_2=\Theta(p^4n^{-2})$. We redefine $\Psi(t)=\mathrm{E}\{\exp(-2it\rho\log\Lambda_n)\}$. Then when $p^3/n^2\to 0$, by Lemma B.2.1, we have

$$\Psi(t)=(1-2it)^{-f/2}\left\{\exp\left[\sum_{l=2}^{3}\varsigma_l\big\{(1-2it)^{-l}-1\big\}+O(p^6n^{-4})\right]\right\}, \tag{B.27}$$

where we use $\varsigma_1=0$. Similarly to (B.19), we have $\Psi(t)=(1-2it)^{-f/2}\{1+V_2(t)\}\{1+V_3(t)\}\{1+O(p^6n^{-4})\}$. Moreover, similarly to (B.20), we obtain

$$\begin{aligned}
&\mathrm{pr}(-2\rho\log\Lambda_n\le x) \qquad\qquad (\text{B.28})\\
&=\left\{\Pr(\chi_f^2\le x)+\sum_{v=1}^{\infty}\frac{\varsigma_2^v}{v!}\sum_{w=0}^{v}\binom{v}{w}\mathrm{pr}(\chi^2_{f+4w}\le x)(-1)^{v-w}\right.\\
&\quad+\sum_{v=1}^{\infty}\frac{\varsigma_3^v}{v!}\sum_{w=0}^{v}\binom{v}{w}\Pr(\chi^2_{f+6w}\le x)(-1)^{v-w}\\
&\quad\left.+\sum_{\substack{v_2\ge1;\ 0\le w_2\le v_2\\ v_3\ge1;\ 0\le w_3\le v_3}}\frac{\varsigma_2^{v_2}\varsigma_3^{v_3}}{v_2!v_3!}\binom{v_2}{w_2}\binom{v_3}{w_3}\Pr(\chi^2_{2f+4w_2+6w_3}\le x)(-1)^{v_2-w_2+v_3-w_3}\right\}\left\{1+O\Big(\frac{p^6}{n^4}\Big)\right\}.
\end{aligned}$$

When $x=\chi_f^2(\alpha)$ and $f\to\infty$, we apply Proposition B.1 with $h=2$ and $h=3$, and Proposition B.2 with $(h_1,h_2)=(2,3)$ to (B.28). Then as $\varsigma_2=\Theta(p^4/n^2)$, $\varsigma_3=O(p^5/n^3)$, and $f=\Theta(p^2)$, we know that when $p\to\infty$ and $p^3/n^2\to 0$,

$$\Pr(-2\rho\log\Lambda_n\le x)=\Pr(\chi_f^2\le x)+\varsigma_2\left\{\Pr(\chi^2_{f+4}\le x)-\Pr(\chi_f^2\le x)\right\}+o(p^3/n^2). \tag{B.29}$$

By (B.26) and (B.29), and applying Lemma B.2.2 with $h=2$, we prove (2), where $\vartheta_2(n,p)=2\varsigma_2/\sqrt{f}$.

B.3. *Proof of Theorem* 2.3 *(III)*

In this section, we prove Theorem 2.3 also by examining the characteristic function of the likelihood ratio test statistic. In particular, motivated by the limit in (B.2), we study the standardized test statistic $(-2\log\Lambda_n+2\mu_n)(2n\sigma_n)^{-1}$, where the values of μ_n and σ_n are given in Theorem 2.3. Under H_0 of the testing problem (III), by (B.1), the characteristic function of $(-2\log\Lambda_n+2\mu_n)/(2n\sigma_n)$ is

$$\begin{aligned}
&\mathrm{E}\left\{\exp\left(is\times\frac{-2\log\Lambda_n+2\mu_n}{2n\sigma_n}\right)\right\} \qquad\qquad (\text{B.30})\\
&=\left(\frac{2e}{n}\right)^{-npti/2}(1-ti)^{-np(1-ti)/2}\frac{\Gamma_p[\{n(1-ti)-1\}/2]}{\Gamma_p\{(n-1)/2\}}\exp\left(\frac{\mu_n si}{n\sigma_n}\right),
\end{aligned}$$

where i denotes the imaginary unit and $t=s/(n\sigma_n)$. Then the proof of Theorem 2.3 utilizes the following inequality result of the characteristic function.

LEMMA B.3.1 (THEOREM 1.4.9 (USHAKOV, 2011)). *Let $G_1(x)$ and $G_0(x)$ be two distribution functions with characteristic functions $\psi_1(s)$ and $\psi_0(s)$, respectively. If $G_0(x)$ has a derivative and $\sup_x G_0'(x)\le a<\infty$, then for any positive T and any $b\ge 1/(2\pi)$,*

$$\sup_x\big|G_1(x)-G_0(x)\big|\le b\int_{-T}^{T}\left|\frac{\psi_1(s)-\psi_0(s)}{s}\right|ds+\frac{c}{T},$$

where c is a constant that depends on a and b.

We next prove (3) and (4) in Theorem 2.3 for the chi-squared approximations without and with the Bartlett correction, respectively.

(i) Chi-squared approximation. We prove (3) by using Lemma B.3.1 to derive an upper bound of the difference $G_1(x) - G_0(x)$, where we consider

$$G_1(x) = \Pr\left(\frac{-2\log\Lambda_n + 2\mu_n}{2n\sigma_n} \le x\right), \qquad G_0(x) = \Phi(x);$$

here $\Phi(x)$ denotes the cumulative distribution function of the standard normal distribution. Then the characteristic function of $G_1(x)$ is $\psi_1(s) =$ (B.30), and the characteristic function of $G_0(x)$ is $\psi_0(s) = \exp(-s^2/2)$. To quantify $\psi_1(s) - \psi_0(s)$, we use the following Lemma B.3.2.

LEMMA B.3.2. *When* $s = o(\min\{(n/p)^{1/2}, f^{1/6}\})$,

$$\log\psi_1(s) - \log\psi_0(s) = O\left(\frac{p}{n}\right)s + \left(\frac{1}{p} + \frac{p}{n}\right)O\left(s^2\right) + O\left(\frac{s^3}{\sqrt{f}}\right). \tag{B.31}$$

Proof. Please see Section D.3.1 on Page 74. □

By Lemmas B.3.1 and B.3.2, we take $T = \min\{(n/p)^{(1-\delta)/2}, f^{(1-\delta)/6}\}$, where $\delta \in (0,1)$ is a small constant, and then

$$\sup_x \left|G_1(x) - G_0(x)\right| \le b\int_{-T}^{T}\psi_0(s)\left\{O\left(\frac{p}{n}\right) + \left(\frac{1}{p} + \frac{p}{n}\right)O\left(s\right) + O\left(\frac{s^2}{\sqrt{f}}\right)\right\}ds + \frac{c}{T}. \tag{B.32}$$

Since $\int_{-T}^{T}\psi_0(s) < \infty$, $\int_{-T}^{T}\psi_0(s)s < \infty$, and $\int_{-T}^{T}\psi_0(s)s^2 < \infty$, by $f = \Theta(p^2)$ and (B.32),

$$\sup_x \left|G_1(x) - G_0(x)\right| = O\left\{\left(\frac{p}{n}\right)^{(1-\delta)/2} + f^{-(1-\delta)/6}\right\}.$$

Consider $x = \{\chi_f^2(\alpha) + 2\mu_n\}(2n\sigma_n)^{-1}$, and then $G_1(x) - G_0(x)$ gives

$$\Pr\left\{-2\log\Lambda_n \le \chi_f^2(\alpha)\right\} - \Phi\left\{\frac{\chi_f^2(\alpha) + 2\mu_n}{2n\sigma_n}\right\} = O\left\{\left(\frac{p}{n}\right)^{(1-\delta)/2} + f^{-(1-\delta)/6}\right\}. \tag{B.33}$$

Then (3) is proved by $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$ and $\Pr\{-2\log\Lambda_n > \chi_f^2(\alpha)\} = 1 - \Pr\{-2\log\Lambda_n \le \chi_f^2(\alpha)\}$.

(ii) Chi-squared approximation with the Bartlett correction. To prove (4), we still use (B.32). Now consider $x = \{\chi_f^2(\alpha) + 2\rho\mu_n\}(2\rho n\sigma_n)^{-1}$, and then $G_1(x) - G_0(x)$ gives

$$\Pr\left\{-2\rho\log\Lambda_n \le \chi_f^2(\alpha)\right\} - \Phi\left\{\frac{\chi_f^2(\alpha) + 2\rho\mu_n}{2\rho n\sigma_n}\right\} = O\left\{\left(\frac{p}{n}\right)^{(1-\delta)/2} + f^{-(1-\delta)/6}\right\}.$$

*Remark B.*3.1. Although Theorem 2.3 is inspired by the limit in (B.2), which was first established in Jiang and Yang (2013), Theorem 2.3 differs from the existing results by further characterizing the convergence rate of (B.2) by Lemma B.3.2. Particularly, Jiang and Yang (2013) proved (B.2) when s is considered fixed and the convergence rate is not examined. On the other hand, Lemma B.3.2 allows s changes with n and p, and the difference between the two characteristic functions is characterized by (B.31). Technically, establishing (B.31) requires a careful investigation of the asymptotic expansion of the gamma functions, where the technical details are given in Sections D.1 and D.3.

*Remark B.*3.2. Since χ_f^2 can be viewed as a summation over f independent χ_1^2 variables, by applying the central limit theorem, we have $\chi_f^2(\alpha) = \sqrt{2f}z_\alpha + f + O(1)$, where z_α denote the upper α-level quantile of the standard normal distribution. For the problem (III), note that μ_n and σ_n in Theorem

2.3 are the same as (B.3) and (B.4), respectively. Then by the proof of (B.7) in Section B.1, we have $2n\sigma_n/\sqrt{2f} = 1 + O(p/n)$. Consequently, when $f \to \infty$ and $p/n \to 0$,

$$\Phi\left\{\frac{\chi_f^2(\alpha) + 2\mu_n}{2n\sigma_n}\right\} = \Phi\left(z_\alpha + \frac{f + 2\mu_n}{2n\sigma_n}\right) + O\left(\frac{1}{\sqrt{f}}\right) + O\left(\frac{p}{n}\right).$$

Moreover, by (B.12), $(f + 2\mu_n)/(2n\sigma_n) \sim -p^2/(6n)$ when $p/n \to 0$. Thus $-(f + 2\mu_n)/(2n\sigma_n) = \sqrt{2}\vartheta_1(n,p) + o(p^{1/d_1}n^{-1})$, which is of the order of $p^{1/d_1}n^{-1}$ with $d_1 = 1/2$. When $p/n^{d_1} \to 0$, by $\alpha = \bar{\Phi}(z_\alpha)$ and Taylor's series of $\bar{\Phi}(\cdot)$ at z_α,

$$\bar{\Phi}\left(z_\alpha + \frac{f + 2\mu_n}{2n\sigma_n}\right) - \alpha = \frac{\vartheta_1(n,p)}{\sqrt{\pi}} \exp\left(-\frac{z_\alpha^2}{2}\right) + o\left(\frac{p^{1/d_1}}{n}\right),$$

which suggests that the first two terms in the right hand side of (3) are consistent with (1). Similarly, for the chi-squared approximation with the Bartlett correction, when $f \to \infty$ and $p/n \to 0$,

$$\Phi\left\{\frac{\chi_f^2(\alpha) + 2\rho\mu_n}{2\rho n\sigma_n}\right\} = \Phi\left(z_\alpha + \frac{f + 2\rho\mu_n}{2\rho n\sigma_n}\right) + O\left(\frac{1}{\sqrt{f}}\right) + O\left(\frac{p}{n}\right).$$

By (B.15), we have $-(f + 2\rho\mu_n)/(2\rho n\sigma_n) = \sqrt{2}\vartheta_2(n,p) + o(p^{2/d_2}n^{-2})$, which is of the order of $p^{2/d_2}n^{-2}$ with $d_2 = 2/3$. Thus when $p^{2/d_2}n^{-2} \to 0$, we also know that the first two terms in the right hand side of (4) are consistent with (2). For other likelihood ratio tests (II)–(VI), similar conclusions also hold by the proofs in Section C.1.

C. PROOFS OF OTHER PROBLEMS

In this section, we provide the proofs of other testing problems following similar arguments to that in Section B. Particularly, for tests (I)–(II) and (IV)–(VII), Theorems 2.1, A.1 and A.4 are proved in Section C.1; Theorems 2.2, A.2 and A.5 are proved in Section C.3, Theorems 2.3, A.3, and A.6 are proved in Section C.4. Propositions A.1 and A.2 are proved in Section C.2.

C.1. *Proof of Theorems* 2.1*, A.*1 *&* A.4

When p is fixed, the chi-squared approximations hold by the classical multivariate analysis (Anderson, 2003; Muirhead, 2009). Therefore, without loss of generality, the proofs below focus on $p \to \infty$. In addition, we note that the analysis of taking subsequences in Section B.1 can be used similarly in the following proofs, and thus we consider without loss of generality that the sequence p/n has a limit below. We next study six likelihood ratio tests in the following subsections separately.

C.1.1. *Proof of Theorem* 2.1 *(I): Testing One-Sample Mean Vector* Similarly to the proof above, we derive the necessary and sufficient conditions for the chi-squared approximations by examining the moment generating functions. Note that testing one-sample mean vector can be viewed as testing coefficient vector μ of the multivariate linear regression $x_i = 1 \times \mu + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \Sigma)$. Motivated by the approximate expansion of multivariate Gamma function in Jiang and Yang (2013), He et al. (2020) studied the moment generating function of the likelihood ratio test in high-dimensional multivariate linear regression. Particularly, by Theorem 3 in He et al. (2020), we know that when $n, p \to \infty$ and $n - p \to \infty$, (B.2) holds with

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

$$\sigma_n^2 = \frac{1}{2}\left\{\log\left(1 - \frac{p}{n}\right) - \log\left(1 - \frac{p}{n-1}\right)\right\}. \tag{C.2}$$

Following the analysis in Section B.1, we know that to derive the necessary and sufficient conditions for the chi-squared approximations without and with the Bartlett correction, it is equivalent to examine (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.1) and σ_n in (C.2).

(I.i) The chi-squared approximation. When $p/n \to 0$, we apply Theorem 1 in He et al. (2020), and know that (B.7)–(B.8) hold if and only if $p^3/n^2 \to 0$. When $p/n \to C \in (0,1]$, we have

$$2\sigma_n^2 = \log\left\{1 + \left(1 - \frac{p}{n-1}\right)^{-1} \frac{p}{n(n-1)}\right\} \sim \frac{C}{n(1-C)},$$

and then $\sqrt{2f}/(2n\sigma_n) = \sqrt{2p}/(2n\sigma_n) \to \sqrt{1-C} < 1$. Therefore (B.7) fails, which suggests that the classical chi-squared approximation fails.

(I.ii) The chi-squared approximation with the Bartlett correction. When $p/n \to 0$, we apply Theorem 2 in He et al. (2020), and know that (B.9)–(B.10) hold if and only if $p^5/n^4 \to 0$. When $p/n \to C \in (0,1]$ and $n - p \to \infty$, we have $\rho \sim 1 - C/2$, and then $\sqrt{2f}/(2n\rho\sigma_n) = (1-C/2)^{-1}\sqrt{2p}/(2n\sigma_n) \to (1 - C/2)^{-1}\sqrt{1-C} < 1$. Therefore (B.9) fails, which suggests that the classical chi-squared approximation with the Bartlett correction fails.

C.1.2. *Proof of Theorem* 2.1 *(II): Testing One-Sample Covariance Matrix* Similarly to the proof in Section B.1, by Theorem 1 in Jiang and Yang (2013) and Jiang and Qi (2015), we know that under the conditions of our Theorem 2.1 and $p \to \infty$, (B.2) holds with

$$\mu_n = -\frac{(n-1)p}{2} - \frac{n-1}{2}(n-p-3/2)\log\left(1 - \frac{p}{n-1}\right), \tag{C.3}$$

$$\sigma_n^2 = -\frac{1}{2}\left\{\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right\} \times \frac{(n-1)^2}{n^2}. \tag{C.4}$$

Following the analysis above, we know that to derive the necessary and sufficient conditions for the chi-squared approximations without and with the Bartlett correction, it is equivalent to examine (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.3) and σ_n in (C.4). As analyzed in Section B.1, it suffices to discuss two cases $\lim_{n\to\infty} p/n = 0$ and $\lim_{n\to\infty} p/n = C \in (0,1]$ below.

(II.i) The chi-squared approximation.

Case (II.i.1) $\lim_{n\to\infty} p/n = 0$. As $\sqrt{2f} \sim p$, and (C.4) and (B.4) are asymptotically the same, by the proof in Section B.1, we know that (B.7) holds under this case. We next show that (B.8) holds if and only if $p^2/n \to 0$. By (B.7) and $\sqrt{2f} \sim p$, (B.8) is equivalent to $p^{-1}(f + 2\mu_n) \to 0$. By Taylor's expansion of μ_n in (C.3), we obtain

$$\mu_n = -\frac{(n-1)p}{2} + \frac{(n-1)}{2}(n-p-3/2)\left\{\frac{p}{n-1} + \frac{p^2}{2(n-1)^2} + \frac{p^3}{3(n-1)^3} + O\left(\frac{p^4}{n^4}\right)\right\}.$$

Through calculations, we obtain

$$\begin{aligned} p^{-1}(f + 2\mu_n) &= p^{-1} \times \left\{-\frac{p^2}{2} + \frac{p^2(n-p)}{2(n-1)} + \frac{p^3 n}{3(n-1)^2} + o(p) + O\left(\frac{p^4}{n^2}\right)\right\} \\ &= p^{-1}\left\{-\frac{p^3}{6n} + o(p) + O\left(\frac{p^4}{n^2}\right)\right\} = -\frac{p^2}{6n}\{1 + o(1)\} + o(1), \end{aligned}$$

which goes to 0 if and only if $p^2/n \to 0$.

Case (II.i.2) $\lim_{n\to\infty} p/n = C \in (0,1]$. Similarly, as (C.4) and (B.4) are asymptotically equal, we can apply the analysis same as Section B.1, and know that the chi-squared approximation fails under this case.

(II.ii) The chi-squared approximation with the Bartlett correction.

Case (II.ii.1) $\lim_{n\to\infty} p/n = 0$. Under this case, we know (B.9) holds since $\rho = 1 + O(p/n) \to 1$ and $p/(2n\sigma_n) \to 1$ as shown above. Given (B.9), to prove (B.10), it is equivalent to prove $p^{-1}(f + 2\rho\mu_n) \to 0$. By Taylor's expansion of μ_n in (C.4), we have

$$\mu_n = -\frac{p(n-1)}{2} + \frac{(n-p-3/2)(n-1)}{2}\left\{\frac{p}{n-1} + \frac{p^2}{2(n-1)^2} + \frac{p^3}{3(n-1)^3} + \frac{p^4}{4(n-1)^4} + O\left(\frac{p^5}{n^5}\right)\right\}.$$

After calculations, we obtain

$$\begin{aligned}2\rho\mu_n = & -p\left(p+\frac{1}{2}\right)+\frac{p^3}{3(n-1)}+\frac{p^2(n-p)}{2(n-1)}-\frac{p^3(n-p)}{6(n-1)^2}\\ & +\frac{p^3(n-p)}{3(n-1)^2}-\frac{p^4 n}{9(n-1)^3}+\frac{p^4 n}{4(n-1)^3}+o(p)+O\left(\frac{p^5}{n^3}\right).\end{aligned}$$

It follows that

$$\begin{aligned}& f+2\rho\mu_n\\ =& -\frac{p^2}{2}+\frac{p^2(n-p)}{2(n-1)}+\frac{p^3}{3(n-1)}-\frac{p^3(n-p)}{6(n-1)^2}+\frac{p^3(n-p)}{3(n-1)^2}+\frac{5p^4 n}{36(n-1)^3}+o(p)+O\left(\frac{p^5}{n^3}\right)\\ =& -\frac{p^4}{36n^2}+o(p)+O\left(\frac{p^5}{n^3}\right).\end{aligned}$$

Therefore $p^{-1}\{f+\mu_n\rho(n-1)\}\to 0$ if and only if $p^3/n^2\to 0$.

Case (II.ii.2) $\lim_{n\to\infty} p/n = C\in(0,1]$. Under this case, we have $\rho\to 1-C/3$. Similarly, as (C.4) (B.4) are asymptotically equal, we can apply the proof same as in Section B.1, and know that the chi-squared approximation with the Bartlett correction also fails under this case.

C.1.3. *Proof of Theorem A.*1 *(IV): Testing the Equality of Several Mean Vectors* Note that testing the equality of several mean vectors can be viewed as testing the coefficient matrix in multivariate linear regression; see, Section 10.7 in Muirhead (2009). Similarly to Section C.1.1, by Theorem 3 in He et al. (2020), we know that when $n,p\to\infty$ and $n-p\to\infty$, (B.2) holds with

$$\begin{aligned}\mu_n = \frac{n}{2}\Bigg\{&(n-p-k-1/2)\log\frac{(n-1-p)(n-k)}{(n-p-k)(n-1)}\\ &+(k-1)\log\frac{(n-1-p)}{(n-1)}+p\log\frac{(n-k)}{(n-1)}\Bigg\},\end{aligned}\tag{C.5}$$

$$\sigma_n^2=\frac{1}{2}\left\{\log\left(1-\frac{p}{n-1}\right)-\log\left(1-\frac{p}{n-k}\right)\right\}.\tag{C.6}$$

Following the analysis in Section B.1, we know to derive the necessary and sufficient conditions for the chi-squared approximations without and with the Bartlett correction, it is equivalent to examine (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.5) and σ_n in (C.6).

(IV.i) The chi-squared approximation. When $p/n\to 0$, we apply Theorem 1 in He et al. (2020), and know that (B.7)–(B.8) hold if and only if $p^3/n^2\to 0$. When $p/n\to C\in(0,1]$ and $n-p\to\infty$, we have $\sigma_n^2\sim C(k-1)/\{2n(1-C)\}$, and then $\sqrt{2f}/(2n\sigma_n)=\sqrt{2(k-1)p}/(2n\sigma_n)\to\sqrt{1-C}<1$. Therefore (B.7) fails, which suggests that the classical chi-squared approximation fails.

(IV.ii) The chi-squared approximation with the Bartlett correction. When $p/n\to 0$, we apply Theorem 2 in He et al. (2020), and know that (B.9)–(B.10) hold if and only if $p^5/n^4\to 0$. When $p/n\to C\in(0,1]$ and $n-p\to\infty$, we have $\rho\sim 1-C/2$, and then $\sqrt{2f}/(2n\rho\sigma_n)=(1-C/2)^{-1}\sqrt{2p}/(2n\sigma_n)\to(1-C/2)^{-1}\sqrt{1-C}<1$. Therefore (B.9) fails, which suggests that the classical chi-squared approximation with the Bartlett correction fails.

C.1.4. *Proof of Theorem A.*1 *(V): Testing the Equality of Several Covariance Matrices* Similarly to the proof in Section B.1, by Theorem 4 in Jiang and Yang (2013) and Jiang and Qi (2015), we know that

under the conditions of Theorem A.1 and $p \to \infty$, (B.2) holds with

$$\mu_n = \frac{1}{4}\left\{(n-k)(2n-2p-2k-1)\log\left(1-\frac{p}{n-k}\right)\right. \tag{C.7}$$
$$\left. -\sum_{i=1}^{k}(n_i-1)(2n_i-2p-3)\log\left(1-\frac{p}{n_i-1}\right)\right\},$$
$$\sigma_n^2 = \frac{(n-k)^2}{2n^2}\left\{\log\left(1-\frac{p}{n-k}\right)-\sum_{i=1}^{k}\left(\frac{n_i-1}{n-k}\right)^2\log\left(1-\frac{p}{n_i-1}\right)\right\}. \tag{C.8}$$

Following the analysis in Section B.1, we next derive the equivalent conditions for (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.7) and σ_n in (C.8).

(V.i) The chi-squared approximation.
Case (V.i.1) $\lim_{n\to\infty} p/n = 0$. Under this case, we show that (B.7) holds. By Taylor's expansion,

$$\begin{aligned}\sigma_n^2 &= \frac{(n-k)^2}{2n^2}\left[-\frac{p}{n-k}-\frac{p^2}{2(n-k)^2}+\sum_{i=1}^{k}\left(\frac{n_i-1}{n-k}\right)^2\left\{\frac{p}{n_i-1}+\frac{p^2}{2(n_i-1)^2}\right\}+O\left(\frac{p^3}{n^3}\right)\right]\\ &= \frac{(n-k)^2}{2n^2}\left\{-\frac{p}{n-k}+\sum_{i=1}^{k}\frac{p(n_i-1)}{(n-k)^2}-\frac{p^2}{2(n-k)^2}+\frac{kp^2}{2(n-k)^2}+O\left(\frac{p^3}{n^3}\right)\right\}\\ &= \frac{(k-1)p^2}{4n^2}\left\{1+o(1)\right\},\end{aligned}$$

where we use $n_i = \Theta(n)$. As $\sqrt{2f} \sim p\sqrt{k-1}$, we have (B.7) holds. Given (B.7), we know that (B.8) is equivalent to $(2f+4\mu_n)/(2p\sqrt{k-1}) \to 0$. Through Taylor's expansion, we obtain

$$\begin{aligned}4\mu_n = &-p(2n-2p-2k-1)-\frac{(n-p)p^2}{n-k}-\frac{2(n-p)p^3}{3(n-k)^2}+o\left(\frac{p^3}{n}\right)+o(p)\\ &+\sum_{i=1}^{k}p(2n_i-2p-3)+\sum_{i=1}^{k}\frac{(n_i-p)\,p^2}{(n_i-1)}+\sum_{i=1}^{k}\frac{2\,(n_i-p)\,p^3}{3\,(n_i-1)^2}+o\left(\frac{p^3}{n}\right)\\ &= p(p-kp-k+1)+\frac{p^3}{3(n-k)}-\sum_{i=1}^{k}\frac{p^3}{3(n_i-1)}+o\left(\frac{p^3}{n}\right)+o(p).\end{aligned}$$

By $f = p(p+1)(k-1)/2$, we have

$$2f+4\mu_n = \frac{p^3}{3}\left(\frac{1}{n-k}-\sum_{i=1}^{k}\frac{1}{n_i-1}\right)+o\left(\frac{p^3}{n}\right)+o(p) = \Theta(p^3/n)+o(p), \tag{C.9}$$

where we use the fact that $(n-k)^{-1}-\sum_{i=1}^{k}(n_i-1)^{-1} > 0$. It follows that $(2f+4\mu_n)/(2p\sqrt{k-1}) = \Theta(p^2/n)$, which converges to 0 if and only if $p^2/n \to 0$.

Case (V.i.2) $\lim_{n\to\infty} p/n = C \in (0,1]$. Under this case, we show that (B.7) and (B.8) do not hold at the same time. Particularly, (B.7) and (B.8) together induce $4(\mu_n+n^2\sigma_n^2)/(2f) \to 0$, which indicates $2(\mu_n+n^2\sigma_n^2)n^{-2} \to 0$, and thus $g_1(C) = 0$, where we define

$$g_1(C) = (2-C)\log(1-C)-\sum_{i=1}^{k}\delta_i(2\delta_i-C)\log(1-C\delta_i^{-1}),$$

and we assume $n_i/n \to \delta_i \in (0,1)$ for $i = 1,\ldots,k$. As $p/n = (p/n_i)\times(n_i/n) < n_i/n$, we have $0 < C \le \delta_i < 1$ for $i = 1,\ldots,k$. We next show that $g_1(C) > 0$ for $C \in (0, \min_{i=1,\ldots,k}\delta_i]$ by taking deriva-

tive of $g_1(C)$. Specifically, by $\sum_{i=1}^k \delta_i = 1$ and calculations, we have

$$g_1'(C) = \sum_{i=1}^k \delta_i \left\{ -\log(1-C) - (1-C)^{-1} + \log(1 - C\delta_i^{-1}) + \delta_i(\delta_i - C)^{-1} \right\},$$
$$g_1''(C) = \sum_{i=1}^k \delta_i \times C \left\{ -(1-C)^{-2} + (\delta_i - C)^{-2} \right\}.$$

When $0 < C \le \delta_i < 1$ for $i = 1, \ldots, k$, we have $g_1''(C) > 0$ and thus $g_1'(C)$ is a monotonically increasing function of C. As $g_1'(0) = 0$, $g_1'(C) > 0$ when $0 < C < 1$ and then $g_1(C)$ is also monotonically increasing. By $g_1(0) = 0$, we further obtain $g_1(C) > 0$ when $0 < C < 1$, which contradicts with $g_1(C) = 0$. As a result, we know (B.7) and (B.8) do not hold simultaneously, which suggests that the chi-squared approximation fails.

(V.ii) The chi-squared approximation with the Bartlett correction. When $\lim_{n\to\infty} p/n = 0$, since $\rho = 1 + O(p/n) \to 1$ and (B.7) is proved above, we know (B.9) holds. Given (B.9), as $f \sim p^2(k-1)/2$, to prove (B.10), it is equivalent to show $(2f + 4\rho\mu_n)/p \to 0$, which is also equivalent to $(2f + 4\mu_n - 4\Delta_n\mu_n)/p \to 0$, where we redefine in this subsection that

$$\Delta_n = \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \times \tilde{D}_{n,1}, \qquad \tilde{D}_{n,1} = \sum_{i=1}^k \frac{1}{n_i - 1} - \frac{1}{n-k}.$$

Similarly to the analysis of (C.9), through Taylor's expansion of μ_n in (C.7), we obtain

$$2f + 4\mu_n = -\frac{p^3}{3} \times \tilde{D}_{n,1} - \frac{p^4}{6} \times \tilde{D}_{n,2} + o\left(\frac{p^4}{n^2}\right) + o(p), \tag{C.10}$$

where $\tilde{D}_{n,2} = \sum_{i=1}^k (n_i - 1)^{-2} - (n-k)^{-2}$. Moreover, by (C.9) and $\Delta_n = O(p/n) = o(1)$, we have

$$4\Delta_n\mu_n = \Delta_n \left(-\frac{p^3}{3} \times \tilde{D}_{n,1} - 2f \right) + o\left(\frac{p^4}{n^2}\right) + o(p), \tag{C.11}$$

Combining (C.10) and (C.11), we have

$$\begin{aligned} 2f + 4\mu_n - 4\Delta_n\mu_n =& -\frac{p^3}{3} \times \tilde{D}_{n,1} - \frac{p^4}{6} \times \tilde{D}_{n,2} + \Delta_n \left(\frac{p^3}{3} \times \tilde{D}_{n,1} + 2f \right) + o\left(\frac{p^4}{n^2}\right) + o(p), \\ =& \frac{p^4}{18(k-1)} \left\{ 2\tilde{D}_{n,1}^2 - 3(k-1)\tilde{D}_{n,2} \right\} + o\left(\frac{p^4}{n^2}\right) + o(p), \end{aligned} \tag{C.12}$$

where we use $\tilde{D}_{n,1} = \Theta(n^{-1})$, $\tilde{D}_{n,2} = \Theta(n^{-2})$, $\Delta_n = p\tilde{D}_{n,1}/\{3(k-1)\} + o(p/n)$, and $2\Delta_n f = p^3\tilde{D}_{n,1}/3 + o(p)$.

We next show that (C.12) $= \Theta(p^4 n^{-2})$. In particular, in this subsection, we redefine $\delta_i = (n_i - 1)/(n-k)$, which satisfies $\sum_{i=1}^k \delta_i = 1$. Then by the definitions of $\tilde{D}_{n,1}$ and $\tilde{D}_{n,2}$, we calculate that

$$\begin{aligned} & (n-k)^2 \times \{2\tilde{D}_{n,1}^2 - 3(k-1)\tilde{D}_{n,2}\} \\ =& (5-3k)\sum_{i=1}^k \delta_i^{-2} + 2\sum_{1\le i\ne j\le k} \delta_i^{-1}\delta_j^{-1} - 4\sum_{i=1}^k \delta_i^{-1} + 3k - 1. \end{aligned} \tag{C.13}$$

As $2\delta_i^{-1}\delta_j^{-1} \le \delta_i^{-2} + \delta_j^{-2}$, we have

$$\begin{aligned}\text{(C.13)} &\le (3-k)\sum_{i=1}^{k}\delta_i^{-2} - 4\sum_{i=1}^{k}\delta_i^{-1} + 3k - 1 \\ &\le (3-k)\sum_{i=1}^{k}\delta_i^{-2} - 4k^2 + 3k - 1, \end{aligned} \tag{C.14}$$

where in the last inequality, we use $\sum_{i=1}^{k}\delta_i^{-1} \ge k^2(\sum_{i=1}^{k}\delta_i)^{-1} = k^2$. Therefore (C.14) < 0 when $k \ge 3$. When $k = 2$, as $\delta_1 + \delta_2 = 1$, we have $\delta_1^{-1} + \delta_2^{-1} = \delta_1^{-1}\delta_2^{-1}$ and (C.13) $= -\sum_{i=1}^{2}\delta_i^{-2} - 2\sum_{i=1}^{2}\delta_i^{-1} + 5$. As $\sum_{i=1}^{2}\delta_i^{-1} \ge 2^2$, (C.13) $< -2 \times 2^2 + 5 < 0$. In summary, we know (C.13) < 0 for $k \ge 2$, and thus (C.12) $= \Theta(p^4 n^{-2})$. If follows that $(2f + 4\rho\mu_n)/p \to 0$ if and only if $p^3/n^2 \to 0$. In summary, we know for testing problem (V), the chi-squared approximation with the Bartlett correction works if and only if $p^3/n^2 \to 0$.

C.1.5. *Proof of Theorem A.1 (VI): Joint Testing the Equality of Several Mean Vectors and Covariance Matrices* Similarly to the proof in Section B.1, by Theorem 3 in Jiang and Yang (2013) and Jiang and Qi (2015), we know that under the conditions of Theorem A.1 and $p \to \infty$, (B.2) holds with

$$\mu_n = \frac{1}{4}\left\{-2kp - \sum_{i=1}^{k}\frac{p}{n_i} - nL_{n,p}(2p - 2n + 3) + \sum_{i=1}^{k} n_i L_{n_i-1,p}(2p - 2n_i + 3)\right\}, \tag{C.15}$$

$$\sigma_n^2 = \frac{1}{2}\left(L_{n,p} - \sum_{i=1}^{k}\frac{n_i^2}{n^2} \times L_{n_i-1,p}\right), \tag{C.16}$$

where $L_{n,p} = \log(1 - p/n)$. Following Section B.1, we next derive the equivalent conditions for (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.15) and σ_n in (C.16).

(VI.i) The chi-squared approximation.
Case (VI.i.1) $\lim_{n\to\infty} p/n = 0$. Under this case, we show that (B.7) holds. As $-\log(1-x) = x + x^2/2 + O(x^3)$ and $n_i = \Theta(n)$, we obtain

$$\begin{aligned} 2\sigma_n^2 = \ &= \sum_{i=1}^{k}\frac{n_i^2}{n^2}\left\{\frac{p}{n_i - 1} + \frac{p^2}{2(n_i-1)^2}\right\} - \frac{p}{n} - \frac{p^2}{2n^2} + O\left(\frac{p^3}{n^3}\right) \\ &= \sum_{i=1}^{k}\frac{n_i^2}{n^2}\left(\frac{p}{n_i} + \frac{p}{n_i^2} + \frac{p^2}{2n_i^2}\right) - \frac{p}{n} - \frac{p^2}{2n^2} + O\left(\frac{p^3}{n^3}\right) \\ &= \frac{kp}{n^2} + \frac{(k-1)p^2}{2n^2} + O\left(\frac{p^3}{n^3}\right), \end{aligned}$$

where in the second equation, we use $(n_i - 1)^{-1} = n_i^{-1} + n_i^{-2} + O(n_i^{-3})$ and $(n_i - 1)^{-2} = n_i^{-2} + O(n_i^{-3})$. It follows that $2n\sigma_n \sim p\sqrt{k-1}$. By $\sqrt{2f} \sim p\sqrt{k-1}$, we have (B.7). Given (B.7), we know that (B.8) is equivalent to $(2f + 4\mu_n)/p \to 0$. As $p/n = o(1)$, through Taylor's expansion, we obtain

$$\begin{aligned} -n(2p - 2n + 3)L_{n,p} &= n(2p - 2n + 3)\left\{\frac{p}{n} + \frac{p^2}{2n^2} + \frac{p^3}{3n^3} + O\left(\frac{p^4}{n^4}\right)\right\} \\ &= p\left\{2p + \frac{p^2}{n} - 2n - p - \frac{2p^2}{3n} + 3 + O\left(\frac{p^3}{n^2}\right) + o(1)\right\} \\ &= p\left\{p + \frac{p^2}{3n} - 2n + 3 + O\left(\frac{p^3}{n^2}\right) + o(1)\right\}. \end{aligned} \tag{C.17}$$

Similarly, by Taylor's expansion and $n_i = \Theta(n)$, we have

$$\begin{aligned} & - n_i(2p - 2n_i + 3)L_{n_i-1,p} \qquad \text{(C.18)} \\ &= n_i(2p - 2n_i + 3)\left\{ \frac{p}{n_i - 1} + \frac{p^2}{2(n_i-1)^2} + \frac{p^3}{3(n_i-1)^3} + O\left(\frac{p^4}{n^4}\right)\right\} \\ &= n_i(2p - 2n_i + 3)\left\{ \frac{p}{n_i} + \frac{p}{n_i^2} + \frac{p^2}{2n_i^2} + \frac{p^3}{3n_i^3} + O\left(\frac{p^4}{n^4}\right) + O\left(\frac{p^2}{n^3}\right)\right\} \\ &= p\left\{ p + \frac{p^2}{3n_i} - 2n_i + 3 - 2 + O\left(\frac{p^3}{n_i^2}\right) + o(1)\right\}, \end{aligned}$$

where in the second equation, we use $(n_i - 1)^{-1} = n_i^{-1} + n_i^{-2} + O(n_i^{-3})$ and $(n_i - 1)^{-a} = n_i^{-a} + O(n_i^{-3})$ for integers $a \geq 2$. Combining (C.17) and (C.18), we obtain

$$\begin{aligned} 2f + 4\mu_n &= 2f - 2kp + p\left\{ (1-k)p + \frac{p^2}{3}\Big(\frac{1}{n} - \sum_{i=1}^{k}\frac{1}{n_i}\Big) + 3 - k\right\} + O\Big(\frac{p^4}{n^2}\Big) + o(p) \qquad \text{(C.19)} \\ &= \frac{p^3}{3}\Big(\frac{1}{n} - \sum_{i=1}^{k}\frac{1}{n_i}\Big) + O\Big(\frac{p^4}{n^2}\Big) + o(p). \end{aligned}$$

As $n^{-1} - \sum_{i=1}^{k} n_i^{-1} = \Theta(n^{-1})$, we have $2f + 4\mu_n = \Theta(p^3 n^{-1})$. Therefore we know $(2f + 4\mu_n)/p \to 0$ if and only if $p^2/n \to 0$.

Case (VI.i.2) $\lim_{n\to\infty} p/n = C \in (0, 1]$. In this subsection, we redefine $\delta_i = n_i/n \in (0, 1)$. Then

$$\frac{4n^2\sigma_n^2}{2f} \to \frac{2}{C^2(k-1)} \times \Big\{ \log(1-C) - \sum_{i=1}^{k} \delta_i^2 \log(1 - C\delta_i^{-1})\Big\},$$

where $0 < C \leq \delta_i < 1$. Therefore (B.7) induces $g_2(C) = 0$, where we define

$$g_2(C) = \log(1-C) - \sum_{i=1}^{k} \delta_i^2 \log(1 - C\delta_i^{-1}) - (k-1)C^2/2.$$

By taking derivative of $g_2(C)$, we obtain $g_2'(0) = 0$, $g_2''(0) = 0$, and

$$g_2'''(C) = \frac{2}{(C-1)^2} - \sum_{i=1}^{k} \frac{2\delta_i^2}{(C-\delta_i)^3} = \sum_{i=1}^{k} \frac{2\delta_i(1-\delta_i)(C^3 - 3\delta_i C + \delta_i^2 + \delta_i)}{(1-C)^3(\delta_i - C)^3}.$$

As $C^3 - 3\delta_i C + \delta_i^2 + \delta_i$ is a monotonically decreasing function of C when $0 < C \leq \delta_i < 1$, and it equals $\delta_i(\delta_i - 1)^2 > 0$ when $C = \delta_i$, we have $g_2'''(C) > 0$ for $0 < C \leq \delta_i$. It follows that $g_2(C)$ is a monotonically increasing function when $0 < C \leq \delta_i < 1$. As $g_2(0) = 0$, we have $g_2(C) > 0$, which contradicts with $g_2(C) = 0$. Therefore, we know that (B.7) does not hold under this case, which implies that the chi-squared approximation fails.

(VI.ii) The chi-squared approximation with the Bartlett correction. When $\lim_{n\to\infty} p/n = 0$, since $\rho = 1 + O(p/n) \to 1$ and (B.7) is proved above, we know (B.9) holds. Given (B.9), as $f \sim p^2(k-1)/2$, to prove (B.10), it is equivalent to show $(2f + 4\rho\mu_n)/p \to 0$, which is equivalent to $(2f + 4\mu_n - 4\Delta_n\mu_n)/p \to 0$, where in this subsection, we redefine

$$\Delta_n = \frac{2p^2 + 9p + 11}{6(p+3)(k-1)} \times D_{n,1}, \qquad D_{n,1} = \sum_{i=1}^{k} \frac{1}{n_i} - \frac{1}{n}.$$

Similarly to (C.17), through Taylor's expansion, we further have

$$n(2p - 2n + 3)r_n^2 = p\left\{ p - 2n + 3 + \frac{p^2}{3n} + \frac{p^3}{6n^2} + O\Big(\frac{p^4}{n^3}\Big) + o(1)\right\}.$$

In addition, similarly to (C.18), we have

$$n_i(2p-2n_i+3)r_{n_i'}^2 = p\left\{p-2n_i+3-2+\frac{p^2}{3n_i}+\frac{p^3}{6n_i^2}+O\Big(\frac{p^4}{n_i^3}\Big)+o(1)\right\}. \tag{C.20}$$

It follows that

$$2f+4\mu_n = -\frac{p^3}{3}D_{n,1}-\frac{p^4}{6}D_{n,2}+O\Big(\frac{p^5}{n^3}\Big)+o(p), \tag{C.21}$$

where $D_{n,2}=\sum_{i=1}^k n_i^{-2}-n^{-2}$. Moreover, by (C.19) and $\Delta_n = O(p/n)=o(1)$,

$$4\Delta_n\mu_n = \Delta_n\Big(-\frac{p^3}{3}D_{n,1}-2f\Big)+O\Big(\frac{p^5}{n^3}\Big)+o(p). \tag{C.22}$$

Combining (C.21) and (C.22), we obtain

$$2f+4\mu_n-4\Delta_n\mu_n = \frac{p^4}{18(k-1)}\{2D_{n,1}^2-3(k-1)D_{n,2}\}+O\Big(\frac{p^5}{n^3}\Big)+o(p), \tag{C.23}$$

where we use $D_{n,1}=\Theta(n^{-1})$, $D_{n,2}=\Theta(n^{-2})$, $\Delta_n = pD_{n,1}/\{3(k-1)\}+o(p/n)$, and $2\Delta_n f = p^3D_{n,1}/3+o(p)$. Following the analysis of (C.13), we know (C.23) $=\Theta(p^4n^{-2})$. Therefore, $(2f+\rho\mu_n)/p\to 0$ if and only if $p^3/n^2\to 0$, which suggests that the chi-squared approximation with the Bartlett correction holds if and only if $p^3/n^2\to 0$.

C.1.6. *Proof of Theorem A.4 (VII): Testing Independence between Multiple Vectors* Similarly to the proof in Section B.1, by Theorem 2 in Jiang and Yang (2013) and Jiang and Qi (2015), we know that under the conditions of Theorem A.4 and $p\to\infty$, (B.2) holds with

$$\mu_n = \frac{n}{2}\left[-\left(n-p-\frac{3}{2}\right)L_{n-1,p}+\sum_{j=1}^k\left\{\left(n-p_j-\frac{3}{2}\right)L_{n-1,p_j}\right\}\right] \tag{C.24}$$

$$\sigma_n^2 = \frac{1}{2}\Big(-L_{n-1,p}+\sum_{j=1}^k L_{n-1,p_j}\Big). \tag{C.25}$$

Following the analysis in Section B.1, we next derive the equivalent conditions for (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.24) and σ_n in (C.25).

(VII.i) The chi-squared approximation.
Case (VI.i.1) $\lim_{n\to\infty}p/n=0$. Under this case, we show that (B.7) holds. Through Taylor's expansion,

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

Recall that $2f=p^2-\sum_{i=1}^k p_i^2$, and thus (B.7) holds. As $f=\Theta(p^2)$ undert the conditions of Theorem A.4, given (B.7), we know (B.8) is equivalent to $(2f+4\mu_n)/p\to 0$. Similarly to the analysis of (C.18), through Taylor's expansion, we have

$$\begin{aligned}n(2n-2p-3)L_{n-1,p} &= p\Big\{p+\frac{p^2}{3n}-2n+1+O\Big(\frac{p^3}{n^2}\Big)+o(1)\Big\},\\ n(2n-2p_i-3)L_{n-1,p_i} &= p_i\Big\{p_i+\frac{p_i^2}{3n}-2n+1+O\Big(\frac{p_i^3}{n^2}\Big)+o(1)\Big\}.\end{aligned}$$

It follows that

$$\begin{aligned}
&2f + 4\mu_n \\
&= p^2 - \sum_{i=1}^{k} p_i^2 - p\Big(p + \frac{p^2}{3n} - 2n + 1\Big) + \sum_{i=1}^{k} p_i\Big(p_i + \frac{p_i^2}{3n} - 2n + 1\Big) + O\Big(\frac{p^4}{n^2}\Big) + o(p) \\
&= \frac{1}{3n}\Big(\sum_{i=1}^{k} p_i^3 - p^3\Big) + O\Big(\frac{p^4}{n^2}\Big) + o(p).
\end{aligned}$$

Under the conditions of Theorem A.4, we have $\sum_{i=1}^{k} p_i^3 - p^3 = \Theta(p^3)$. Thus $(2f + 4\mu_n)/p \to 0$ if and only if $p^2/n \to 0$, which suggests that the chi-squared approximation holds if and only if $p^2/n \to 0$.

Case (VI.i.2) $\lim_{n\to\infty} p/n = C \in (0, 1]$. Under this case, we show that (B.7) and (B.8) do not hold at the same time. Particularly, as $f = \Theta(p^2)$ and $p/n \to C \in (0, 1]$, (B.7) induces $(2n^2\sigma_n^2 - f)/n^2 \to 0$, and (B.8) induces $(f + 2\mu_n)/n^2 \to 0$. Therefore, (B.7) and (B.8) together give $(2n^2\sigma_n^2 + 2\mu_n)/n^2 \to 0$. Suppose $\lim_{n\to\infty} p_i/n = C_i \in (0, 1)$. It follows that $\sum_{i=1}^{k} C_i = C$, and

$$(2n^2\sigma_n^2 + 2\mu_n)/n^2 \to -g_3(C) + \sum_{i=1}^{k} g_3(C_i), \tag{C.26}$$

where $g_3(C) = (2 - C)\log(1 - C)$. Note that $g_3(C)$ is a strictly concave function of $C \in (0, 1]$ and $g(0) = 0$. By the property of strictly concave function, we have

$$\sum_{i=1}^{k} g_3(C_i) = \sum_{i=1}^{k} g_3(C \times C_i/C) > \sum_{i=1}^{k} g_3(C) \times C_i/C = g_3(C),$$

where we use $\sum_{i=1}^{k} C_i = C$. Therefore when $C \in (0, 1]$, the right hand side of (C.26) > 0, which contradicts with $(2n^2\sigma_n^2 + 2\mu_n)/n^2 \to 0$. We thus know that (B.7) and (B.8) do not hold simultaneously, which suggests that the chi-squared approximation fails.

(VII.ii) The chi-squared approximation with the Bartlett correction. When $\lim_{n\to\infty} p/n = 0$, since $\rho = 1 + O(p/n) \to 1$ and (B.7) is proved above, we know (B.9) holds. Given (B.9), as $f = \Theta(p^2)$, to prove (B.10), it is equivalent to show $(2f + 4\rho\mu_n)/p \to 0$, which is equivalent to $(2f + 4\mu_n - 4\Delta_n\mu_n)/p \to 0$, where in this subsection, we redefine

$$\Delta_n = \frac{2 \times D_{p,3} + 9 \times D_{p,2}}{6n \times D_{p,2}}, \quad D_{p,3} = p^3 - \sum_{i=1}^{k} p_i^3, \quad D_{p,2} = p^2 - \sum_{i=1}^{k} p_i^2.$$

Similarly to (C.20), through Taylor's expansion, we further obtain

$$\begin{aligned}
n(2n - 2p - 3)L_{n-1,p} &= p\Big\{p - 2n + 1 + \frac{p^2}{3n} + \frac{p^3}{6n^2} + O\Big(\frac{p^4}{n^3}\Big) + o(1)\Big\}, \\
n(2n - 2p_i - 3)L_{n-1,p_i} &= p_i\Big\{p_i - 2n + 1 + \frac{p_i^2}{3n} + \frac{p_i^3}{6n^2} + O\Big(\frac{p_i^4}{n^3}\Big) + o(1)\Big\}.
\end{aligned}$$

It follows that

$$2f + 4\mu_n = -\frac{1}{3n}D_{p,3} - \frac{1}{6n^2}D_{p,4} + O\Big(\frac{p^5}{n^3}\Big) + o(p), \tag{C.27}$$

where $D_{p,4} = p^4 - \sum_{i=1}^{k} p_i^4$. Moreover, as $\Delta_n = \Theta(p/n)$, by (C.27) and $2f = D_{p,2}$, we have

$$4\Delta_n\mu_n = \Delta_n\Big(-\frac{1}{3n}D_{p,3} - D_{p,2}\Big) + O\Big(\frac{p^5}{n^3}\Big) + o(p).$$

As $\Delta_n = D_{p,3}/(3nD_{p,2}) + O(n^{-1})$, we calculate that

$$\begin{aligned} & 2f + 4\mu_n - 4\Delta_n\mu_n \qquad\qquad \text{(C.28)} \\ = & -\frac{1}{3n}D_{p,3} - \frac{1}{6n^2}D_{p,4} + \frac{D_{p,3}}{3nD_{p,2}}\Big(\frac{1}{3n}D_{p,3} + D_{p,2}\Big) + O\Big(\frac{p^5}{n^3}\Big) + o(p) \\ = & -\frac{1}{18n^2 D_{p,2}}(3D_{p,4}D_{p,2} - 2D_{p,3}^2) + O\Big(\frac{p^5}{n^3}\Big) + o(p). \end{aligned}$$

We next prove (C.28) $= \Theta(p^4 n^{-2})$ by showing $3D_{p,4}D_{p,2} - 2D_{p,3}^2 = \Theta(p^6)$. Specifically, by the definitions of $D_{p,2}$, $D_{p,3}$, and $D_{p,4}$, we write

$$\begin{aligned} & 3D_{p,4}D_{p,2} - 2D_{p,3}^2 \qquad\qquad \text{(C.29)} \\ = & \, p^4\Big(p^2 - \sum_{i=1}^k p_i^2\Big) + 2p^3\Big(-p\sum_{i=1}^k p_i^2 + \sum_{i=1}^k p_i^3\Big) + 2p^2\Big(p\sum_{i=1}^k p_i^3 - \sum_{i=1}^k p_i^4\Big) \\ & + \Big(-p^2 + \sum_{i=1}^k p_i^2\Big)\sum_{i=1}^k p_i^4 + 2\left\{\sum_{i=1}^k p_i^2 \sum_{i=1}^k p_i^4 - \Big(\sum_{i=1}^k p_i^3\Big)^2\right\}. \end{aligned}$$

Using $p = \sum_{i=1}^k p_i$, we obtain

$$p\sum_{i=1}^k p_i^\alpha - \sum_{i=1}^k p_i^{\alpha+1} = \sum_{i\neq j} p_i p_j^\alpha, \qquad p\sum_{i\neq j} p_i p_j - 2\sum_{i\neq j} p_i^2 p_j = \sum_{i\neq j\neq l} p_i p_j p_l, \qquad \text{(C.30)}$$

where integer $1 \le \alpha \le 5$, and we use $\sum_{i\neq j}$ and $\sum_{i\neq j\neq l}$ to denote the summation $\sum_{1\le i\neq j\le k}$ and $\sum_{1\le i\neq j\neq l\le k}$ for simplicity. By (C.30), we calculate that

$$\begin{aligned} \text{(C.29)} &= p^3 \sum_{i\neq j\neq l} p_i p_j p_l + 2p^2 \sum_{i\neq j} p_i^3 p_j - \sum_{i\neq j} p_i p_j \sum_{l=1}^k p_l^4 - 2\sum_{i\neq j} p_i^3 p_j^3 + 2\sum_{i\neq j} p_i^2 p_j^4 \\ &> 2p^2 \sum_{i\neq j} p_i^3 p_j - \sum_{i\neq j} p_i p_j \sum_{l=1}^k p_l^4 - 2\sum_{i\neq j} p_i^3 p_j^3 \\ &= 2\Big(\sum_{i=1}^k p_i^2 + \sum_{i\neq j} p_i p_j\Big)\sum_{i\neq j} p_i^3 p_j - 2\sum_{i\neq j} p_i p_j^5 - \sum_{i\neq j\neq l} p_i p_j p_l^4 - 2\sum_{i\neq j} p_i^3 p_j^3 > 0. \end{aligned}$$

Therefore (C.29) $= \Theta(p^6)$ and then (C.28) $= \Theta(p^4 n^{-2})$. Thus $(2f + 4\rho\mu_n)/p \to 0$ if and only if $p^3/n^2 \to 0$, which suggests that the chi-squared approximation with the Bartlett correction holds if and only if $p^3/n^2 \to 0$.

C.2. *Proofs of Propositions A.1 & A.2*

This section proves Propositions A.1 and A.2 following similar arguments to that in Sections B.1 and C.1. In particular, we consider without loss of generality that $p \to \infty$ and p/n has a limit.

C.2.1. *Proof of Proposition A.1* Following the analysis in Section C.1.3, we know that when $n, p \to \infty$, $n - k \to \infty$, and $n - p \to \infty$, (B.2) holds with μ_n in (C.5) and σ_n^2 (C.6). Moreover, to derive the necessary and sufficient conditions for the chi-squared approximations without and with the Bartlett correction, it is equivalent to examine (B.7)–(B.8) and (B.9)–(B.10), respectively, with μ_n in (C.5) and σ_n in (C.6).

(i) The chi-squared approximation. (i.1) When $p/n \to 0$ and $k/n \to 0$, we apply Theorem 1 in He et al. (2020), and know that (B.7)–(B.8) hold if and only if $\sqrt{pk}(p+k)/n \to 0$. (i.2) When $p/n \to C \in (0, 1]$ and $k/n \to 0$, we have $f \sim C(k-1)n$ and $2\sigma_n^2 \sim C(k-1)/\{n(1-C)\}$. It follows that

$\sqrt{2f}/(2n\sigma_n) \sim \sqrt{1-C} < 1$. Thus (B.7) fails, which suggests that the chi-squared approximation fails. (i.3) When $p/n \to 0$ and $k/n \to C \in (0,1]$, by applying the symmetric substitution technique in Section 10.4 of Muirhead (2009), we can switch k and p and analyze similarly as in the case (i.2) above. Therefore we know the chi-squared approximation also fails here. (i.4) When $p/n \to C_1 \in (0,1]$ and $k/n \to C_2 \in (0,1]$, we know $0 < C_1 + C_2 \leq 1$ as $p + k < n$. By the constraint, it then suffices to consider $C_1, C_2 \in (0,1)$. Note that $2\sigma_n^2 \sim \log\{(1-C_1)(1-C_2)\} - \log(1-C_1-C_2)$ and $2f/n^2 \sim 2C_1C_2$. Thus (B.7) induces $g_4(C_1, C_2) = 0$ where $g_4(C_1, C_2) = C_1C_2 - \log\{(1-C_1)(1-C_2)\} + \log(1-C_1-C_2)$. If $C_1 + C_2 = 1$, $g_4(C_1, C_2) \to -\infty$. We next consider $0 < C_1 + C_2 < 1$. By calculations, we have $g_4(0, C_2) = 0$, and

$$\frac{\mathrm{d}}{\mathrm{d}C_1} g_4(C_1, C_2) = \frac{C_2\{(C_1-1)(C_1+C_2) - C_1\}}{(1-C_1)(1-C_1-C_2)} < 0,$$

where we use $C_1, C_2 \in (0,1)$ and $0 < C_1 + C_2 < 1$. Similarly to the previous analyses, we know that $g_4(C_1, C_2)$ is monotonically decreasing for $C_1 \in (0,1)$ and thus $g_4(C_1, C_2) < 0$, as $C_1 \in (0,1)$ and $g_4(0, C_2) = 0$. Therefore (B.7) fails, which suggests that the classical chi-squared approximation fails.

(ii) The chi-squared approximation with the Bartlett correction. (ii.1) When $p/n \to 0$ and $k/n \to 0$, we apply Theorem 2 in He et al. (2020), and know that (B.9)–(B.10) hold if and only if $\sqrt{pk}(p^2+k^2)/n^2 \to 0$. (ii.2) When $p/n \to C \in (0,1]$ and $k/n \to 0$, we have $\rho \sim 1 - C/2$, and the proof of part (IV.ii) in Section C.1.3 can be applied here similarly. Thus the chi-squared approximation fails. (ii.3) When $p/n \to 0$ and $k/n \to C \in (0,1]$, we know the chi-squared approximation also fails by switching k and p symmetrically as in the case (i.3) above. (ii.4) When $p/n \to C_1 \in (0,1]$ and $k/n \to C_2 \in (0,1]$, we know $0 < C_1 + C_2 \leq 1$ as $p + k < n$. Similarly to the case (i.4) above, we consider $C_1, C_2 \in (0,1)$ and $C_1 + C_2 < 1$. Here $\rho \sim 1 - (C_1 + C_2)/2$ and then (B.9) induces $g_5(C_1, C_2) = 0$, where $g_5(C_1, C_2) = 2C_1C_2 - (2 - C_1 - C_2)[\log\{(1-C_1)(1-C_2)\} - \log(1-C_1-C_2)]$. By calculations, we have $g_5(0, C_2) = 0$, and

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}C_1} g_5(C_1, C_2)|_{C_1=0} = -C_2/(1-C_2) < 0, \\
&\frac{\mathrm{d}^2}{\mathrm{d}^2C_1} g_5(C_1, C_2) = -\frac{C_2\{(C_1+C_2)(C_2-2)+2\}}{(1-C_1)^2(1-C_1-C_2)^2} < 0,
\end{aligned}$$

where we use $(C_1 + C_2)(C_2 - 2) + 2 > 0$ as $0 < C_1 + C_2 < 1$ and $-2 < C_2 - 2 < -1$. Similarly to the analysis above, we know that $g_5(C_1, C_2) < 0$ and thus (B.9) fails, which suggests that the chi-squared approximation with the Bartlett correction fails.

C.2.2. *Proof of Proposition A.2* *(i) The chi-squared approximation.* (i.1) When $p_1/n \to 0$ and $p_2/n \to 0$, we apply Theorem 1 in He et al. (2020), and know that (B.7)–(B.8) hold if and only if $\sqrt{p_1p_2}(p_1+p_2)/n \to 0$. (i.2) When $p_1/n \to C \in (0,1]$ and $p_2/n \to 0$, we have $2f \sim Cnp_2$ and $2\sigma_n^2 \sim Cp_2/\{2n(1-C)\}$. Then $\sqrt{2f}/(2n\sigma_n) \sim \sqrt{1-C} < 1$ suggesting the failure of (B.7) and thus the chi-squared approximation fails. (i.3) When $p_1/n \to 0$ and $p_2/n \to C \in (0,1]$, the chi-squared approximation also fails by the symmetric substitution technique in Section C.2.1. (i.4) When $p_1/n \to C_1 \in (0,1]$ and $p_2/n \to C_2 \in (0,1]$, we have $2\sigma_n^2 \sim \log\{(1-C_1)(1-C_2)\} - \log(1-C_1-C_2)$ and $2f/n^2 \sim C_1C_2$. It follows that the analysis in case (i.4) of Section C.2.1 can be applied similarly, and we obtain the same conclusion, that is, (B.7) fails and then the chi-squared approximation fails.

(ii) The chi-squared approximation with the Bartlett correction. (ii.1) When $p_1/n \to 0$ and $p_2/n \to 0$, we apply Theorem 2 in He et al. (2020), and know that (B.9)–(B.10) hold if and only if $\sqrt{p_1p_2}(p_1^2+p_2^2)/n^2 \to 0$. (ii.2) When $p_1/n \to C \in (0,1]$ and $p_2/n \to 0$, we have $\rho \sim 1 - C/2$, and then $\sqrt{2f}/(2n\rho\sigma_n) = (1-C/2)^{-1}\sqrt{2p}/(2n\sigma_n) \to (1-C/2)^{-1}\sqrt{1-C} < 1$. Therefore (B.9) fails, which suggests that the classical chi-squared approximation with the Bartlett correction fails. (ii.3) When $p_1/n \to 0$ and $p_2/n \to C \in (0,1]$, similar conclusion holds by the symmetric substitution technique as above. (ii.4) When $p_1/n \to C_1 \in (0,1]$ and $p_2/n \to C_2 \in (0,1]$, we have $\rho \sim 1 - (C_1 + C_2)/2$. It fol-

lows that the analysis in case (ii.4) of Section C.2.1 can be applied similarly. Then we obtain the same conclusion, that is, (B.9) fails and the chi-squared approximation with the Bartlett correction fails.

C.3. *Proofs of Theorems* 2.2, *A.*2 *&* A.5

In this section, we prove the results for other testing problems in Theorems 2.2, A.2 & A.5 following similar analysis to that in Section B.2. Particularly, for each test, we consider the characteristic function of $-2\eta \log \Lambda_n$ when $\eta = 1$ and ρ; here ρ denotes the corresponding Bartlett correction factor of each test.

By Eq. (20)–(23) in Section 8.2.4 of Muirhead (2009), we know that for the testing problems (I)–(II) and (IV)–(VII), the characteristic functions of the likelihood ratio test statistics take the following general form:

$$\log \mathrm{E}\{\exp(-2it\eta \log \Lambda_n)\} = \varphi(t) - \varphi(0), \tag{C.31}$$

where

$$\begin{aligned}\varphi(t) = 2it\eta &\left(\sum_{k=1}^{K_1} \xi_{1,k} \log \xi_{1,k} - \sum_{j=1}^{K_2} \xi_{2,j} \log \xi_{2,j} \right) \\ &+ \sum_{k=1}^{K_1} \log \Gamma\big\{\eta\xi_{1,k}(1-2it) + \tau_{1,k} + \upsilon_{1,k}\big\} - \sum_{j=1}^{K_2} \log \Gamma\big\{\eta\xi_{2,j}(1-2it) + \tau_{2,j} + \upsilon_{2,j}\big\},\end{aligned}$$

i denotes the imaginary unit, $\tau_{1,k} = (1-\eta)\xi_{1,k}$, and $\tau_{2,j} = (1-\eta)\xi_{2,j}$. We next consider $\eta = 1$ and ρ for the chi-squared approximation without and with the Bartlett correction, respectively. The values of ρ, K_1, K_2, $\xi_{1,k}$, $\xi_{2,j}$, $\upsilon_{1,k}$, and $\upsilon_{2,j}$ depend on the testing problem, and thus take different values in the following subsections. Moreover, by Muirhead (2009), in each problem, we have $\sum_{k=1}^{K_1} \xi_{1,k} = \sum_{j=1}^{K_2} \xi_{2,k}$, the degrees of freedom f is

$$f = -2\left\{ \sum_{k=1}^{K_1} \upsilon_{1,k} - \sum_{j=1}^{K_2} \upsilon_{2,j} - \frac{1}{2}(K_1 - K_2) \right\}, \tag{C.32}$$

and the Bartlett correction ρ takes the value

$$\rho = 1 - \frac{1}{f}\left\{ \sum_{k=1}^{K_1} \frac{\upsilon_{1,k}^2 - \upsilon_{1,k} + \frac{1}{6}}{\xi_{1,k}} - \sum_{j=1}^{K_2} \frac{\upsilon_{2,j}^2 - \upsilon_{2,j} + \frac{1}{6}}{\xi_{2,j}} \right\}. \tag{C.33}$$

In the following proofs, we use Lemma C.3.1 below to obtain an asymptotic expansion of each characteristic function.

LEMMA C.3.1. *For a finite integer L, when $\eta = 1$ or ρ, $p/n \to 0$, and $R_{n,L}$ (in* (C.34) *below) converges to* 0*,*

$$\log \mathrm{E}\{\exp(-2it\eta \log \Lambda_n)\} = -\frac{f}{2}\log(1-2it) + \sum_{l=1}^{L-1} \varsigma_l\big\{(1-2it)^{-l} - 1\big\} + R_{n,L},$$

where

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)}\left\{ \sum_{k=1}^{K_1} \frac{B_{l+1}(\tau_{1,k} + \upsilon_{1,k})}{(\eta \times \xi_{1,k})^l} - \sum_{j=1}^{K_2} \frac{B_{l+1}(\tau_{2,j} + \upsilon_{2,j})}{(\eta \times \xi_{2,j})^l} \right\},$$

$B_{l+1}(\cdot)$ denotes the $(l+1)$-th Bernoulli polynomial; see, e.g., Eq. (25) in Section 8.2.4 of Muirhead (2009), and $R_{n,L}$ denotes the remainder which is of the order of

$$R_{n,L} = O\bigg(\sum_{k=1}^{K_1} \frac{|\tau_{1,k} + \upsilon_{1,k}|^{L+1}}{|\eta\xi_{1,k}|^L} + \sum_{j=1}^{K_2} \frac{|\tau_{2,j} + \upsilon_{2,j}|^{L+1}}{|\eta\xi_{2,j}|^L} \bigg). \tag{C.34}$$

Proof. Please see Section D.2.16 on Page 73. □

We next examine each testing problem based on Lemma C.3.1.

C.3.1. *Proof of Theorem* 2.2 *(I): Testing One-Sample Mean Vector* Recall that in Section C.1.1, we mention that testing one-sample mean vector can be viewed as testing coefficient vector of a multivariate linear regression model. By Section 10.5 in Muirhead (2009), we know that in this problem, $K_1 = 1$, $K_2 = 1$, $\xi_{1,1} = n/2$, $\xi_{2,1} = n/2$, $\upsilon_{1,1} = -p/2$, $\upsilon_{2,1} = 0$, $f = p$ and $\rho = 1 - (p/2+1)/n$. We next discuss the chi-squared approximation without and with the Bartlett correction, respectively.

(i) Chi-squared approximation. Consider $\rho = 1$ and $p^3/n^2 \to 0$. Then $\tau_{1,1} = \tau_{2,1} = 0$,

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \times \frac{1}{(n/2)^l} \Big\{ B_{l+1}\Big(-\frac{p}{2}\Big) - B_{l+1}(0) \Big\}, \tag{C.35}$$

and for any finite integer L, $R_{n,L} = O(p^{L+1}n^{-L})$. Since $B_{l+1}(\cdot)$ is a polynomial of order $l+1$, then $\varsigma_l = O(p^{l+1}/n^l)$. By Lemma C.3.1, when $p^3/n^2 \to 0$, $R_{n,3} = O(p^4 n^{-3}) \to 0$, and

$$\begin{aligned} \mathrm{E}\{\exp(-2it\log\Lambda_n)\} &= (1-2it)^{-\frac{f}{2}} \prod_{l=1}^{2} \exp\Big[\varsigma_l\big\{(1-2it)^{-l} - 1\big\}\Big]\big\{1 + O(p^4 n^{-3})\big\} \\ &= (1-2it)^{-\frac{f}{2}} \big\{1 + V_1(t) + V_2(t) + V_1(t)V_2(t)\big\}\big\{1 + O(p^4 n^{-3})\big\}, \end{aligned}$$

where $V_l(t)$ is defined as in (B.18) on Page 26. Then similarly to the proof in Section B.2, by the inversion property of the characteristic function, we obtain

$$\begin{aligned} &\Pr(-2\log\Lambda_n \le x) \\ &= \Bigg\{ \Pr(\chi_f^2 \le x) + \sum_{v=1}^{\infty} \frac{\varsigma_1^v}{v!} \sum_{w=0}^{v} \binom{v}{w} \Pr(\chi_{f+2w}^2 \le x)(-1)^{v-w} \\ &\quad + \sum_{v=1}^{\infty} \frac{\varsigma_2^v}{v!} \sum_{w=0}^{v} \binom{v}{w} \Pr(\chi_{f+4w}^2 \le x)(-1)^{v-w} \\ &\quad + \sum_{\substack{v_1\ge 1;\ 0\le w_1\le v_1 \\ v_2\ge 1;\ 0\le w_2\le v_2}} \frac{\varsigma_1^{v_1}\varsigma_2^{v_2}}{v_1!v_2!} \binom{v_1}{w_1}\binom{v_2}{w_2} \Pr(\chi_{2f+2w_1+4w_2}^2 \le x)(-1)^{v_1-w_1+v_2-w_2} \Bigg\} \Big\{ 1 + O\Big(\frac{p^4}{n^3}\Big) \Big\}. \end{aligned} \tag{C.36}$$

When $x = \chi_f^2(\alpha)$, by Propositions B.1 and B.2, and $\varsigma_l = O(p^{l+1}/n^l)$, we have

$$\Pr(-2\log\Lambda_n \le x) = \Pr(\chi_f^2 \le x) + \varsigma_1 \big\{ \Pr(\chi_{f+2}^2 \le x) - \Pr(\chi_f^2 \le x) \big\} + o(p^{3/2}/n).$$

Particularly, by Lemma B.2.2,

$$\Pr(\chi_{f+2}^2 \le x) - \Pr(\chi_f^2 \le x) = -\frac{1}{\sqrt{f\pi}} \exp\Big(-\frac{z_\alpha^2}{2}\Big) \Big\{ 1 + O(f^{-1/2}) \Big\},$$

and we compute $\varsigma_1 = (p^2+2p)/(4n)$. In Theorem 2.2, we have $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) Chi-squared approximation with the Bartlett correction. By choosing the Bartlett correction factor ρ as in (C.33), we have $\varsigma_1 = 0$; see, e.g., Section 8.2.4 in Muirhead (2009). Specifically, in this problem,

$\rho = 1-(p+2)/(2n)$, $\rho\xi_{1,1} = \rho\xi_{2,1} = n/2-(p+2)/4$, $\tau_{1,1} = \tau_{2,1} = (p+2)/4$, $\upsilon_{1,1} = -p/2$, $\upsilon_{2,1} = 0$, and then

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)(\rho\times\xi_{1,1})^l}\left\{B_{l+1}\left(-\frac{p-2}{4}\right) - B_{l+1}\left(\frac{p+2}{4}\right)\right\}.$$

We calculate $\varsigma_2 = p(p^2-4)\{48(\rho n)^2\}^{-1}$, $\varsigma_3 = 0$, and $\varsigma_l = O(p^{l+1}n^{-l})$ for $l \geq 4$. Similarly to the proof in Section B.2, when $p^5/n^4 \to 0$, we have

$$\mathrm{E}\{\exp(-2it\rho\log\Lambda_n)\} = (1-2it)^{-\frac{f}{2}}\big\{1+V_2(t)+V_4(t)+V_2(t)V_4(t)\big\}\big\{1+O(p^6/n^5)\big\},$$

and thus

$$\begin{aligned}
&\Pr(-2\rho\log\Lambda_n \leq x) \qquad\qquad (C.37)\\
&= \Bigg\{\Pr(\chi^2_f \leq x) + \sum_{v=1}^{\infty}\frac{\varsigma_2^v}{v!}\sum_{w=0}^{v}\binom{v}{w}\Pr(\chi^2_{f+4w}\leq x)(-1)^{v-w}\\
&\quad + \sum_{v=1}^{\infty}\frac{\varsigma_4^v}{v!}\sum_{w=0}^{v}\binom{v}{w}\Pr(\chi^2_{f+8w}\leq x)(-1)^{v-w}\\
&\quad + \sum_{\substack{v_2\geq 1;\ 0\leq w_2\leq v_2\\ v_4\geq 1;\ 0\leq w_4\leq v_4}}\frac{\varsigma_2^{v_2}\varsigma_4^{v_4}}{v_2!v_4!}\binom{v_2}{w_2}\binom{v_4}{w_4}\Pr(\chi^2_{2f+4w_2+8w_4}\leq x)(-1)^{v_2-w_2+v_4-w_4}\Bigg\}\Big\{1+O\Big(\frac{p^6}{n^5}\Big)\Big\}.
\end{aligned}$$

Note that $\varsigma_2 = \Theta(p^3n^{-2})$ and $\varsigma_4 = \Theta(p^5n^{-4})$. By applying proposition B.1 with $h=2$,

$$\sum_{v=1}^{\infty}\frac{\varsigma_2^v}{v!}\sum_{w=0}^{v}\binom{v}{w}\Pr(\chi^2_{f+4w}\leq x)(-1)^{v-w} = \sum_{v=1}^{\infty}\Big\{O\big(\varsigma_2 p^{-1/2}\big)\Big\}^v = \Theta(p^{5/2}n^{-2}).$$

By applying proposition B.1 with $h=4$, we have

$$\sum_{v=1}^{\infty}\frac{\varsigma_4^v}{v!}\sum_{w=0}^{v}\binom{v}{w}\Pr(\chi^2_{f+8w}\leq x)(-1)^{v-w} = \sum_{v=1}^{\infty}\Big\{O\big(\varsigma_4 p^{-1/2}\big)\Big\}^v = O\big(p^{9/2}n^{-4}\big) = o(p^{5/2}n^{-2}),$$

and

$$\begin{aligned}
&\sum_{\substack{v_2\geq 1;\ 0\leq w_2\leq v_2\\ v_4\geq 1;\ 0\leq w_4\leq v_4}}\frac{\varsigma_2^{v_2}\varsigma_4^{v_4}}{v_2!v_4!}\binom{v_2}{w_2}\binom{v_4}{w_4}\Pr(\chi^2_{2f+4w_2+8w_4}\leq x)(-1)^{v_2-w_2+v_4-w_4}\\
&= \sum_{v_2\geq 1}\big\{O\big(\varsigma_2 p^{-1/2}\big)\big\}^{v_2}\sum_{v_4\geq 1}\frac{\{O(\varsigma_4)\}^{v_4}}{v_4!} = o\big(p^{5/2}n^{-2}\big).
\end{aligned}$$

In summary, by (C.37),

$$\Pr(-2\rho\log\Lambda_n\leq x) = \Pr(\chi^2_f\leq x) + \varsigma_2\Big\{\Pr(\chi^2_{f+4}\leq x) - \Pr(\chi^2_f\leq x)\Big\} + o\big(p^{5/2}n^{-2}\big).$$

Particularly, by Lemma B.2.2,

$$\Pr(\chi^2_{f+4}\leq x) - \Pr(\chi^2_f\leq x) = \ -\frac{2}{\sqrt{f\pi}}\exp\left(-\frac{z_\alpha^2}{2}\right)\Big\{1+O(f^{-1/2})\Big\}.$$

In Theorem 2.2 (I), $\vartheta_2(n,p) = 2\varsigma_2/\sqrt{f}$.

C.3.2. *Proof of Theorem* 2.2 *(II): Testing One-Sample Covariance Matrix* In this problem, by Section 8.3.3 in Muirhead (2009), we know $f=(p+2)(p-1)/2$, and

- $K_1 = p,\ K_2 = 1$;

- $\xi_{1,k} = (n-1)/2, \quad \upsilon_{1,k} = -(k-1)/2$ for $k = 1, \ldots, K_1$;
- $\xi_{2,1} = p(n-1)/2, \ \ \upsilon_{2,1} = 0.$

(i) Chi-squared approximation. Consider $\rho = 1$ and $p^2/n \to 0$. Then $\tau_{1,k} = 0$ for $k = 1, \ldots, K_1$, $\tau_{2,1} = 0$, and

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \left\{ \sum_{k=1}^{p} \left(\frac{2}{n-1} \right)^l B_{l+1}\left(-\frac{k-1}{2} \right) - \frac{2}{p(n-1)} B_{l+1}(0) \right\},$$

which satisfies $\varsigma_l = O(p^{l+2}/n^l)$. By Lemma C.3.1,

$$\mathrm{E}\{\exp(-2it \log \Lambda_n)\} = (1-2it)^{-\frac{f}{2}} \big\{ 1 + V_1(t) + V_2(t) + V_1(t)V_2(t) \big\} \big\{ 1 + O(p^5/n^3) \big\},$$

where $V_l(t)$ is defined as in (B.18). Similarly to Section B.2, by the inversion property of the characteristic functions, and Propositions B.1 and B.2, we obtain (B.22). We calculate

$$\begin{aligned} \varsigma_1 &= \frac{1}{2} \left[\sum_{k=1}^{p} \frac{2}{n-1} \left\{ \left(-\frac{k-1}{2} \right)^2 - \left(-\frac{k-1}{2} \right) + \frac{1}{6} \right\} - \frac{2}{p(n-1)} \times \frac{1}{6} \right] \\ &= \frac{2p^3 + 3p^2 - p - 4/p}{24(n-1)}. \end{aligned}$$

The conclusion then follows by Lemma B.2.2 and $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) Chi-squared approximation with the Bartlett correction. In this problem, consider

$$\rho = 1 - \frac{2p^2 + p + 2}{6p(n-1)},$$

and $p^3/n^2 \to 0$. Then $\tau_{1,k} = (2p^2 + p + 2)/(12p)$ for $k = 1, \ldots, p$, and $\tau_{2,1} = (2p^2 + p + 2)/12$. It follows that

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \left\{ \frac{\rho(n-1)}{2} \right\}^{-l} \left\{ \sum_{k=1}^{p} B_{l+1}\left(\frac{2p^2 + p + 2}{12p} - \frac{k-1}{2} \right) - p^{-l} B_{l+1}\left(\frac{2p^2 + p + 2}{12} \right) \right\}.$$

In particular, we calculate

$$\varsigma_2 = \frac{(p-2)(p-1)(p+2)}{288p^2\rho^2(n-1)^2}(2p^3 + 6p^2 + 3p + 2).$$

Similarly to Section B.2, by the inversion property of the characteristic functions, and Propositions B.1 and B.2, we obtain (B.29). The conclusion then follows by Lemma B.2.2 and $\vartheta_2(n,p) = 2\varsigma_2/\sqrt{f}$.

C.3.3. *Proof of Theorem A.2 (IV): Testing the Equality of Several Mean Vectors* Recall that in Section C.1.3, we show that this testing problem can be viewed as testing the coefficient matrix in multivariate linear regression. Then by Eq. (3) in Section 10.5.3 in Muirhead (2009), we know that in this problem, $f = (k-1)p$, and

- $K_1 = k-1$, $K_2 = k-1$;
- $\xi_{1,j_1} = n/2, \ \upsilon_{1,j_1} = -(j_1 + p)/2, \quad j_1 = 1, \ldots, k-1;$
- $\xi_{2,j_2} = n/2, \ \upsilon_{2,j_2} = -j_2/2, \quad j_2 = 1, \ldots, k-1.$

(i) Chi-squared approximation. It follows that

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \left(\frac{2}{n} \right)^l \left\{ \sum_{j_1=1}^{k-1} B_{l+1}\left(-\frac{j_1 + p}{2} \right) - \sum_{j_2=1}^{k-1} B_{l+1}\left(-\frac{j_2}{2} \right) \right\},$$

which is $O(p^{l+1}n^{-l})$ when k is finite. In particular, we calculate

$$\varsigma_1 = \frac{p(k-1)(p+2+k)}{4n}.$$

Applying similar analysis to that in Section C.3.1, the conclusion follows by $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) Chi-squared approximation with the Bartlett correction. In this problem,

$$\rho = 1 - \frac{1}{2n}(p+k+2).$$

It follows that

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)}\left(\frac{2}{\rho n}\right)^l \left[\sum_{j_1=1}^{k-1} B_{l+1}\left\{\frac{(1-\rho)n-(j_1+p)}{2}\right\} - \sum_{j_2=1}^{k-1} B_{l+1}\left\{\frac{(1-\rho)n-j_2}{2}\right\}\right].$$

We calculate that

$$\varsigma_2 = \frac{(k-1)p(p^2+k^2-2k-4)}{48\rho^2 n^2}.$$

Similarly to Section C.3.1, the conclusion then follows by $\vartheta_2(n,p) = 2\varsigma_2/\sqrt{f}$.

C.3.4. *Proof of Theorem A.2 (V): Testing the Equality of Several Covariance Matrices* In this problem, by Section 8.2.4 in Muirhead (2009), we have $f = p(p+1)(k-1)/2$, and

- $K_1 = kp,\ K_2 = p$;
- $\xi_{1,j_1} = (n_r - 1)/2, \quad j_1 = (r-1)p+1,\ldots,rp, \quad (r = 1,\ldots,k)$;
- $\upsilon_{1,j_1} = -(r-1)/2, \quad j_1 = r, p+r,\ldots,(k-1)p+r, \ (r = 1,\ldots,p)$;
- $\xi_{2,j_2} = (n-k)/2, \upsilon_{2,j_2} = -(j_2-1)/2, \quad j_2 = 1,\ldots,p.$

(i) Chi-squared approximation. Consider $\rho = 1$ and $p^2/n \to 0$. Then

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)}\left[\sum_{r_1=1}^{k}\sum_{r_2=1}^{p}\left(\frac{2}{n_{r_1}-1}\right)^l B_{l+1}\left(-\frac{r_2-1}{2}\right) - \sum_{j=1}^{p}\left(\frac{2}{n-k}\right)^l B_{l+1}\left(-\frac{j-1}{2}\right)\right],$$

which satisfies $\varsigma_l = O(p^{l+2}/n^l)$. Particularly,

$$\varsigma_1 = \left(\sum_{i=1}^{k}\frac{1}{n_i-1} - \frac{1}{n-k}\right)\frac{1}{24}p(2p^2+3p-1).$$

Following similar analysis to that in Section B.2, the conclusion then follows by $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) Chi-squared approximation with the Bartlett correction. In this problem,

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

and we consider $p^3/n^2 \to 0$. In this problem,

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)}\left[\sum_{r_1=1}^{k}\sum_{r_2=1}^{p}\frac{B_{l+1}\{(1-\rho)(n_{r_1}-1)/2-(r_2-1)/2\}}{\{\rho(n_{r_1}-1)/2\}^l} - \sum_{j=1}^{p}\frac{B_{l+1}\{(1-\rho)(n-k)/2-(j-1)/2\}}{\{\rho(n-k)/2\}^l}\right].$$

Note that $(1-\rho)(n-k)$ and $(1-\rho)(n_{r_1}-1)$ are of the order of $\Theta(p)$, $B_{l+1}(\cdot)$ is a polynomial of order $l+1$, and k is finite. Then for $l \geq 2$, $\varsigma_l = O(p^{l+2}/n^l)$. In particular, we calculate

$$\varsigma_2 = \frac{p(p+1)}{48\rho^2}\left[(p-1)(p+2)\left\{\sum_{i_1=1}^{k}\frac{1}{(n_{i_1}-1)^2} - \frac{1}{(n-k)^2}\right\} - 6(k-1)(1-\rho)^2\right].$$

Similarly to Section B.2, the conclusion then follows by $\vartheta_2(n,p) = 2\varsigma_2/\sqrt{f}$.

C.3.5. *Proof of Theorem A.2 (VI): Joint Testing the Equality of Several Mean Vectors and Covariance Matrices* In this problem, by Section 10.8.2 in Muirhead (2009), we have $f = (k-1)p(p+3)/2$, and

- $K_1 = kp$, $K_2 = p$;
- $\xi_{1,j_1} = n_r/2, \quad j_1 = (r-1)p+1, \ \ldots, \ rp, \ \ (r = 1, \ldots, k)$;
- $\upsilon_{1,j_1} = -r/2, \quad j_1 = r, \ p+r, \ldots, \ (k-1)p+r, \ (r = 1, \ldots, p)$;
- $\xi_{2,j_2} = n/2, \ \upsilon_{2,j_2} = -j_2/2, \quad (j_2 = 1, \ldots, p)$.

(i) Chi-squared approximation. Consider $\rho = 1$ and $p^2/n \to 0$. It follows that

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)}\left\{\sum_{r_1=1}^{k}\sum_{r_2=1}^{p}\frac{B_{l+1}(-r_2/2)}{(n_{r_1}/2)^l} - \sum_{j=1}^{p}\frac{B_{l+1}(-j/2)}{(n/2)^l}\right\}.$$

Particularly, we compute

$$\varsigma_1 = \left(\sum_{r=1}^{k}\frac{1}{n_r} - \frac{1}{n}\right)\frac{1}{24}p\left(2p^2+9p+11\right).$$

Following similar analysis to that in Section B.2, the conclusion then follows by $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) Chi-squared approximation with the Bartlett correction. In this problem,

$$\rho = 1 - \left(\sum_{r=1}^{k}\frac{1}{n_r} - \frac{1}{n}\right)\frac{\left(2p^2+9p+11\right)}{6(k-1)(p+3)}.$$

It follows that $\varsigma_1 = 0$ and for $l \geq 2$,

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)}\left\{\sum_{r_1=1}^{k}\sum_{r_2=1}^{p}\frac{B_{l+1}\{(1-\rho)n_{r_1}/2 - r_2/2\}}{(\rho n_{r_1}/2)^l} - \sum_{j=1}^{p}\frac{B_{l+1}\{(1-\rho)n/2 - j/2\}}{(\rho n/2)^l}\right\}.$$

Particularly, we calculate

$$\varsigma_2 = \frac{1}{\rho^2}\left\{\frac{p(p+1)(p+2)(p+3)}{48}\left(\sum_{i=1}^{k}\frac{1}{n_i^2} - \frac{1}{n^2}\right) - \frac{p(k-1)(p+3)}{8}(1-\rho)^2\right\}.$$

Applying similar analysis to that in Section B.2, the conclusion then follows by $\vartheta_2(n,p) = 2\varsigma_2/\sqrt{f}$.

C.3.6. *Proof of Theorem A.5 (VII): Testing Independence between Multiple Vectors* In this problem, by Section 11.2.4 in Muirhead (2009), we have $f = (p^2 - \sum_{j=1}^{k} p_j^2)/2$, and

- $K_1 = p, \ K_2 = p$;
- $\xi_{1,j_1} = n/2, \ \upsilon_{1,j_1} = -j_1/2, \quad j_1 = 1, \ldots, p$;
- $\xi_{2,\ p_1+\ldots+p_{r-1}+j_2} = n/2, \ \upsilon_{2,\ p_1+\ldots+p_{r-1}+j_2} = -j_2/2, \quad r = 1, \ldots, k, \ \ j_2 = 1, \ldots, p_r$.

(i) Chi-squared approximation. Consider $\rho = 1$ and $p^2/n \to 0$. It follows that

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \left\{ \sum_{j_1=1}^{p} \frac{B_{l+1}(-j_1/2)}{(n/2)^l} - \sum_{r=1}^{k} \sum_{j_2=1}^{p_r} \frac{B_{l+1}(-j_2/2)}{(n/2)^l} \right\}.$$

Particularly,

$$\varsigma_1 = \frac{2(p^3 - \sum_{j=1}^{k} p_j^3) + 9(p^2 - \sum_{j=1}^{k} p_j^2)}{24n}.$$

Following similar analysis to that in Section B.2, the conclusion then follows by $\vartheta_1(n,p) = \varsigma_1/\sqrt{f}$.

(ii) Chi-squared approximation with the Bartlett correction. In this problem,

$$\rho = 1 - \frac{2D_{p,3} + 9D_{p,2}}{6nD_{p,2}}$$

where $D_{p,r} = p^r - \sum_{j=1}^{k} p_j^r$. Then

$$\varsigma_l = \frac{(-1)^{l+1}}{l(l+1)} \left\{ \sum_{j_1=1}^{p} \frac{B_{l+1}\{(1-\rho)n/2 - j_1/2\}}{(\rho n/2)^l} - \sum_{r=1}^{k} \sum_{j_2=1}^{p_r} \frac{B_{l+1}\{(1-\rho)n/2 - j_2/2\}}{(\rho n/2)^l} \right\}.$$

In particular, we calculate

$$\varsigma_2 = \frac{1}{(\rho n)^2} \left(\frac{1}{48} D_{p,4} - \frac{5D_{p,2}}{96} - \frac{D_{p,3}^2}{72D_{p,2}} \right).$$

Applying similar analysis to that in Section B.2, the conclusion then follows by $\vartheta_2(n,p) = 2\varsigma_2/\sqrt{f}$.

C.4. *Proofs of Theorems* 2.3*, A.3, & A.6*

In this section, we prove other problems in Theorems 2.3, A.3, & A.6 similarly as in Section B.3. Specifically, we still define $\psi_0(s) = \exp(-s^2/2)$, and we let $\psi_1(s)$ be the characteristic function of $(-2\log\Lambda_n + 2\mu_n)/(2n\sigma_n)$, where Λ_n denotes the corresponding likelihood ratio test statistic, and μ_n and σ_n take the corresponding values given in Theorems 2.3, A.3, & A.6. By the analysis in Section B.3, we know that it suffices to prove the results similar to Lemma B.3.2 on Page 29. In particular, in the following subsections, we prove that under H_0 of each test, when $s = o(\min\{(n/p)^{1/2}, f^{1/6}\})$, the characteristic functions satisfy

$$\log\psi_1(s) - \log\psi_0(s) = O\left(\frac{p}{n} + \frac{1}{\sqrt{f}}\right) s + \left(\frac{1}{p} + \frac{p}{n}\right) O\left(s^2\right) + O\left(\frac{s^3}{\sqrt{f}}\right). \tag{C.38}$$

C.4.1. *Proof of Theorem* 2.3 *(I): Testing One-Sample Mean Vector* Recall that in Section C.1.1, we mention that testing one-sample mean vector can be viewed as testing coefficient vector of a multivariate linear regression model. By Section 10.5.3 in Muirhead (2009), we have

$$\log\psi_1(s) = \log\frac{\Gamma\left\{\frac{1}{2}n(1-ti) - \frac{1}{2}p\right\}}{\Gamma\left\{\frac{1}{2}(n-p)\right\}} - \log\frac{\Gamma\left\{\frac{1}{2}n(1-ti)\right\}}{\Gamma\left(\frac{1}{2}n\right)} + \frac{\mu_n si}{n\sigma_n},$$

where $t = s/(n\sigma_n)$. By (B.7), $t = s/(n\sigma_n) = O(s/\sqrt{f})$. By Lemma D.1.3 (on Page 53),

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

Similarly, we have

$$\log\frac{\Gamma\{\frac{1}{2}n(1-ti)\}}{\Gamma(\frac{1}{2}n)}=\left\{\frac{n(1-ti)}{2}\right\}\log\left\{\frac{n(1-ti)}{2}\right\}+\frac{1}{2}nti-\frac{n}{2}\log\left(\frac{n}{2}\right)+\frac{ti}{2}+O\left(\frac{t}{n}+t^2\right).$$

It follows that

$$\log\psi_1(s)=g_0\bigg(-\frac{nti}{2}\bigg)-g_0(0)+\frac{\mu_n si}{n\sigma_n}+O\left(\frac{pt}{n}+t^2\right),$$

where we define in this subsection that $g_0(z)=\{(n-p)/2+z\}\log\{(n-p)/2+z\}-(n/2+z)\log(n/2+z)$. Following the proof of Lemma D.3.3 (see Section D.3.4 on Page 77), we similarly obtain

$$g_0\bigg(-\frac{nti}{2}\bigg)-g_0(0)=g_0^{(1)}(0)\times\bigg(-\frac{nti}{2}\bigg)-\frac{g_0^{(2)}(0)}{2}\frac{n^2t^2}{4}+O(pt^3),$$

where

$$g_0^{(1)}(0)=\log\left(1-\frac{p}{n}\right),\qquad g_0^{(2)}(0)=\frac{2p}{n(n-p)}.$$

Recall that $2n\sigma_n/\sqrt{2f}\to 1$ by (B.7). Then by Taylor's series and $f=p$,

$$g_0^{(2)}(0)n^2=4n^2\sigma_n^2\left\{1+O\left(\frac{p}{n}\right)\right\}=4n^2\sigma_n^2+O\left(\frac{p^2}{n}\right).$$

Moreover, by Taylor's series, we have $ng_0^{(1)}(0)-2\mu_n=O\left(p/n\right)$. In summary, by $t=s/(n\sigma_n)$ and $n\sigma_n=\Theta(\sqrt{p})$, we obtain

$$\log\psi_1(s)=-\frac{\mu_n si}{n\sigma_n}-\frac{4n^2\sigma_n^2}{2}\frac{s^2}{4(n\sigma_n)^2}+\frac{\mu_n si}{n\sigma_n}+O\bigg(\frac{ps}{n}\bigg)+O\bigg(\frac{p}{n}+\frac{1}{p}\bigg)s^2+O\bigg(\frac{s^3}{\sqrt{p}}\bigg).$$

Then (C.38) is proved.

C.4.2. *Proof of Theorem* 2.3 *(II): Testing One-Sample Covariance Matrix* By Corollary 8.3.6 in Muirhead (2009), we have

$$\log\psi_1(s)=-\frac{p(n-1)ti}{2}\log p+\log\frac{\Gamma_p\{\frac{1}{2}(n-1)(1-ti)\}}{\Gamma_p\{\frac{1}{2}(n-1)\}}+\log\frac{\Gamma\{\frac{1}{2}p(n-1)\}}{\Gamma\{\frac{1}{2}p(n-1)(1-ti)\}}+\mu_n ti.$$

By (B.7) and $f=\Theta(p^2)$, $n\sigma_n=\Theta(p)$. Then as $t=s/(n\sigma_n)$, the conditions in Lemma D.3.1 (on Page 74) are satisfied and we have

$$\begin{aligned}\log\frac{\Gamma_p\{(n-1)(1-ti)/2\}}{\Gamma_p\{(n-1)/2\}}=&-\frac{(n-1)\beta_{n,1}ti}{2}+\frac{(n-1)^2\beta_{n,2}t^2}{4}+\beta_{n,3}\bigg\{-\frac{(n-1)ti}{2}\bigg\}\\&+O\left(\frac{p^2t}{n}\right)+\left(\frac{1}{p}+\frac{p}{n}\right)O\big(p^2t^2\big)+O\big(p^2t^3\big),\end{aligned}$$

where $\beta_{n,1}$, $\beta_{n,2}$, and $\beta_{n,3}(\cdot)$ are defined in Lemma D.3.1. In addition, we can apply Lemma D.1.3 and obtain

$$\begin{aligned}\log\frac{\Gamma\{p(n-1)/2\}}{\Gamma\{p(n-1)(1-ti)/2\}}=&-p\bigg\{\frac{n-1}{2}(1-ti)\bigg\}\log\bigg[p\bigg\{\frac{n-1}{2}(1-ti)\bigg\}\bigg]\\&+\frac{p(n-1)}{2}\log\frac{p(n-1)}{2}-\frac{p(n-1)ti}{2}-ti+O\left(\frac{t}{pn}+t^2\right).\end{aligned}$$

By the definition of $\beta_{n,3}(\cdot)$ in Lemma D.3.1, we have

$$\log\frac{\Gamma\{p(n-1)/2\}}{\Gamma\{p(n-1)/2-pnti/2\}}=-\beta_{n,3}\bigg\{-\frac{(n-1)ti}{2}\bigg\}-\frac{p(n-1)ti(1-\log p)}{2}+O\left(t+t^2\right).$$

Since $\mu_n = (\beta_{n,1} + p)(n-1)/2$, $2n^2\sigma^2 = \beta_{n,2}(n-1)^2$, $t = s/(n\sigma_n)$, and $n\sigma_n = \Theta(p)$,

$$\log \psi_1(s) - \log \psi_0(s) = O\bigg(\frac{p}{n} + \frac{1}{p}\bigg)s + O\bigg(\frac{1}{p} + \frac{p}{n}\bigg)s^2 + O\bigg(\frac{s^3}{p}\bigg).$$

C.4.3. *Proof of Theorem A.3 (IV): Testing the Equality of Several Mean Vectors* By (C.31) and the analysis in Section C.3.3, we have

$$\log \psi_1(s) = \sum_{j=1}^{k-1}\bigg[\log \frac{\Gamma\big\{\frac{1}{2}(n-j-p) - \frac{1}{2}nti\big\}}{\Gamma\big\{\frac{1}{2}(n-j-p)\big\}} - \log \frac{\Gamma\big\{\frac{1}{2}(n-j) - \frac{1}{2}nti\big\}}{\Gamma\big\{\frac{1}{2}(n-j)\big\}}\bigg] + \frac{\mu_n si}{n\sigma_n},$$

where $t = s/(n\sigma_n)$. By Lemma D.1.3,

$$\begin{aligned}\log \frac{\Gamma\big\{\frac{1}{2}(n-j-p) - \frac{1}{2}nti\big\}}{\Gamma\big\{\frac{1}{2}(n-j-p)\big\}} &= \bigg\{\frac{1}{2}(n-j-p) - \frac{1}{2}nti\bigg\}\log\bigg\{\frac{1}{2}(n-j-p) - \frac{1}{2}nti\bigg\} \\ &\quad - \frac{n-j-p}{2}\log\frac{n-j-p}{2} + \frac{nti}{2} + O(t+t^2).\end{aligned}$$

Applying similar analysis, we obtain

$$\log \frac{\Gamma\big\{\frac{1}{2}(n-j-nti)\big\}}{\Gamma\big\{\frac{1}{2}(n-j)\big\}} = \bigg(\frac{n-j-nti}{2}\bigg)\log\bigg(\frac{n-j-nti}{2}\bigg) - \frac{n-j}{2}\log\frac{n-j}{2} + \frac{nti}{2} + O(t+t^2).$$

It follows that $\log \psi_1(s) = \sum_{j=1}^{k-1}\{g_j(nti/2) - g_j(0)\} + \mu_n si/(n\sigma_n) + O(t+t^2)$, where we define in this subsection that

$$g_j(z) = \bigg(\frac{n-j-p}{2} - z\bigg)\log\bigg(\frac{n-j-p}{2} - z\bigg) - \bigg(\frac{n-j}{2} - z\bigg)\log\bigg(\frac{n-j}{2} - z\bigg).$$

Following similar proof to that of Lemma D.3.3 (see Section D.3.4), we obtain

$$\sum_{j=1}^{k-1}\{g_j(nti) - g_j(0)\} = \sum_{j=1}^{k-1} g_j^{(1)}(0)\frac{nti}{2} - \frac{n^2t^2}{8}\sum_{j=1}^{k-1} g_j^{(2)}(0) + O(pt^3), \tag{C.39}$$

where

$$g_j^{(1)}(0) = \log\bigg(\frac{n-j}{2}\bigg) - \log\bigg(\frac{n-j-p}{2}\bigg), \qquad g_j^{(2)}(0) = \frac{2}{n-j-p} - \frac{2}{n-j}.$$

Note that

$$\frac{1}{2}\sum_{j=1}^{k-1} g_j^{(2)}(0) = \sum_{j=1}^{k-1}\frac{p}{(n-j-p)(n-j)} = \frac{p(k-1)}{(n-p-1)n}\left\{1 + O\left(\frac{k}{n}\right)\right\},$$

and

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

Thus $\sum_{j=1}^{k-1} g_j^{(2)}(0)(4\sigma_n^2)^{-1} = 1 + O(n^{-1})$. In addition,

$$\sum_{j=1}^{k-1} g_j^{(1)}(0) = \log\frac{\Gamma(n-1)}{\Gamma(n-k)} - \log\frac{\Gamma(n-p-1)}{\Gamma(n-p-k)}.$$

We then apply Lemma D.1.1 to expand the $\log\Gamma(\cdot)$ function, and calculate

$$\begin{aligned}\sum_{j=1}^{k-1} g_j^{(1)}(0) = & -\left(n-p-k-\frac{1}{2}\right)\left\{\log\left(1-\frac{p}{n-1}\right)-\log\left(1-\frac{p}{n-k}\right)\right\} \\ & -p\log\left(1-\frac{k-1}{n-1}\right)-(k-1)\log\left(1-\frac{p}{n-1}\right)+O(n^{-1}).\end{aligned}$$

Therefore $\sum_{j=1}^{k-1} g_j^{(1)}(0) = -\mu_n/n + O(n^{-1})$. Then by (C.39), $t = s/(n\sigma_n)$, $n\sigma_n = \Theta(f^{1/2})$, and $f = \Theta(p)$, we have

$$\begin{aligned}\log\psi_1(s) &= \left\{-\mu_n/n + O(n^{-1})\right\}nti - \frac{n^2\sigma_n^2 t^2}{2}\left\{1+O(n^{-1})\right\} + \mu_n ti + O\left(t+t^2+pt^3\right) \\ &= -\frac{s^2}{2} + O\left(\frac{1}{\sqrt{f}}\right)s + O\left(\frac{p}{n}+\frac{1}{f}\right)s^2 + O\left(\frac{s^3}{\sqrt{f}}\right).\end{aligned}$$

By $\log\psi_0(s) = -s^2/2$, (C.38) is proved.

C.4.4. *Proof of Theorem A.3 (V): Testing the Equality of Several Covariance Matrices* By (C.31) and the analysis in Section C.3.4, we have

$$\begin{aligned}\log\psi_1(s) = & \log\frac{\Gamma_p\left\{\frac{1}{2}(n-k)\right\}}{\Gamma_p\left\{\frac{1}{2}(n-k)(1-ti)\right\}} + \sum_{j=1}^{k}\log\frac{\Gamma_p\left\{\frac{1}{2}(n_j-1)(1-ti)\right\}}{\Gamma_p\left\{\frac{1}{2}(n_j-1)\right\}} \\ & -p\left\{(n-k)\log(n-k)-\sum_{j=1}^{k}(n_j-1)\log(n_j-1)\right\}\frac{ti}{2}+\frac{\mu_n si}{n\sigma_n},\end{aligned}$$

where $t = s/(n\sigma_n)$. By Lemma D.3.1, we can expand $\log\Gamma_p(\cdot)$ and obtain

$$\log\psi_1(s) = -\mu_n ti - \frac{n^2\sigma_n^2 t^2}{2} + \mu_n ti + R_n(t), \tag{C.40}$$

where the calculations of μ_n and σ_n are similar to that in Section A.5 of Jiang and Qi (2015), and thus the details are skipped here. In (C.40), $R_n(t)$ denotes the remainder term of the expansion. Since Lemma D.3.1 is used, we know that the remainder term satisfies

$$R_n(t) = O\left(\frac{p}{n}\right)s + \left(\frac{1}{p}+\frac{p}{n}\right)s^2 + O\left(\frac{s^3}{p}\right).$$

By $t = s/(n\sigma_n)$ and (C.40), (C.38) is obtained.

C.4.5. *Proof of Theorem A.3 (VI): Joint Testing the Equality of Several Mean Vectors and Covariance Matrices* By Corollary 10.8.3 in Muirhead (2009),

$$\begin{aligned}\log\psi_1(s) = & \log\frac{\Gamma_p\left\{\frac{1}{2}(n-1)\right\}}{\Gamma_p\left\{\frac{1}{2}(n-1)-\frac{1}{2}nti\right\}} + \sum_{j=1}^{k}\log\frac{\Gamma_p\left\{\frac{1}{2}(n_j-1)-\frac{1}{2}n_j ti\right\}}{\Gamma_p\left\{\frac{1}{2}(n_j-1)\right\}} \\ & -p\left(n\log n-\sum_{j=1}^{k}n_j\log n_j\right)\frac{ti}{2}+\frac{\mu_n si}{n\sigma_n},\end{aligned}$$

where $t = s/(n\sigma_n)$. By Lemma D.3.1,

$$\log \frac{\Gamma_p\{\frac{1}{2}(n_j-1)-\frac{1}{2}n_j ti\}}{\Gamma_p\{\frac{1}{2}(n_j-1)\}} = \left[2pn_j + \left(n_j - p - \frac{3}{2}\right) n_j \log\left(1-\frac{p}{n_j-1}\right)\right]\frac{ti}{2} + \left\{\frac{p}{n_j-1} + \log\left(1-\frac{p}{n_j-1}\right)\right\}\frac{n_j^2 t^2}{4} + \varrho_{n_j}(t) + R_n(t), \tag{C.41}$$

where for an integer l, we define

$$\varrho_l(t) = p\left\{\left(\frac{l-1}{2}+\frac{lt}{2}\right)\log\left(\frac{l-1}{2}+\frac{lt}{2}\right) - \frac{l-1}{2}\log\frac{l-1}{2}\right\}, \tag{C.42}$$

and $R_n(t)$ denotes the remainder term and it is of the order of

$$R_n(t) = O\left(\frac{pt}{n}\right) + O\left(\frac{1}{p}+\frac{p}{n}\right)p^2t^2 + O\big(p^2t^3\big). \tag{C.43}$$

In addition, to evaluate $\log\psi_1(s)$, we also use Lemma C.4.1 below.

LEMMA C.4.1. *Under the conditions of Theorem A.*3*, as $p/n \to 0$ and $t = s/(n\sigma_n) = O(s/\sqrt{f})$,*

$$n^2t^2\log\left(1-\frac{p}{n-1}\right) = n^2t^2\log\left(1-\frac{p}{n}\right) + O\left(\frac{p}{n}\right)t^2, \tag{C.44}$$

$$\left\{(n-p-3/2)n\log\left(1-\frac{p}{n-1}\right)\right\}t = \left\{(n-p-3/2)n\log\left(1-\frac{p}{n}\right)\right\}t - pt + O\left(\frac{p}{n}\right)t.$$

Moreover, for $\varrho_l(t)$ defined in (C.42)*, we have*

$$-\varrho_n(t) + \sum_{j=1}^k \varrho_{n_j}(t) = \left(1-k-n\log n + \sum_{j=1}^k n_j\log n_j\right)\frac{tp}{2} + O\left(\frac{pt}{n}+pt^2\right). \tag{C.45}$$

Proof. Please see Section D.3.5 on Page 77. □

By Lemma C.4.1 and the expansions of gamma functions in (C.41), we calculate

$$\begin{aligned}
&\log\psi_1(s) \qquad\qquad (C.46)\\
&= \left\{p - \left(n-p-\frac{3}{2}\right)n\log\left(1-\frac{p}{n}\right) + \sum_{j=1}^k\left(n_j-p-\frac{3}{2}\right)n_j\log\left(1-\frac{p}{n_j-1}\right)\right\}\frac{ti}{2}\\
&\quad - \left(n^2L_{n,p} - \sum_{j=1}^k n_j^2 L_{n_j-1,p}\right)\frac{t^2}{4} - p\left\{(1-k) - n\log n + \sum_{j=1}^k n_j\log n_j\right\}\frac{ti}{2}\\
&\quad - p\left(n\log n - \sum_{j=1}^k n_j\log n_j\right)\frac{ti}{2} + \frac{\mu_n si}{n\sigma_n} + R_n(t),
\end{aligned}$$

where $R_n(t)$ denotes the remainder term of (C.46), which is of the order same as that in (C.43), whereas we mention that the exact value of $R_n(t)$ can change. Then we obtain (C.38) by $t = s/(n\sigma_n)$ and $n\sigma_n = \Theta(f^{1/2})$.

C.4.6. *Proof of Theorem A.*6 *(VII): Testing Independence between Multiple Vectors* By Theorem 11.2.3 in Muirhead (2009), we know

$$\log\psi_1(s) = \log\frac{\Gamma_p\{\frac{1}{2}(n-1)-\frac{1}{2}nti\}}{\Gamma_p\{\frac{1}{2}(n-1)\}} + \sum_{j=1}^k \frac{\Gamma_{p_j}\{\frac{1}{2}(n-1)\}}{\Gamma_{p_j}\{\frac{1}{2}(n-1)-\frac{1}{2}nti\}} + \frac{\mu_n si}{n\sigma_n},$$

where $t = s/(n\sigma_n)$. By Lemma D.3.1, we can expand $\log\Gamma_p(\cdot)$ and obtain

$$\begin{aligned}\log\psi_1(s) = &\left[2p + \left(n - p - \frac{3}{2}\right)L_{n-1,p} - \sum_{j=1}^{k}\left\{2p_j + \left(n - p_j - \frac{3}{2}\right)L_{n-1,p_j}\right\}\right]\frac{nti}{2} \\ &+ \left\{\frac{p}{n-1} + L_{n-1,p} - \sum_{j=1}^{k}\left(\frac{p_j}{n-1} + L_{n-1,p_j}\right)\right\}\frac{n^2t^2}{4} \\ &+ \left(p - \sum_{j=1}^{k} p_j\right)\left\{\frac{n(1-ti)}{2}\log\frac{n(1-ti)}{2} - \frac{n}{2}\log\frac{n}{2}\right\} + \frac{\mu_n si}{n\sigma_n} + R_n(t),\end{aligned}$$

where $R_n(t)$ denotes the remainder term and its order satisfies

$$R_n(t) = O\left(\frac{pt}{n}\right) + O\left(\frac{1}{p} + \frac{p}{n}\right)p^2t^2 + O\big(p^2t^3\big).$$

Then we obtain (C.38) by noticing $p - \sum_{j=1}^{k} p_j = 0$ and $t = O(s/p)$.

D. PROOFS OF ASSISTED LEMMAS

D.1. *Results on Asymptotic Expansions of the Gamma Functions*

In this section, we provide some results on asymptotic expansions of the gamma functions, which are repeatedly used in the proofs. We first give the following Lemma D.1.1 on the expansion of $\log\Gamma(z)$, which also provides the basis for other lemmas below. Lemma D.1.1 and its proof can be found in 12.33 of Whittaker and Watson (1996).

LEMMA D.1.1. *Suppose that a complex number z satisfies Re$(z) \geq \epsilon_1 > 0$ and $|\arg(z)| \leq \pi/2 - \epsilon_2$ with $\epsilon_1 > 0$ and $0 < \epsilon_2 < \pi/4$ being given in advance. When $|z| \to \infty$, and an even integer L, we have*

$$\log\Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + \log\sqrt{2\pi} + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(0)}{l(l+1)z^l} + R_L(z), \tag{D.1}$$

where $B_{l+1}(\cdot)$ represents the Bernoulli polynomial of order $l+1$, and

$$|R_L(z)| = O\left(\frac{|B_{L+2}(0)|}{(L+1)(L+2)|z|^{L+1}}\right).$$

Particularly, we know $B_l(0) = 0$ when l is odd and $l \geq 3$.

In Lemma D.1.1, if we take $L = 2$ and z as a real number, by $B_2(0) = 1/6$, we have

$$\log\Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + \log\sqrt{2\pi} + \frac{1}{12z} + O(z^{-2}). \tag{D.2}$$

Given Lemma D.1.1, we next prove two additional lemmas on asymptotic expansions of the gamma functions.

LEMMA D.1.2. *Suppose a complex number $z + a$ satisfies Re$(z + a) \geq \epsilon_1 > 0$ and $|\arg(z + a)| \leq \pi/2 - \epsilon_2$ with $\epsilon_1 > 0$ and $0 < \epsilon_2 \leq \pi/4$ being given in advance. Assume $|a| \to \infty$ as $|z| \to \infty$ and $|a| = o(|z|)$. For a finite even L, when $|a|^{L+1}/|z|^L \to 0$,*

$$\log\Gamma(z + a) = \left(z + a - \frac{1}{2}\right)\log z - z + \log\sqrt{2\pi} + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(a)}{l(l+1)z^l} + O\left(\frac{|a|^{L+1}}{|z|^L}\right).$$

Proof. Please see Section D.1.1 on Page 53. □

LEMMA D.1.3. *For a real number $x \to \infty$ and a real number $b = o(x)$,*

$$\log \frac{\Gamma(x+bi)}{\Gamma(x)} = (x+bi)\log(x+bi) - x\log x - bi - \frac{bi}{2x} + O\bigg(\frac{b+b^2}{x^2}\bigg),$$

where i denotes the imaginary unit.

Proof. Please see Section D.1.2 on Page 54. □

D.1.1. *Proof of Lemma D.*1.2 *(on Page 52)* By (D.1), for a finite even L, we have

$$\begin{aligned}
&\log\Gamma(z+a) \qquad\qquad (D.3)\\
&= \left(z+a-\frac{1}{2}\right)\log(z+a) - (z+a) + \log\sqrt{2\pi} + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(0)}{l(l+1)(z+a)^l} + O\big(|z+a|^{-L-1}\big)\\
&= \left(z+a-\frac{1}{2}\right)\log z - z + \left(z+a-\frac{1}{2}\right)\log\left(1+\frac{a}{z}\right) - a + \log\sqrt{2\pi}\\
&\quad + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(0)}{l(l+1)z^l}\left(1+\frac{a}{z}\right)^{-l} + O\big(|z+a|^{-L-1}\big).
\end{aligned}$$

By Taylor's expansion,

$$\left(z+a-\frac{1}{2}\right)\log\left(1+\frac{a}{z}\right) - a = \sum_{k=1}^{L-1}\frac{(-1)^{k+1}}{z^k}\left\{\frac{a^{k+1}}{k(k+1)} - \frac{1}{2k}a^k\right\} + O\left(\frac{|a|^{L+1}}{|z|^L}\right). \qquad (D.4)$$

Note that $B_0(0) = 1$ and $B_1(0) = -1/2$. Thus

$$(D.4) = \sum_{k=1}^{L-1}\frac{(-1)^{k+1}}{k(k+1)z^k}\left\{B_0(0)a^{k+1} + \binom{k+1}{1}B_1(0)a^k\right\} + O\left(\frac{|a|^{L+1}}{|z|^L}\right). \qquad (D.5)$$

In addition, by Taylor's expansion, when L is finite,

$$\begin{aligned}
&\sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(0)}{l(l+1)z^l}\left(1+\frac{a}{z}\right)^{-l} \qquad (D.6)\\
&= \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(0)}{l(l+1)z^l}\left\{\sum_{s=0}^{L-1-l}(-1)^s\binom{l+s-1}{s}\frac{a^s}{z^s} + O\left(|a/z|^{L-l}\right)\right\}\\
&= \sum_{k=1}^{L-1}\sum_{t=1}^{k}\frac{(-1)^{k+1}B_{t+1}(0)}{t(t+1)z^k}\frac{(k-1)!}{(t-1)!(k-t)!}a^{k-t} + O\left(|a/z|^L\right)\\
&= \sum_{k=1}^{L-1}\sum_{t=2}^{k+1}\frac{(-1)^{k+1}B_t(0)}{k(k+1)z^k}\binom{k+1}{t}a^{k+1-t} + O\left(|a/z|^L\right).
\end{aligned}$$

Combining (D.3), (D.5), and (D.6), we obtain

$$\begin{aligned}
\log\Gamma(z+a) = &\left(z+a-\frac{1}{2}\right)\log z - z + \log\sqrt{2\pi}\\
&+ \sum_{k=1}^{L-1}\frac{(-1)^{k+1}}{k(k+1)z^k}\left\{\sum_{t=0}^{k+1}\binom{k+1}{t}B_t(0)a^{k+1-t}\right\} + O\left(\frac{|a|^{L+1}}{|z|^L}\right).
\end{aligned}$$

By the property of the Bernoulli polynomials, $B_{k+1}(a) = \sum_{t=0}^{k+1} B_t(0) a^{k+1-t}$; see, e.g., Eq. (13) on Page 21 in Luke (1969). Therefore the lemma is proved.

D.1.2. *Proof of Lemma D.*1.3 *(on Page 53)* By Binet's second formula of the gamma function, it can be obtained that for a complex number z with positive real part, and any integer $L \geq 1$,

$$\log\Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + \log\sqrt{2\pi} + \sum_{l=1}^{L} \frac{B_{2l}(0)}{(2l-1)(2l)z^{2l-1}} + \frac{2(-1)^L}{z^{2L-1}} \int_0^\infty \int_0^t \frac{u^{2L}\mathrm{d}u}{u^2+z^2} \frac{\mathrm{d}t}{e^{2\pi t}-1};$$

please see Page 252 in Whittaker and Watson (1996) for details. Take $L = 1$, and by $B_2(0) = 1/6$, we have

$$\log\Gamma(x) = \left(x - \frac{1}{2}\right)\log x - x + \log\sqrt{2\pi} + \frac{1}{12x} - \frac{2}{x}\int_0^\infty \left(\int_0^t \frac{u^2}{x^2+u^2}\mathrm{d}u\right)\frac{\mathrm{d}t}{e^{2\pi t}-1}.$$

Similarly, we have

$$\begin{aligned}\log\Gamma(x+bi) = &\left(x+bi-\frac{1}{2}\right)\log(x+bi) - (x+bi) + \log\sqrt{2\pi} \\ &+ \frac{1}{12(x+bi)} - \frac{2}{x+bi}\int_0^\infty \left(\int_0^t \frac{u^2}{(x+bi)^2+u^2}\mathrm{d}u\right)\frac{\mathrm{d}t}{e^{2\pi t}-1}.\end{aligned}$$

It follows that

$$\begin{aligned}&\log\frac{\Gamma(x+bi)}{\Gamma(x)} \\ &= (x+bi)\log(x+bi) - x\log x - bi - \frac{1}{2}\log\left(1+\frac{bi}{x}\right) + \frac{1}{12}\left(\frac{1}{x+bi} - \frac{1}{x}\right) + \tilde{R}_2,\end{aligned} \tag{D.7}$$

where

$$\tilde{R}_2 = -2\int_0^\infty \int_0^t \left[\frac{u^2}{(x+bi)\{(x+bi)^2+u^2\}} - \frac{u^2}{x(x^2+u^2)}\right]\mathrm{d}u\frac{\mathrm{d}t}{e^{2\pi t}-1}.$$

To evaluate $\tilde{R}_2$, we note that

$$\begin{aligned}&\frac{u^2}{(x+bi)\{(x+bi)^2+u^2\}} - \frac{u^2}{x(x^2+u^2)} \\ =& -\frac{u^2}{x^3} \times \frac{2b_x i - b_x^2 + b_x i\{(1+b_x i)^2+u_x^2\}}{(1+b_x i)\{(1+b_x i)^2+u_x^2\}(1+u_x^2)} \\ =& -\frac{u^2 b_x}{x^3(1+b_x i)(1+u_x^2)} \times \left[\frac{2i-b_x}{(1+b_x i)^2+u_x^2} + i\right],\end{aligned}$$

where for easy presentation, we let $b_x = b/x$ and $u_x = u/x$. Since $b = o(x)$, $|(1+b_x i)^{-1}|$ is bounded. Moreover, we also know $(1+u_x^2)^{-1}$ and $|\{(1+b_x i)^2+u_x^2\}^{-1}|$ are bounded. It follows that there exists a constant C such that

$$|\tilde{R}_2| \leq \frac{Cb_x}{x^3}\int_0^\infty \left(\int_0^t u^2\mathrm{d}u\right)\frac{\mathrm{d}t}{e^{2\pi t}-1} = O\left(\frac{b}{x^4}\right),$$

where we use $\int_0^\infty t^3(e^{2\pi t}-1)^{-1}\mathrm{d}t$ is a constant; see 7.2 in Whittaker and Watson (1996). Lemma D.1.3 is then obtained by (D.7) and

$$\log\left(1+\frac{bi}{x}\right) = \frac{bi}{x} + O\left(\frac{b^2}{x^2}\right), \qquad \frac{1}{x+bi} - \frac{1}{x} = O\left(\frac{b}{x^2}\right).$$

D.2. *Lemmas for Theorems* 2.2*, A.*2 *&* A.5

D.2.1. *Proof of Lemma B.*2.1 *(on Page 25)* By (B.1), we can write

$$\log \mathrm{E}\{\exp(-2it\eta \log \Lambda_n)\} = G_1 + G_2 + G_3,$$

where in this subsection, we let

$$G_1 = -i\eta npt \log\left(\frac{2e}{n}\right), \quad G_2 = -\frac{np}{2}(1-2i\eta t)\log(1-2i\eta t),$$
$$G_3 = \log\Gamma_p\left(\frac{n-1}{2} - \eta nit\right) - \log\Gamma_p\Big(\frac{n-1}{2}\Big).$$

By the property of multivariate gamma function; see, e.g., Theorem 2.1.12 in Muirhead (2009), we obtain

$$\begin{aligned} G_3 &= \sum_{j=1}^{p} \log\Gamma\left\{\frac{n}{2}(1-2\eta it) - \frac{j}{2}\right\} - \sum_{j=1}^{p}\log\Gamma\left(\frac{n}{2}-\frac{j}{2}\right) \\ &= \sum_{j=1}^{p}\left[\log\Gamma\left\{\frac{\eta n}{2}(1-2it) + \frac{n(1-\eta)-j}{2}\right\} - \log\Gamma\left\{\frac{\eta n}{2} + \frac{n(1-\eta)-j}{2}\right\}\right]. \end{aligned}$$

We first examine G_3. When $\eta = 1$ or $\eta = \rho$, for $1 \le j \le p$, $n(1-\eta) - j = O(p)$ and $\eta n = \Theta(n)$. As $p = o(n)$, $|\{n(1-\eta)-j\}\{\eta n(1-2it)\}^{-1}| = O(p/n) = o(1)$. Then we can apply Lemma D.1.2 on Page 52, and obtain

$$\begin{aligned} &\log\Gamma\left\{\frac{\eta n}{2}(1-2it) + \frac{n(1-\eta)-j}{2}\right\} \\ &= \left\{\frac{\eta n}{2}(1-2it) + \frac{n(1-\eta)-j-1}{2}\right\}\log\left\{\frac{\eta n}{2}(1-2it)\right\} - \frac{\eta n}{2}(1-2it) + \log\sqrt{2\pi} \\ &\quad + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}}{l(l+1)}B_{l+1}\left\{\frac{n(1-\eta)}{2} - \frac{j}{2}\right\}\left\{\frac{\eta n}{2}(1-2it)\right\}^{-l} + O\left(\frac{p^{L+1}}{n^L}\right), \end{aligned}$$

and

$$\begin{aligned} &\log\Gamma\left\{\frac{\eta n}{2} + \frac{n(1-\eta)-j}{2}\right\} \\ &= \left\{\frac{\eta n}{2} + \frac{n(1-\eta)-j-1}{2}\right\}\log\frac{\eta n}{2} - \frac{\eta n}{2} + \log\sqrt{2\pi} \\ &\quad + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}}{l(l+1)}B_{l+1}\left\{\frac{n(1-\eta)}{2} - \frac{j}{2}\right\}\left(\frac{\eta n}{2}\right)^{-l} + O\left(\frac{p^{L+1}}{n^L}\right). \end{aligned}$$

It follows that

$$\begin{aligned} G_3 = &-\eta pnti\log\left(\frac{n}{2e}\right) - p\eta nit\log\eta + \frac{pn}{2}(1-2i\eta t)\log(1-2it) - \sum_{j=1}^{p}\frac{j+1}{2}\log(1-2it) \\ &+ \sum_{l=1}^{L-1}\frac{(-1)^{l+1}}{l(l+1)}\sum_{j=1}^{p}B_{l+1}\left\{\frac{n(1-\eta)}{2} - \frac{j}{2}\right\}\left(\frac{\eta n}{2}\right)^{-l}\left\{(1-2it)^{-l} - 1\right\} + O\left(\frac{p^{L+2}}{n^L}\right). \end{aligned}$$

We next examine G_2. By $1-2i\eta t=\eta(1-2it)+1-\eta$, and Taylor's expansion,

$$\begin{aligned}&(1-2i\eta t)\log(1-2i\eta t)\\ =&\{\eta(1-2it)+1-\eta\}\log\{\eta(1-2it)\}\\ &+1-\eta+(1-\eta)\sum_{l=1}^{L-1}\frac{(-1)^{l+1}}{l(l+1)}\Big(\frac{1-\eta}{\eta}\Big)^{l}(1-2it)^{-l}+O\big\{(1-\eta)^{L+1}\big\}.\end{aligned}$$

As $\log(1)=(1-2i\eta\times 0)\log(1-2i\eta\times 0)=0$, by applying Taylor's expansion similarly as above,

$$\begin{aligned}&(1-2i\eta t)\log(1-2i\eta t)-\log(1)\\ =&\ -2i\eta t\log\eta(1-2it)+\log(1-2it)\\ &+(1-\eta)\sum_{l=1}^{L-1}\frac{(-1)^{l+1}}{l(l+1)}\Big(\frac{1-\eta}{\eta}\Big)^{l}\big\{(1-2it)^{-l}-1\big\}+O\{(1-\eta)^{L+1}\}.\end{aligned}$$

As $(1-\eta)/\eta=\{(1-\eta)n/2\}/(\eta n/2)$,

$$\begin{aligned}G_2=&-\sum_{l=1}^{L-1}\frac{(-1)^{l+1}}{l(l+1)}\sum_{j=1}^{p}\left\{\frac{(1-\eta)n}{2}\right\}^{l+1}\Big(\frac{\eta n}{2}\Big)^{-l}\big\{(1-2it)^{-l}-1\big\}\\ &+i\eta npt\log\eta(1-2it)-\frac{np}{2}\log(1-2it)+O\big\{(1-\eta)^{L+1}pn\big\}.\end{aligned}$$

In summary, as $1-\eta=O(p/n)$ when $\eta=1$ or ρ, we have

$$G_1+G_2+G_3=-\sum_{j=1}^{p}\frac{j+1}{2}\log(1-2it)+\sum_{l=1}^{L-1}\varsigma_l\big\{(1-2it)^{-l}-1\big\}+O\Big(\frac{p^{L+2}}{n^L}\Big),$$

where

$$\varsigma_l=\frac{(-1)^{l+1}}{l(l+1)}\sum_{j=1}^{p}\left[B_{l+1}\left\{\frac{(1-\eta)n}{2}-\frac{j}{2}\right\}-\left\{\frac{(1-\eta)n}{2}\right\}^{l+1}\right]\Big(\frac{\eta n}{2}\Big)^{-l}.$$

Particularly, as $B_{l+1}(\cdot)$ is a polynomial of order $l+1$ and $(1-\eta)n=O(p)$, we have $\varsigma_l=O(p^{l+2}n^{-l})$.

D.2.2. *Notation of the finite difference and computation rules* In the following, we prove Propositions B.1 and B.2 and Lemma B.2.2 based on the calculus of the finite difference. To facilitate the proofs, we introduce some notation. Given x, define a function with respect to the degrees of freedom f as $F_x(f)=P(\chi_f^2\le x)$. Let Δ_{2h} represent a forward difference operator with step $2h$, that is, $\Delta_{2h}(F_x,f)=F_x(f+2h)-F_x(f)$. For an integer $v\ge 1$, it follows that the v-th order forward difference is

$$\Delta_{2h}^{v}(F_x,f)=\sum_{w=0}^{v}\binom{v}{w}(-1)^{v-w}F(f+2hw),$$

where $\Delta_{2h}^{1}(F_x,f)=\Delta_{2h}(F_x,f)$. Particularly, when $h=1$, we have

$$\Delta_{2}^{v}(F_x,f)=\sum_{w=0}^{v}\binom{v}{w}(-1)^{v-w}P(\chi_{f+2w}^2\le x);$$

when $h=2$,

$$\Delta_{4}^{v}(F_x,f)=\sum_{w=0}^{v}\binom{v}{w}(-1)^{v-w}P(\chi_{f+4w}^2\le x).$$

In the following proofs, we use several rules of the finite difference operator listed in Lemmas D.2.1–D.2.3 below, which can be found in Section 3.7 of Zwillinger (2002).

LEMMA D.2.1 (LEIBNIZ RULE). *For two functions $F(f)$ and $G(f)$, and two positive integers v and h,*

$$\Delta_h^v(FG, f) = \sum_{w=0}^{v} \binom{v}{w} \Delta_h^w(F, f)\Delta_h^{v-w}(G, f + hw).$$

LEMMA D.2.2 (LINEARITY RULE). *For two constants C_1 and C_2, two functions $F(f)$ and $G(f)$, and two positive integers v and h, the linear combination $C_1F(f) + C_2G(f)$ satisfies*

$$\Delta_h^v(C_1F + C_2G, f) = C_1\Delta_h^v(F) + C_2\Delta_h^v(G).$$

LEMMA D.2.3. *For a function $F(f)$ and positive integers v_1, v_2, h_1, and h_2,*

$$\Delta_{h_2}^{v_2}\Delta_{h_1}^{v_1}(F, f) = \Delta_{h_1}^{v_1}\Delta_{h_2}^{v_2}(F, f) = \Delta_{h_2}^{v_2}\Delta_{h_1}^{v_1-1}\{\Delta_{h_1}(F, f)\} = \Delta_{h_1}^{v_1}\Delta_{h_2}^{v_2-1}\{\Delta_{h_2}(F, f)\}.$$

Based on the notation and lemmas on the finite difference, we first prove Lemma B.2.2 in Section D.2.3, and then use Lemma B.2.2 to prove Propositions B.1 and B.2 in Sections D.2.4 and D.2.5, respectively.

D.2.3. *Proof of Lemma B.2.2 (on Page 27)* We prove (B.24) in Lemma B.2.2 from the cumulative distribution function of the chi-squared distribution. In particular, by the probability density of χ_f^2, we have

$$\Pr\left(\chi_f^2 \le x\right) = \frac{\gamma(f/2, x/2)}{\Gamma(f/2)},$$

where $\gamma(m, x)$ is the lower incomplete gamma function defined as $\gamma(m, x) = \int_0^x t^{m-1}e^{-t}\mathrm{d}t$, and $\Gamma(m)$ is the gamma function defined as $\Gamma(m) = \int_0^\infty t^{m-1}e^{-t}\mathrm{d}t$; see, e.g., Section 6.2 in Press et al. (1992). Thus for an integer h,

$$\Delta_{2h}^1(F_x, f) = \frac{\Gamma\left(\frac{f}{2}\right)\gamma\left(\frac{f}{2} + h, \frac{x}{2}\right) - \Gamma\left(\frac{f}{2} + h\right)\gamma\left(\frac{f}{2}, \frac{x}{2}\right)}{\Gamma\left(\frac{f}{2} + h\right)\Gamma\left(\frac{f}{2}\right)},$$

where $\Delta_{2h}^1(F_x, f) = \Pr(\chi_{f+2h}^2 \le x) - \Pr(\chi_f^2 \le x)$ following the notation in Section D.2.2. By integration by parts, we have

$$\Gamma(m + 1) = m\Gamma(m), \quad \text{and then} \quad \Gamma(m + h) = \prod_{k=1}^{h}(m + h - k)\Gamma(m). \tag{D.8}$$

Similarly, we have $\gamma(m + 1, x) = m\gamma(m, x) - x^m e^{-x}$, and then

$$\gamma(m + h, x) = \prod_{k=1}^{h}(m + h - k)\gamma(m, x) - \sum_{k=1}^{h}\prod_{t=1}^{k-1}(m + h - t)x^{m+h-k}e^{-x};$$

this recurrence formulas can also be found in Sections 6.3 and 6.5 in Abramowitz and Stegun (1970). It follows that

$$\Delta_{2h}^1(F_x, f) = -\frac{\sum_{k=1}^{h}\prod_{t=1}^{k-1}(f/2 + h - t)(x/2)^{\frac{f}{2}+h-k}e^{-x/2}}{\prod_{t=1}^{h}(f/2 + h - t) \times \Gamma(f/2)} = -\sum_{k=1}^{h}\frac{(x/2)^{\frac{f}{2}+h-k}\,e^{-x/2}}{\Gamma\left(f/2 + h - k + 1\right)}.$$

Therefore (B.24) is proved.

We next prove (B.25) in Lemma B.2.2 based on (B.24) by discussing $h \in \{1, 2, 3, 4\}$, respectively. *(1).* We first consider $h = 1$. Under this case,

$$\Delta_2^1(F_x, f) = -\frac{(x/2)^{f/2}e^{-x/2}}{\Gamma(f/2 + 1)}. \tag{D.9}$$

By (D.2), as $f \to \infty$, $\Gamma(f/2) = (f/2)^{f/2-1/2}e^{-f/2}\sqrt{2\pi}\{1+O(f^{-1})\}$. Moreover, by $\Gamma(f/2+1) = \Gamma(f/2)f/2$, we have

$$\begin{aligned}\frac{1}{\Gamma(f/2+1)}(x/2)^{f/2}e^{-x/2} &= \frac{1}{\sqrt{f\pi}}\left(\frac{x}{f}\right)^{f/2}\exp\left\{\frac{f-x}{2}+O(f^{-1})\right\}\\ &= \frac{1}{\sqrt{f\pi}}\exp\left\{\frac{f-x}{2}+\frac{f}{2}\log\left(1+\frac{x-f}{f}\right)+O(f^{-1})\right\}.\end{aligned}$$

When $x = \chi^2_f(\alpha)$, we have $x = f+\sqrt{2f}\{z_\alpha + O(f^{-1/2})\}$ by (B.6). Then by Taylor's series,

$$\Delta_2^1(F_x,f) = \frac{1}{\sqrt{f\pi}}\exp\left\{-\frac{(x-f)^2}{4f}+O(f^{-1/2})\right\} = \frac{1}{\sqrt{f\pi}}\exp\left(-\frac{z_\alpha^2}{2}\right)\{1+O(f^{-1/2})\}.$$

(2). When $h=2$, by (B.24), (D.8), and $x = f+\sqrt{2f}\{z_\alpha+O(f^{-1/2})\}$, we have

$$\Delta_4^1(F_x,f) = -\frac{\frac{x}{2}}{\frac{f}{2}+1}\times\Delta_2^1(F_x,f)+\Delta_2^1(F_x,f) = -\frac{2}{\sqrt{f\pi}}\exp\left(-\frac{z_\alpha^2}{2}\right)\{1+O(f^{-1/2})\}.$$

(3). When $h=3$, similarly by (B.24), (D.8), and $x = f+\sqrt{2f}\{z_\alpha+O(f^{-1/2})\}$, we have

$$\Delta_6^1(F_x,f) = -\frac{(\frac{x}{2})^2}{(\frac{f}{2}+2)(\frac{f}{2}+1)}\Delta_2^1(F_x,f)+\Delta_4^1(F_x,f) = -\frac{3}{\sqrt{f\pi}}\exp\left(-\frac{z_\alpha^2}{2}\right)\{1+O(f^{-1/2})\}.$$

(4). When $h=4$, similarly by (B.24), (D.8), and $x = f+\sqrt{2f}\{z_\alpha+O(f^{-1/2})\}$, we have

$$\begin{aligned}\Delta_8^1(F_x,f) &= -\frac{(\frac{x}{2})^3}{(\frac{f}{2}+3)(\frac{f}{2}+2)(\frac{f}{2}+1)}\Delta_2^1(F_x,f)+\Delta_6^1(F_x,f)\\ &= -\frac{4}{\sqrt{f\pi}}\exp\left(-\frac{z_\alpha^2}{2}\right)\{1+O(f^{-1/2})\}.\end{aligned}$$

In summary, (B.25) is proved.

D.2.4. *Proof of Proposition B.*1 *(on Page 27)* We prove Proposition B.1 based on the notation in Section D.2.2 and Lemma B.2.2, which is proved in Section D.2.3 above. Particularly, we write the left hand side of (B.21) as $\Delta_{2h}^v(F_x,f)$ below. By (B.25), we know (B.21) holds for $v=1$ and $h\in\{1,2,3,4\}$. We next prove (B.21) for $v\geq 2$ when $h\in\{1,2,3,4\}$, respectively.

(Part I) Proof for $h=1$. When $v=2$, by (B.24), we have

$$\Delta_2^2(F_x,f) = -\frac{1}{\Gamma(\frac{f}{2}+2)}\left(\frac{x}{2}\right)^{\frac{f}{2}+1}e^{-x/2}+\frac{1}{\Gamma(\frac{f}{2}+1)}\left(\frac{x}{2}\right)^{\frac{f}{2}}e^{-x/2}.$$

Then we can write $\Delta_2^2(F_x,f) = A_1(f)Q_1(f)$, where we define

$$Q_1(f) = \Delta_2^1(F_x,f), \quad \text{and} \quad A_1(f) = x/(f+2)-1. \tag{D.10}$$

Note that $Q_1(f) = O(f^{-1/2})$ by (B.25), and $A_1(f) = O(f^{-1/2})$ by (B.6) when $x=\chi^2_f(\alpha)$. Therefore, (B.21) holds for $h=1$ and $v=2$.

We next prove (B.21) for $h=1$ and $v>2$ by the mathematical induction. Assume that there exists some constant C such that uniformly for integers $1\leq k\leq v-1$,

$$\Delta_2^k(F_x,f) = O(k!C^kf^{-k/2}),$$

that is, uniformly for integers $1\leq k\leq v-1$,

$$\Delta_2^{k-1}(Q_1,f) = O(k!C^kf^{-k/2}). \tag{D.11}$$

We next prove $\Delta_2^v(F_x, f) = O(v!C^v f^{-v/2})$. By the definition of $Q_1(f)$ and $A_1(f)$, we have

$$\Delta_2^v(F_x, f) = \Delta_2^{v-1}(Q_1, f) = \Delta_2^{v-2}(A_1 Q_1, f).$$

By Lemma D.2.1,

$$\Delta_2^{v-2}(A_1 Q_1, f) = \sum_{w=0}^{v-2} \binom{v-2}{w} \Delta_2^w(A_1, f) \Delta_2^{v-2-w}(Q_1, f + 2w). \tag{D.12}$$

To evaluate (D.12), by (D.11), for $0 \le w \le v-2$, we have

$$\Delta_2^{v-2-w}(Q_1, f + 2w) = O\big\{(v-w-1)!C^{v-w-1} f^{-(v-w-1)/2}\big\}.$$

In addition, to evaluate $\Delta_2^w(A_1, f)$ in (D.12), we use the following Lemma D.2.4.

LEMMA D.2.4. *When $x = \chi_f^2(\alpha)$ and $f \to \infty$, $A_1(f) = \sqrt{2} z_\alpha f^{-1/2}\{1 + O(f^{-1})\}$, and for any integer $w \ge 1$,*

$$\Delta_2^w(A_1, f) = x \times (-1)^w 2^w w! \frac{1}{\prod_{k=1}^{w+1}(f + 2k)}. \tag{D.13}$$

Thus there exists a constant C such that (D.13) *is of the order of $O(w!C^w f^{-w})$ as $f \to \infty$ uniformly for $w \ge 1$.*

Proof. Please see Section D.2.6 on Page 67. □

By Lemma D.2.4, (D.12) gives that as $f \to \infty$,

$$\begin{aligned}
&\Delta_2^{v-2}(A_1 Q_1, f) \qquad\qquad (D.14)\\
&= O(f^{-1/2}) \times O\big\{(v-1)!C^{v-1} f^{-(v-1)/2}\big\} \\
&\quad + \sum_{w=1}^{v-2} \binom{v-2}{w} O\big(w!C^w f^{-w}\big) \times O\big\{(v-w-1)!C^{v-w-1} f^{-(v-w-1)/2}\big\} \\
&= (v-1)!C^{v-1} O(f^{-v/2}) + \sum_{w=1}^{v-2} (v-2)!(v-w-1) C^{v-1} O\{f^{-(w-1)/2} \times f^{-v/2}\} \\
&= O(v!C^v f^{-v/2}),
\end{aligned}$$

where in the last equation, we use $v - w - 1 \le v - 1$ and $O\{f^{-(w-1)/2}\} = O(1)$ when $w \ge 1$. We note that there exists a constant C such that the last equation in (D.14) holds uniformly for $v \ge 1$. In summary, we obtain (B.21) for $h = 1$.

(Part II) Proof for $h = 2$. By (B.24), (D.9) and (D.10),

$$\Delta_4^1(F_x, f) = Q_2(f) + Q_1(f), \tag{D.15}$$

where we define

$$Q_2(f) = -\frac{1}{\Gamma(\frac{f}{2} + 2)} \left(\frac{x}{2}\right)^{\frac{f}{2}+1} e^{-x/2}.$$

Then by (D.15) and Lemma D.2.2, we have

$$\Delta_4^v(F_x, f) = \Delta_4^{v-1}(Q_2, f) + \Delta_4^{v-1}(Q_1, f).$$

Therefore, to prove (B.21) for $h = 2$, it suffices to prove

$$\begin{aligned}
\Delta_4^{v-1}(Q_1, f) &= O(v!C^v f^{-v/2}), \\
\Delta_4^{v-1}(Q_2, f) &= O(v!C^v f^{-v/2}). \qquad\qquad (D.16)
\end{aligned}$$

As $Q_1(f) = Q_2(f-2)$, it suffices to prove (D.16), and we next use the mathematical induction. Note that (D.16) holds for $v = 1$ since $\Delta_4^0(Q_2, f) = Q_2(f) = O(f^{-1/2})$ by the proof of (B.25). In addition, for $v = 2$, we have

$$\Delta_4^1(Q_2, f) = Q_2(f+4) - Q_2(f) = A_2(f)Q_2(f), \tag{D.17}$$

where

$$A_2(f) = \frac{(\frac{x}{2})^2}{(\frac{f}{2}+3)(\frac{f}{2}+2)} - 1.$$

Note that $Q_2(f) = O(f^{-1/2})$, and when $x = \chi_f^2(\alpha)$, we have $A_2(f) = O(f^{-1/2})$ by (B.6). Therefore, $\Delta_4^1(Q_2, f) = O(f^{-1})$, i.e., (D.16) holds for $v = 2$. For $v \geq 3$, we next use the mathematical induction, where we assume for integers $0 \leq w \leq v-2$,

$$\Delta_4^w(Q_2, f) = O\{(w+1)!C^{w+1}f^{-(w+1)/2}\}, \tag{D.18}$$

and prove (D.16). By (D.17), $\Delta_4^{v-1}(Q_2, f) = \Delta_4^{v-2}(A_2Q_2, f)$. Then by Lemma D.2.1,

$$\Delta_4^{v-2}(A_2Q_2, f) = \sum_{w=0}^{v-2} \binom{v-2}{w} \Delta_4^w(A_2, f)\Delta_4^{v-2-w}(Q_2, f+4w). \tag{D.19}$$

We next prove (D.16) by (D.18), (D.19) and the following Lemma D.2.5.

LEMMA D.2.5. *When $x = \chi_f^2(\alpha)$, $A_2(f) = 2\sqrt{2}z_\alpha f^{-1/2}\{1 + O(f^{-1/2})\}$. Moreover, there exists a constant C such that uniformly for any integer $w \geq 1$,*

$$\Delta_4^w(A_2, f) = O\bigg\{(w+1)!C^w \prod_{t=1}^{w}(f+2t)^{-1}\bigg\}. \tag{D.20}$$

Proof. Please see Section D.2.7 on Page 67. □

By Lemma D.2.5 and (D.18), we have

$$\begin{aligned}
\text{(D.19)} =& O(f^{-1/2}) \times O\big\{(v-1)!C^{v-1}f^{-(v-1)/2}\big\} \\
&+ \sum_{w=1}^{v-2} \binom{v-2}{w} O\bigg\{(w+1)!C^w \prod_{t=1}^{w}(f+2t)^{-1}(v-1-w)!C^{(v-1-w)}f^{-\frac{v-1-w}{2}}\bigg\} \\
=& O\big\{(v-1)!C^{v-1}f^{-\frac{v}{2}}\big\} + \sum_{w=1}^{v-2} O\big\{(v-2)!(v-1-w)C^{v-1}f^{-\frac{v}{2}}\big\}\frac{(w+1)f^{\frac{w+1}{2}}}{\prod_{t=1}^{w}(f+2t)}.
\end{aligned} \tag{D.21}$$

To evaluate (D.21), we note that when $w = 1$ and 2, $(w+1)f^{(w+1)/2}\{\prod_{t=1}^{w}(f+2t)\}^{-1} = O(f^{(1-w)/2})$; when $w \geq 3$, as $f \to \infty$,

$$\frac{(w+1)f^{(w+1)/2}}{\prod_{t=1}^{w}(f+2t)} \leq \frac{w+1}{2w}\frac{f^{(w+1)/2}}{f^{w-1}} = O(1)$$

uniformly over $w \geq 3$. Moreover, by $\sum_{w=1}^{v-2}(v-2)!(v-1-w) \leq v!$, we obtain (D.19) $= O(v!C^v f^{-v/2})$.

(Part III) Proof for $h = 3$. By (B.24),

$$\Delta_6^1(F_x, f) = Q_3(f) + Q_2(f) + Q_1(f), \tag{D.22}$$

where we define

$$Q_3(f) = -\frac{1}{\Gamma(\frac{f}{2}+3)}\left(\frac{x}{2}\right)^{\frac{f}{2}+2} e^{-x/2}.$$

Then by (D.22) and Lemma D.2.2,

$$\Delta_6^v(F_x, f) = \Delta_6^{v-1}(Q_3, f) + \Delta_6^{v-1}(Q_2, f) + \Delta_6^{v-1}(Q_1, f).$$

Since $Q_2(f) = Q_3(f-2)$ and $Q_1(f) = Q_3(f-4)$, it suffices to prove

$$\Delta_6^{v-1}(Q_3, f) = O(v! C^v f^{-v/2}). \tag{D.23}$$

We next prove (D.23) by the mathematical induction. Note that (D.23) holds for $v = 1$ since $\Delta_6^0(Q_3, f) = Q_3(f) = O(f^{-1/2})$ by the proof of (B.25) in Section D.2.3. In addition, for $v = 2$,

$$\Delta_6^1(Q_3, f) = Q_3(f+6) - Q_3(f) = A_3(f) Q_3(f), \tag{D.24}$$

where

$$A_3(f) = \prod_{k=1}^{3} A_{3,k}(f) - 1, \qquad A_{3,k}(f) = \frac{x}{f + 4 + 2k}.$$

Note that $A_3(f) = O(f^{-1/2})$ when $x = \chi_f^2(\alpha)$ by (B.6). Moreover, as $Q_3(f) = O(f^{-1/2})$, $\Delta_6^1(Q_3, f) = O(f^{-1})$, i.e., (D.23) holds for $v = 2$. For $v \geq 3$, we next use the mathematical induction, where we assume for integers $0 \leq w \leq v-2$,

$$\Delta_6^w(Q_3, f) = O\{(w+1)! C^{w+1} f^{-(w+1)/2}\}, \tag{D.25}$$

and prove (D.23). By (D.24), $\Delta_6^{v-1}(Q_3, f) = \Delta_6^{v-2}(A_3 Q_3, f)$. Then by Lemma D.2.1,

$$\Delta_6^{v-2}(A_3 Q_3, f) = \sum_{w=0}^{v-2} \binom{v-2}{w} \Delta_6^w(A_3, f) \Delta_6^{v-2-w}(Q_3, f + 6w). \tag{D.26}$$

We next prove (D.26) by (D.25) and the following Lemma D.2.6.

LEMMA D.2.6. *When $x = \chi_f^2(\alpha)$, $A_3(f) = 3\sqrt{2} z_\alpha f^{-1/2}\{1 + O(f^{-1/2})\}$. Moreover, there exists a constant C such that uniformly for any integer $w \geq 1$,*

$$\Delta_6^w(A_3, f) = O\bigg\{(w+2)! C^w \prod_{t=1}^{w} (f + 2t)^{-1}\bigg\}.$$

Proof. Please see Section D.2.8 on Page 68. □

Then by (D.25) and Lemma D.2.6,

$$\begin{aligned}
\text{(D.26)} = {} & O(f^{-1/2}) \times O\big\{(v-1)! C^{v-1} f^{-(v-1)/2}\big\} \\
& + \sum_{w=1}^{v-2} \binom{v-2}{w} O\bigg\{(w+2)! C^w \prod_{t=1}^{w} (f+2t)^{-1} (v-1-w)! C^{v-1-w} f^{-(v-1-w)/2}\bigg\} \\
= {} & O\big\{(v-1)! C^{v-1} f^{-v/2}\big\} + \sum_{w=1}^{v-2} O\big\{(v-1)! C^v f^{-v/2}\big\} \frac{(w+2)(w+1) f^{(w+1)/2}}{\prod_{t=1}^{w}(f+2t)}.
\end{aligned}$$

Note that when $w \leq 4$, $(w+2)(w+1) f^{(w+1)/2} \prod_{t=1}^{w} (f+2t)^{-1} = O\{f^{(1-w)/2}\}$; when $w \geq 5$,

$$\frac{(w+2)(w+1) f^{(w+1)/2}}{\prod_{t=1}^{w}(f+2t)} \leq \frac{(w+2)(w+1)}{w(w-1)} f^{(5-w)/2} = O(1)$$

as $f \to \infty$ uniformly over $w \geq 5$. It follows that (D.26) $= O(v! C^v f^{-v})$ and thus (D.23) is proved.

(Part IV) Proof for $h = 4$. By (B.24),

$$\Delta_8^1(F_x, f) = Q_4(f) + Q_3(f) + Q_2(f) + Q_1(f), \tag{D.27}$$

where we define

$$Q_4(f) = -\frac{1}{\Gamma(\frac{f}{2}+4)}\left(\frac{x}{2}\right)^{\frac{f}{2}+3} e^{-x/2}.$$

Then by (D.27) and Lemma D.2.1,

$$\Delta_8^v(F_x, f) = \Delta_8^{v-1}(Q_4, f) + \Delta_8^{v-1}(Q_3, f) + \Delta_8^{v-1}(Q_2, f) + \Delta_8^{v-1}(Q_1, f).$$

Since $Q_3(f) = Q_4(f-2)$, $Q_2(f) = Q_4(f-4)$, and $Q_1(f) = Q_4(f-6)$, it suffices to prove

$$\Delta_8^{v-1}(Q_4, f) = O(v! C^v f^{-v/2}). \tag{D.28}$$

We next prove (D.28) by the mathematical induction. Note that (D.28) holds for $v = 1$ since $\Delta_8^0(Q_4, f) = Q_4(f) = O(f^{-1/2})$ by the proof of (B.25) in Section D.2.3. In addition, for $v = 2$, we have

$$\Delta_8^1(Q_4, f) = Q_4(f+8) - Q_4(f) = A_4(f)Q_4(f), \tag{D.29}$$

where

$$A_4(f) = \prod_{k=1}^{4} A_{4,k}(f) - 1, \qquad A_{4,k}(f) = \frac{x}{f+6+2k}.$$

Note that $A_4(f) = O(f^{-1/2})$ as $x = f + \sqrt{2f}\{z_\alpha + O(f^{-1/2})\}$. Moreover, as $Q_4(f) = O(f^{-1/2})$, $\Delta_8^1(Q_4, f) = O(f^{-1})$, i.e., (D.28) holds for $v = 2$. For $v \geq 3$, we next use the mathematical induction, where we assume for integers $0 \leq w \leq v-2$,

$$\Delta_8^w(Q_4, f) = O\{(w+1)! C^{w+1} f^{-(w+1)/2}\}, \tag{D.30}$$

and prove (D.28). By (D.29), $\Delta_8^{v-1}(Q_4, f) = \Delta_8^{v-2}(A_4 Q_4, f)$. Then by Lemma D.2.1,

$$\Delta_8^{v-2}(A_4 Q_4, f) = \sum_{w=0}^{v-2} \binom{v-2}{w} \Delta_8^w(A_4, f) \Delta_8^{v-2-w}(Q_4, f+8w). \tag{D.31}$$

We next prove (D.31) by (D.30), (D.31) and the following Lemma D.2.7.

LEMMA D.2.7. *When $x = \chi_f^2(\alpha)$, $A_4(f) = 4\sqrt{2} z_\alpha f^{-1/2}\{1 + O(f^{-1/2})\}$. Moreover, there exists a constant C such that as $f \to \infty$,*

$$\Delta_8^w(A_4, f) = O\left\{(w+3)! C^w \prod_{t=1}^{w} (f+2t)^{-1}\right\}$$

holds uniformly for any integer $w \geq 1$

Proof. Please see Section D.2.9 on Page 68. □

Then by (D.30) and Lemma D.2.7,

$$\begin{aligned}
\text{(D.31)} = {} & O(f^{-1/2}) \times O\{(v-1)! C^{v-1} f^{-(v-1)/2}\} \\
& + \sum_{w=1}^{v-2} \binom{v-2}{w} O\left\{(w+3)! C^w \prod_{t=1}^{w} (f+2t)^{-1} (v-1-w)! C^{v-1-w} f^{-(v-1-w)/2}\right\} \\
= {} & O\{(v-1)! C^{v-1} f^{-v/2}\} + \sum_{w=1}^{v-2} O\{(v-1)! C^v f^{-v/2}\} \frac{(w+3)(w+2)(w+1) f^{\frac{w+1}{2}}}{\prod_{t=1}^{w}(f+2t)}.
\end{aligned}$$

Note that when $w \le 6$, $(w+3)(w+2)(w+1)f^{(w+1)/2}\prod_{t=1}^{w}(f+2t)^{-1} = O\{f^{(1-w)/2}\}$; when $w \ge 7$, as $f \to \infty$,

$$\frac{(w+3)(w+2)(w+1)f^{(w+1)/2}}{\prod_{t=1}^{w}(f+2t)} \le \frac{(w+3)(w+2)(w+1)}{w(w-1)(w-2)} f^{(7-w)/2} = O(1)$$

holds uniformly over $w \ge 7$. It follows that (D.31) $= O(v!C^v f^{-v})$ and thus (D.28) is proved.

D.2.5. *Proof of Proposition B.2 (on Page 27)* Similar to the proof of Proposition B.1 in Section D.2.5, we prove Proposition B.2 using the notation in Section D.2.2 and Lemma B.2.2. We next discuss $(h_1, h_2) = (1, 2)$ and $(h_1, h_2) = (2, 3)$ in *(Part I)* and *(Part II)* below, respectively.

(Part I) Proof for $h_1 = 1$ and $h_2 = 2$. Based on the notation in Section D.2.2, it is equivalent to prove that there exists some constant C such that when $x = \chi^2_f(\alpha)$, as $f \to \infty$,

$$\Delta_4^{v_2}\Delta_2^{v_1}(F_x, f) = O\big\{v_1!v_2!C^{v_1+v_2}f^{-(v_1+v_2)/2}\big\}, \tag{D.32}$$

uniformly for integers $v_1, v_2 \ge 1$.

When $v_1 = 0$ or $v_2 = 0$, (D.32) holds by Proposition B.1. When $v_1 = v_2 = 1$, by (D.9), we have

$$\Delta_4^1\Delta_2^1(F_x, f) = -\frac{1}{\Gamma(\frac{f}{2}+3)}\Big(\frac{x}{2}\Big)^{\frac{f}{2}+2}e^{-x/2} + \frac{1}{\Gamma(\frac{f}{2}+1)}\Big(\frac{x}{2}\Big)^{\frac{f}{2}}e^{-x/2} = D_{2,4}(f)\Delta_2^1(F_x, f),$$

where

$$D_{2,4}(f) = \frac{x^2}{(f+4)(f+2)} - 1. \tag{D.33}$$

As $D_{2,4}(f) = O(f^{-1/2})$ and $\Delta_2^1(F_x, f) = O(f^{-1/2})$, (D.32) holds for $v_1 = v_2 = 1$. We next prove (D.32) by the mathematical induction. Particularly, we assume for integers $s_1 \le v_1$ and $s_2 \le v_2$,

$$\Delta_4^{s_2}\Delta_2^{s_1}(F_x, f) = O\big\{s_1!s_2!C^{s_1+s_2}f^{-(s_1+s_2)/2}\big\}, \tag{D.34}$$

and prove that (D.34) also holds for $(s_1, s_2) = (v_1+1, v_2)$ and $(s_1, s_2) = (v_1, v_2+1)$, i.e., $\Delta_4^{v_2}\Delta_2^{v_1+1}(F_x, f)$ and $\Delta_4^{v_2+1}\Delta_2^{v_1}(F_x, f)$, respectively.

Step I.1. $\Delta_4^{v_2}\Delta_2^{v_1+1}(F_x, f)$. Recall that we define $Q_1(f) = \Delta_2^1(F_x, f)$. It follows that (D.34) gives that for integers $s_1 \le v_1 - 1$ and $s_2 \le v_2$

$$\Delta_4^{s_2}\Delta_2^{s_1}(Q_1, f) = O\big\{(s_1+1)!s_2!C^{s_1+s_2+1}f^{-(s_1+s_2+1)/2}\big\}. \tag{D.35}$$

It is then equivalent to prove that (D.35) holds for $(s_1, s_2) = (v_1, v_2)$, i.e., $\Delta_4^{v_2}\Delta_2^{v_1}(Q_1, f)$. By $\Delta_2^1(Q_1, f) = A_1(f)Q_1(f)$, (see the definitions in (D.10)), and Lemmas D.2.1 and D.2.2,

$$\begin{aligned}
&\Delta_4^{v_2}\Delta_2^{v_1}(Q_1, f)\\
&= \sum_{w_1=0}^{v_1-1}\binom{v_1-1}{w_1}\Delta_4^{v_2}\Big\{\Delta_2^{w_1}(A_1, f)\Delta_2^{v_1-1-w_1}(Q_1, f+2w_1)\Big\}\\
&= \sum_{w_1=0}^{v_1-1}\binom{v_1-1}{w_1}\sum_{w_2=0}^{v_2}\binom{v_2}{w_2}\Delta_4^{w_2}\Delta_2^{w_1}(A_1, f)\Delta_4^{v_2-w_2}\Delta_2^{v_1-1-w_1}(Q_1, f+2w_1+4w_2).
\end{aligned} \tag{D.36}$$

To evaluate (D.36), we use the following Lemma D.2.8.

LEMMA D.2.8. *For two integers w_1 and w_2 satisfying $w_1 + w_2 \ge 1$, there exists some constant C such that as $f \to \infty$,*

$$\Delta_4^{w_2}\Delta_2^{w_1}(A_1, f) = (w_1+w_2)!O\left(C^{w_1+w_2}\prod_{k=1}^{w_1+w_2}\frac{1}{f+2k}\right)$$

uniformly over $w_1 + w_2 \geq 1$.

Proof. Please see Section D.2.10 on Page 69. □

By Lemma D.2.8 and the assumption (D.35), we have

$$\begin{aligned}(\text{D.36}) = &\sum_{w_1=0}^{v_1-1}\binom{v_1-1}{w_1}\sum_{w_2=0}^{v_2}\binom{v_2}{w_2}(w_1+w_2)!O\left(\prod_{k=1}^{w_1+w_2}\frac{1}{f+2k}\right)C^{v_1+v_2+1}\\ &\times(v_2-w_2)!(v_1-w_1)!O\big\{(f+2w_1+4w_2)^{-(v_1-w_1+v_2-w_2)/2}\big\}\\ =&\sum_{w_1=0}^{v_1-1}(v_1-1)!(v_1-w_1)\sum_{w_2=0}^{v_2}v_2!C^{v_1+v_2+1}O\{f^{-(v_1+v_2+1)/2}\}\\ &\times\frac{(w_1+w_2)!}{w_1!w_2!}\prod_{k=1}^{w_1+w_2}\frac{1}{f+2k}\times f^{(w_1+w_2+1)/2}.\end{aligned}$$

We next use the following Lemma D.2.9.

LEMMA D.2.9. *For integers* w_1, w_2, *and* f,

$$\frac{(w_1+w_2)!}{w_1!w_2!}\prod_{k=1}^{w_1+w_2}\frac{1}{f+2k}\times f^{(w_1+w_2+1)/2} = O\big\{2^{-(w_1+w_2-1)/2}\big\}.$$

Proof. Please see Section D.2.11 on Page 70. □

It follows that by Lemma D.2.9,

$$\begin{aligned}(\text{D.36}) &= \sum_{w_1=0}^{v_1-1}(v_1-1)!(v_1-w_1)\sum_{w_2=0}^{v_2}v_2!\frac{1}{(\sqrt{2})^{w_1+w_2-1}}C^{v_1+v_2+1}O\{f^{-(v_1+v_2+1)/2}\}\\ &= O\{v_1!v_2!C^{v_1+v_2+1}f^{-(v_1+v_2+1)/2}\},\end{aligned}\tag{D.37}$$

which is $O\{(v_1+1)!v_2!C^{v_1+v_2+1}f^{-(v_1+v_2+1)/2}\}$ as $v_1 < v_1+1$. Therefore, we obtain $\Delta_4^{v_2}\Delta_2^{v_1}(Q_1,f) = O\{(v_1+1)!v_2!C^{v_1+v_2+1}f^{-(v_1+v_2+1)/2}\}$.

Step I.2. $\Delta_4^{v_2+1}\Delta_2^{v_1}(F_x,f)$. By (D.15),

$$\Delta_4^{v_2+1}\Delta_2^{v_1}(F_x,f) = \Delta_2^{v_1}\Delta_4^{v_2+1}(F_x,f) = \Delta_2^{v_1}\Delta_4^{v_2}(Q_2,f) + \Delta_4^{v_2}\Delta_2^{v_1+1}(F_x,f).$$

By (D.37), we have $\Delta_4^{v_2}\Delta_2^{v_1+1}(F_x,f) = O\{v_1!v_2!C^{v_1+v_2+1}f^{-(v_1+v_2+1)/2}\}$. Therefore, it remains to prove $\Delta_2^{v_1}\Delta_4^{v_2}(Q_2,f) = O\{v_1!(v_2+1)!C^{v_1+v_2+1}f^{-(v_1+v_2+1)/2}\}$. By (D.17) and Lemma D.2.1,

$$\begin{aligned}&\Delta_2^{v_1}\Delta_4^{v_2}(Q_2,f)\\ =&\Delta_2^{v_1}\left\{\sum_{w_2=0}^{v_2-1}\binom{v_2-1}{w_2}\Delta_4^{w_2}(A_2,f)\Delta_4^{v_2-1-w_2}(Q_2,f+4w_2)\right\}\\ =&\sum_{w_2=0}^{v_2-1}\binom{v_2-1}{w_2}\sum_{w_1=0}^{v_1}\binom{v_1}{w_1}\Delta_2^{w_1}\Delta_4^{w_2}(A_2,f)\Delta_2^{v_1-w_1}\Delta_4^{v_2-1-w_2}(Q_2,f+4w_2+2w_1).\end{aligned}\tag{D.38}$$

To evaluate (D.38) through the mathematical induction, by (D.15) and (D.34), we can assume that or integers $s_1 \leq v_1$ and $s_2 \leq v_2 - 1$,

$$\Delta_2^{s_1}\Delta_4^{s_2}(Q_2,f) = O\big\{s_1!(s_2+1)!C^{s_1+s_2+1}f^{-(s_1+s_2+1)/2}\big\}.\tag{D.39}$$

In addition, we use the following Lemma D.2.10.

LEMMA D.2.10. *For two integers w_1 and w_2 satisfying $w_1 + w_2 \geq 1$,*

$$\Delta_4^{w_2}\Delta_2^{w_1}(A_2, f) = (w_1 + w_2 + 1)!O\left(C^{w_1+w_2+1}\prod_{k=1}^{w_1+w_2}\frac{1}{f+2k}\right).$$

Proof. Please see Section D.2.12 on Page 71. □

Combining (D.39) and Lemma D.2.10, we obtain $\Delta_2^{v_1}\Delta_4^{v_2}(Q_2, f) = O\{v_1!v_2!C^{v_1+v_2+1}f^{-(v_1+v_2+1)/2}\}$ similarly to (D.37) in *Step I.1*. As $v_2 < v_2 + 1$, we have $\Delta_2^{v_1}\Delta_4^{v_2}(Q_2, f) = O\{v_1!(v_2+1)!C^{v_1+v_2+1} \times f^{-(v_1+v_2+1)/2}\}$.

(Part II) Proof for $h_1 = 2$ and $h_2 = 3$. In this part, we prove

$$\Delta_6^{v_2}\Delta_4^{v_1}(F_x, f) = O\big\{v_1!v_2!C^{v_1+v_2}f^{-(v_1+v_2)/2}\big\}, \tag{D.40}$$

as $f \to \infty$ and uniformly for integers $v_1, v_2 \geq 1$.

When $v_1 = 0$ or $v_2 = 0$, (D.40) holds by Proposition B.1. When $v_1 = v_2 = 1$, note that $\Delta_4^1(F_x, f) = Q_1(f) + Q_2(f)$ by (D.15). Then we have $\Delta_6^1\Delta_4^1(F_x, f) = \Delta_6^1(Q_1, f) + \Delta_6^1(Q_2, f)$. Particularly,

$$\Delta_6^1(Q_1, f) = D_{2,6}(f)Q_1(f), \quad D_{2,6}(f) = \frac{x^3}{(f+6)(f+4)(f+2)} - 1; \tag{D.41}$$
$$\Delta_6^1(Q_2, f) = D_{4,6}(f)Q_2(f), \quad D_{4,6}(f) = \frac{x^3}{(f+8)(f+6)(f+4)} - 1.$$

By the proof of (B.25), $Q_1(f) = O(f^{-1/2})$ and $Q_2(f) = O(f^{-1/2})$. In addition, for $x = \chi_f^2(\alpha)$, by (B.6), $D_{2,6}(f) = O(f^{-1/2})$ and $D_{4,6}(f) = O(f^{-1/2})$. Therefore, (D.40) holds for $v_1 = 1$ and $v_2 = 1$. When $v_1 > 1$ or $v_2 > 1$, by (D.15),

$$\Delta_6^{v_2}\Delta_4^{v_1}(F_x, f) = \Delta_6^{v_2}\Delta_4^{v_1-1}(Q_1, f) + \Delta_6^{v_2}\Delta_4^{v_1-1}(Q_2, f).$$

It suffices to prove

$$\Delta_6^{v_2}\Delta_4^{v_1-1}(Q_1, f) = O\big\{v_1!v_2!C^{v_1+v_2}f^{-(v_1+v_2)/2}\big\}, \tag{D.42}$$
$$\Delta_6^{v_2}\Delta_4^{v_1-1}(Q_2, f) = O\big\{v_1!v_2!C^{v_1+v_2}f^{-(v_1+v_2)/2}\big\}. \tag{D.43}$$

We next prove (D.42) and (D.43) by the mathematical induction, respectively.

First, to prove (D.42), we apply the mathematical induction considering increasing v_1 and v_2 in the following *Step II.1* and *Step II.2*, respectively.

Step II.1. We assume for $0 \leq s_1 \leq v_1 - 2$ and $0 \leq s_2 \leq v_2$,

$$\Delta_6^{s_2}\Delta_4^{s_1}(Q_1, f) = O\{(s_1+1)!s_2!C^{s_1+s_2+1}f^{-(s_1+s_2+1)/2}\}, \tag{D.44}$$

and then prove (D.42). Note that $\Delta_4^1(Q_1, f) = D_{2,4}(f)Q_1(f)$, where $D_{2,4}(f)$ is defined in (D.33). Then by the Leibniz rule in Lemma D.2.1,

$$\begin{aligned}
&\Delta_6^{v_2}\Delta_4^{v_1-1}(Q_1, f)\\
&= \Delta_6^{v_2}\Delta_4^{v_1-2}(D_{2,4}Q_1, f)\\
&= \sum_{k_2=0}^{v_2}\sum_{k_1=0}^{v_1-2}\binom{v_1-2}{k_1}\binom{v_2}{k_2}\Delta_6^{k_2}\Delta_4^{k_1}(D_{2,4}, f) \times \Delta_6^{v_2-k_2}\Delta_4^{v_1-2-k_1}(Q_1, f+4k_1+6k_2).
\end{aligned} \tag{D.45}$$

To evaluate (D.45), we use the following Lemma D.2.11.

LEMMA D.2.11. *For integers $k_1 + k_2 \geq 1$, there exists some constant C such that*

$$\Delta_6^{k_2}\Delta_4^{k_1}(D_{2,4}, f) = (k_1 + k_2 + 1)!O\left(C^{k_1+k_2}\prod_{t=1}^{k_1+k_2}\frac{1}{f+2t}\right),$$

as $f \to \infty$ and uniformly over $k_1 + k_2 \geq 1$.

Proof. Please see Section D.2.13 on Page 72. □

Then applying similar analysis to that of (D.36) and (D.37) in *Part I* above, we obtain (D.42) by the assumption (D.44) and Lemma D.2.11.

Step II.2. We assume for $0 \leq s_1 \leq v_1 - 1$ and $0 \leq s_2 \leq v_2 - 1$, (D.44) holds, and then prove (D.42). By (D.41) and the Leibniz rule in Lemma D.2.1,

$$\begin{aligned}
&\Delta_6^{v_2}\Delta_4^{v_1-1}(Q_1, f) \qquad (D.46)\\
&= \Delta_4^{v_1-1}\Delta_6^{v_2-1}(D_{2,6}Q_1, f)\\
&= \sum_{k_1=0}^{v_1-1}\sum_{k_2=0}^{v_2-1}\binom{v_1-1}{k_1}\binom{v_2-1}{k_2}\Delta_6^{k_2}\Delta_4^{k_1}(D_{2,6}, f) \times \Delta_6^{v_2-1-k_2}\Delta_4^{v_1-1-k_1}(Q_1, f + 4k_1 + 6k_2).
\end{aligned}$$

Similarly to the analysis of (D.45), we use the following Lemma D.2.12 to evaluate (D.46).

LEMMA D.2.12. *For integers $k_1 + k_2 \geq 1$, there exists a constant C such that*

$$\Delta_6^{k_2}\Delta_4^{k_1}(D_{2,6}, f) = (k_1 + k_2 + 2)!O\left(C^{k_1+k_2}\prod_{t=1}^{k_1+k_2}\frac{1}{f+2t}\right),$$

as $f \to \infty$ and uniformly over $k_1 + k_2 \geq 1$.

Proof. Please see Section D.2.14 on Page 72. □

Since we assume (D.44) holds for $0 \leq s_1 \leq v_1 - 1$ and $0 \leq s_2 \leq v_2 - 1$, then by Lemma D.2.12,

$$\begin{aligned}
(D.46) = &\sum_{k_1=0}^{v_1-1}\sum_{k_2=0}^{v_2-1}\binom{v_1-1}{k_1}\binom{v_2-1}{k_2}(k_1 + k_2 + 2)!(v_2 - 1 - k_2)!(v_1 - k_1)!\\
&\times C^{v_1+v_2}f^{-(v_1+v_2)/2}O\left(f^{-(k_1+k_2+1)/2}\prod_{t=1}^{k_1+k_2}\frac{1}{f+2t}\right)\\
= &\, C^{v_1+v_2}f^{-(v_1+v_2)/2}(v_1-1)!(v_2-1)!\sum_{k_2=0}^{v_2}\sum_{k_1=0}^{v_1-1}(v_1 - k_1)\\
&\times \frac{(k_1 + k_2 + 2)!}{k_1!k_2!}O\left(f^{-(k_1+k_2+1)/2}\prod_{t=1}^{k_1+k_2}\frac{1}{f+2t}\right).
\end{aligned}$$

We next use the following Lemma D.2.13 to evaluate (D.46).

LEMMA D.2.13. *For integers $k_1 + k_2 \geq 1$, as $f \to \infty$,*

$$\frac{(k_1 + k_2 + 2)!}{k_1!k_2!}O\left\{f^{-(k_1+k_2+1)/2}\prod_{t=1}^{k_1+k_2}\frac{1}{f+2t}\right\} = O\{2^{-(k_1+k_2-1)/2}\}.$$

Proof. Please see Section D.2.15 on Page 72. □

Then by Lemma D.2.13, we obtain $\Delta_6^{v_2}\Delta_4^{v_1-1}(Q_1, f) = O\{v_1!v_2!C^{v_1+v_2}f^{-(v_1+v_2)/2}\}$ similarly to (D.37). In summary, combining *Step II.1* and *Step II.2*, we finish the proof of (D.42).

Second, to prove (D.43), we can use the mathematical induction similarly to the proof of (D.42). The analysis would be very similar and the details are thus skipped.

D.2.6. *Proof of Lemma D.2.4 (on Page 59)* When $x = \chi^2_f(\alpha)$, by (B.6), we have $x = f + \sqrt{2f}\{z_\alpha + O(f^{-1/2})\}$, and then $A_1(f) = \sqrt{2}z_\alpha f^{-1/2}\{1 + O(f^{-1})\}$. We next prove (D.13) by the mathematical induction. For $w = 1$, we compute

$$\Delta_2^1(A_1, f) = A_1(f+2) - A_1(f) = -x \times 2 \times \frac{1}{(f+2)(f+4)}.$$

Therefore (D.13) holds when $w = 1$. We next assume (D.13) holds, and prove the conclusion holds for $\Delta_2^{w+1}(A_1, f)$. Particularly,

$$\begin{aligned}\Delta_2^{w+1}(A_1, f) &= x \times (-1)^w 2^w w! \left\{ \frac{1}{\prod_{k=2}^{w+2}(f+2k)} - \frac{1}{\prod_{k=1}^{w+1}(f+2k)} \right\} \\ &= x \times (-1)^{w+1} 2^{w+1}(w+1)! \frac{1}{\prod_{k=1}^{w+2}(f+2k)}.\end{aligned}$$

In summary, Lemma D.2.4 is proved.

D.2.7. *Proof of Lemma D.2.5 (on Page 60)* When $x = \chi^2_f(\alpha)$, by (B.6), we have $x = f + \sqrt{2f}\{z_\alpha + O(f^{-1/2})\}$, and then $A_2(f) = 2\sqrt{2}z_\alpha f^{-1/2}\{1 + O(f^{-1})\}$. We next prove (D.20). Note that we can write $A_2(f) = A_{2,1}(f)A_{2,2}(f) - 1$, where we define

$$A_{2,1}(f) = \frac{x}{f+4} \quad \text{and} \quad A_{2,2}(f) = \frac{x}{f+6}.$$

By Lemmas D.2.1 and D.2.2, when $w \geq 1$,

$$\Delta_4^w(A_2, f) = \Delta_4^w(A_{2,1}A_{2,2}, f) = \sum_{k=0}^{w} \binom{w}{k} \Delta_4^k(A_{2,1}, f)\Delta_4^{w-k}(A_{2,2}, f+4k). \tag{D.47}$$

To prove (D.47) $= O(w!C^w f^{-w})$, we next evaluate $\Delta_4^k(A_{2,1}, f)$ and $\Delta_4^{w-k}(A_{2,2}, f+4k)$.

In particular, we prove that

$$\Delta_4^k(A_{2,1}, f) = (-1)^k 4^k k! x \times \frac{1}{\prod_{t=1}^{k+1}(f+4t)} \tag{D.48}$$

by the mathematical induction. When $k = 1$,

$$\Delta_4^1(A_{2,1}, f) = \frac{x}{f+8} - \frac{x}{f+4} = \frac{x \times (-4)}{(f+4)(f+8)}.$$

Thus (D.48) holds for $k = 1$. We next assume (D.48) holds and prove the conclusion for $\Delta_4^{k+1}(A_{2,1}, f)$. Specifically,

$$\begin{aligned}\Delta_4^{k+1}(A_{2,1}, f) &= (-1)^k 4^k k! x \left\{ \frac{1}{\prod_{t=2}^{k+2}(f+4t)} - \frac{1}{\prod_{t=1}^{k+1}(f+4t)} \right\} \\ &= (-1)^{k+1} 4^{k+1}(k+1)! x \frac{1}{\prod_{t=1}^{k+2}(f+4t)}.\end{aligned}$$

In summary, (D.48) is proved. Moreover, as $A_{2,2}(f) = A_{2,1}(f+2)$, we have

$$\Delta_4^k(A_{2,2}, f) = \Delta_4^k(A_{2,1}, f+2) = (-1)^k 4^k k! x \frac{1}{\prod_{t=1}^{k+1}(f+2+4t)}.$$

It follows that $\Delta_4^{w-k}(A_{2,2}, f+4k) = (-1)^{w-k}4^{w-k}(w-k)!x\{\prod_{t=k+1}^{w+1}(f+2+4t)\}^{-1}$. Then by (D.47), there exists a constant C such that

$$\left|\Delta_4^w(A_{2,1}A_{2,2}, f)\right| = \left|\sum_{k=0}^{w}\binom{w}{k}\frac{(-4)^w k!(w-k)!x^2}{\prod_{t=1}^{k+1}(f+4t)\prod_{t=k+1}^{w+1}(f+2+4t)}\right| \le w!C^w\sum_{k=0}^{w}\frac{x^2}{\prod_{t=1}^{w+2}(f+2t)}.$$

As $x = \chi_f^2(\alpha) = O(f)$, we obtain that (D.20) holds as $f \to \infty$ and uniformly for any integer $w \ge 1$.

D.2.8. *Proof of Lemma D.2.6 (on Page 61)* When $x = \chi_f^2(\alpha)$, by (B.6), we have $x = f + \sqrt{2f}\{z_\alpha + O(f^{-1/2})\}$, and then $A_3(f) = 3\sqrt{2}z_\alpha f^{-1/2}\{1 + O(f^{-1/2})\}$. We next consider $\Delta_6^w(A_3, f)$ for $w \ge 1$. As $A_3(f) = \prod_{l=1}^3 A_{3,l}(f) - 1$,

$$\Delta_6^w(A_3, f) = \sum_{k_1=0}^{w}\sum_{k_2=0}^{k_1}\binom{k_1}{k_2}\binom{w}{k_1}\Delta_6^{k_2}(A_{3,1}, f)\Delta_6^{k_1-k_2}(A_{3,2}, f+6k_2)\Delta_6^{w-k_1}(A_{3,3}, f+6k_1).$$

Similarly to the proofs of Lemma D.2.4 in Section D.2.6, for $A_{3,l}(f)$, $l \in \{1,2,3\}$, we can obtain that for any integer $w \ge 1$ and $l \in \{1,2,3\}$

$$\Delta_6^w(A_{3,l}, f) = (-6)^w w!x \times \frac{1}{\prod_{t=0}^{w}(f+4+2l+6t)}.$$

It follows that

$$\begin{aligned}\Delta_6^w(A_3, f) = &\sum_{k_1=0}^{w}\sum_{k_2=0}^{k_1}\binom{k_1}{k_2}\binom{w}{k_1}(-6)^w k_2!(k_1-k_2)!(w-k_1)! \\ &\times x^3\left\{\prod_{t=1}^{k_2+1}(f+6t)\prod_{t=k_2+1}^{k_1+1}(f+6t+2)\prod_{t=k_1+1}^{w+1}(f+6t+4)\right\}^{-1}.\end{aligned}$$

As $\binom{k_1}{k_2}\binom{w}{k_1}k_2!(k_1-k_2)!(w-k_1)! = w!$, $\sum_{k_1=0}^{w}\sum_{k_2=0}^{k_1}1 \le (w+1)^2$, and $x = \chi_f^2(\alpha) = O(f)$, there exists a constant C such that as $f \to \infty$ and uniformly over $w \ge 1$,

$$\Delta_6^w(A_3, f) = O\left\{(w+2)!C^w\prod_{t=1}^{w}(f+2t)^{-1}\right\}.$$

D.2.9. *Proof of Lemma D.2.7 (on Page 62)* When $x = \chi_f^2(\alpha)$, by (B.6), we have $x = f + \sqrt{2f}\{z_\alpha + O(f^{-1/2})\}$, and then $A_4(f) = 4\sqrt{2}z_\alpha f^{-1/2}\{1 + O(f^{-1/2})\}$. We next prove the conclusion for $w \ge 1$. As $A_4(f) = \prod_{l=1}^4 A_{4,l}(f) - 1$,

$$\begin{aligned}\Delta_8^w(A_4, f) = &\sum_{k_1=0}^{w}\sum_{k_2=0}^{k_1}\sum_{k_3=0}^{k_2}\binom{w}{k_1}\binom{k_1}{k_2}\binom{k_2}{k_3}\Delta_8^{k_3}(A_{4,1}, f) \times \Delta_8^{k_2-k_3}(A_{4,2}, f+8k_3) \\ &\times \Delta_8^{k_1-k_2}(A_{4,3}, f+8k_2) \times \Delta_8^{w-k_1}(A_{4,4}, f+8k_1).\end{aligned}$$

Similarly to the proof of Lemma D.2.4 in Section D.2.6, for $A_{4,l}(f)$, $l \in \{1,2,3,4\}$, we can obtain that for any integer $w \ge 1$,

$$\Delta_8^w(A_{4,l}, f) = (-8)^w w!x \times \frac{1}{\prod_{t=0}^{w}(f+6+2l+8t)}.$$

It follows that

$$\Delta_8^w(A_4,f)=\sum_{k_1=0}^{w}\sum_{k_2=0}^{k_1}\sum_{k_3=0}^{k_2}\binom{w}{k_1}\binom{k_1}{k_2}\binom{k_2}{k_3}(-8)^w k_3!(k_2-k_3)!(k_1-k_2)!(w-k_1)!$$
$$\times x^4\Bigg\{\prod_{t=1}^{k_3+1}(f+8t)\prod_{t=k_3+1}^{k_2+1}(f+8t+2)\prod_{t=k_2+1}^{k_1+1}(f+8t+4)\prod_{k_1+1}^{w+1}(f+8t+6)\Bigg\}^{-1}.$$

As $\binom{w}{k_1}\binom{k_1}{k_2}\binom{k_2}{k_3}k_3!(k_2-k_3)!(k_1-k_2)!(w-k_1)!=w!$, $\sum_{k_1=0}^{w}\sum_{k_2=0}^{k_1}\sum_{k_3=0}^{k_2}1\le(w+1)^3$, and $x=O(f)$, there exists a constant C such that

$$\Delta_8^w(A_4,f)=O\Bigg\{(w+3)!C^w\prod_{t=1}^{w}(f+2t)^{-1}\Bigg\}.$$

D.2.10. *Proof of Lemma D.*2.8 *(on Page 63)* By the proof of Lemma D.2.4, we have

$$\Delta_2^{w_1}(A_1,f)=(-1)^{w_1}2^{w_1}w_1!x\prod_{s=1}^{w_1+1}A_{1,s}(f),$$

where $A_{1,s}(f)=1/(f+2s)$. It follows that

$$\Delta_4^{w_2}\left\{\Delta_2^{w_1}(A_1,f)\right\}=x(-2)^{w_1}w_1!\Delta_4^{w_2}\Bigg\{\prod_{s=1}^{w_1+1}A_{1,s}(f)\Bigg\}.\tag{D.49}$$

To prove Lemma D.2.8, by $x=\chi_f^2(\alpha)=O(f)$ and (D.49), it suffices to prove

$$\Delta_4^{w_2}\Bigg\{\prod_{s=1}^{w_1+1}A_{1,s}(f)\Bigg\}=\frac{(w_1+w_2)!}{w_1!}O\left\{C^{w_1+w_2}\prod_{s=1}^{w_1+w_2+1}(f+2s)^{-1}\right\}.\tag{D.50}$$

We next prove (D.50) by the mathematical induction. Consider $w_1=0$ first. Similarly to the proof of Lemma D.2.4, for each integer $1\le s\le w_1+1$, we have

$$\Delta_4^{w_2}(A_{1,s},f)=w_2!(-4)^{w_2}\prod_{k=0}^{w_2}(f+2s+4k).\tag{D.51}$$

Thus (D.50) holds for $w_1=0$. We then assume for integers $1\le l\le w_1$,

$$\Delta_4^{w_2}\Bigg\{\prod_{s=1}^{l}A_{1,s}(f)\Bigg\}=\frac{(w_2+l-1)!}{(l-1)!}O\left\{\prod_{k=1}^{w_2+l}(f+2k)^{-1}\right\},\tag{D.52}$$

and prove (D.50). By the Leibniz rule in Lemma D.2.1,

$$\Delta_4^{w_2}\Bigg\{\prod_{s=1}^{w_1+1}A_{1,s}(f)\Bigg\}=\sum_{k_2=0}^{w_2}\binom{w_2}{k_2}\Delta_4^{k_2}\Bigg\{\prod_{s_1=1}^{w_1}A_{1,s_1}(f)\Bigg\}\Delta_4^{w_2-k_2}(A_{1,w_1+1},f+4k_2).\tag{D.53}$$

Then by (D.51) and (D.52), we obtain

$$\begin{aligned}(\text{D.53}) &= \sum_{k_2=0}^{w_2}\binom{w_2}{k_2}\frac{(k_2+w_1-1)!}{(w_1-1)!}O\left(C^{w_1+k_2}\prod_{s_1=1}^{w_1+k_2}\frac{1}{f+2s_1}\right)\\ &\quad\times O\left\{(w_2-k_2)!C^{w_2-k_2}\prod_{s_2=0}^{w_2-k_2}\frac{1}{f+4k_2+2(w_1+1)+4s_2}\right\}\\ &= C^{w_1+w_2}\sum_{k_2=0}^{w_2}w_2!\binom{k_2+w_1-1}{k_2}O\left(\prod_{s=1}^{w_1+w_2+1}\frac{1}{f+2s}\right).\end{aligned}$$

By the hockey-stick identity, $\sum_{k_2=0}^{w_2}\binom{k_2+w_1-1}{k_2}=\binom{w_1+w_2}{w_2}$. Therefore, (D.50) is proved and then (D.49) follows.

D.2.11. *Proof of Lemma D.2.9 (on Page 64)* We next prove Lemma D.2.9 by discussing the cases when w_1+w_2 is odd and even, respectively.
(1) When w_1+w_2 is odd, $(w_1+w_2+1)/2$ is an integer, and then

$$\begin{aligned}(w_1+w_2)!\prod_{k=1}^{w_1+w_2}\frac{1}{f+2k}\times f^{(w_1+w_2+1)/2} &\le (w_1+w_2)!\prod_{k=(w_1+w_2+1)/2+1}^{w_1+w_2}\frac{1}{2k}\\ &\le 2^{-(w_1+w_2-1)/2}\prod_{k=1}^{(w_1+w_2+1)/2}k.\end{aligned}$$

To prove Lemma D.2.9, it now suffices to prove that there exists a constant C such that

$$\frac{1}{w_1!w_2!}\prod_{k=1}^{(w_1+w_2+1)/2}k\le C. \tag{D.54}$$

To prove (D.54), we use the following Lemma D.2.14.

LEMMA D.2.14 (FACTORIAL BOUND). *For any integer $w\ge 1$,*

$$\left(\frac{w}{e}\right)^w e\le w!\le\left(\frac{w+1}{e}\right)^{w+1}e.$$

Proof. This is a known bound on factorial in literature, and is obtained by $\int_1^w\ln x\mathrm{d}x\le\sum_{x=1}^w\ln x\le\int_0^w\ln(x+1)\mathrm{d}x$. □

Assume without loss of generality that $w_2\ge w_1$, and then by Lemma D.2.14,

$$\begin{aligned}&\frac{1}{w_1!w_2!}\prod_{k=1}^{(w_1+w_2+1)/2}k\\ &\le\frac{1}{e}\left(\frac{e}{w_1}\right)^{w_1}\left(\frac{e}{w_2}\right)^{w_2}\left(\frac{w_1+w_2+3}{2e}\right)^{(w_1+w_2+3)/2}\\ &=\frac{1}{e}\left(\frac{e^2}{w_1w_2}\frac{w_1+w_2+3}{2e}\right)^{w_1}\left(\frac{e^2}{w_2^2}\frac{w_1+w_2+3}{2e}\right)^{(w_2-w_1)/2}\left(\frac{w_1+w_2+3}{2e}\right)^{3/2}.\end{aligned}\tag{D.55}$$

As $w_1+w_2+3\le 4w_2$, there exists a constant C such that

$$(\text{D.55})\le C\left(\frac{2e}{w_1}\right)^{w_1}\left(\frac{2e}{w_2}\right)^{(w_2-w_1)/2}(w_1+w_2+3)^{3/2}.$$

When $w_2 - w_1 \geq 3$,

$$\text{(D.55)} \leq C \left(\frac{2e}{w_1}\right)^{w_1} \left(\frac{2e}{w_2}\right)^{(w_2-w_1-3)/2} \left\{\frac{2e(w_1+w_2+3)}{w_2}\right\}^{3/2},$$

which is bounded. When $0 \leq w_2 - w_1 \leq 2$,

$$\text{(D.55)} \leq C \left(\frac{2e}{w_1}\right)^{w_1} (2w_1+5)^{3/2},$$

which is also bounded. In summary, (D.55) is bounded.

(2) When $w_1 + w_2$ is even, similarly, we have

$$(w_1+w_2)! \prod_{k=1}^{w_1+w_2} \frac{1}{f+2k} \times f^{(w_1+w_2+1)/2} \leq 2^{-(w_1+w_2)/2-1} \prod_{k=1}^{(w_1+w_2)/2+1} k.$$

To prove Lemma D.2.9, it now suffices to prove that there exists a constant C such that

$$\frac{1}{w_1!w_2!} \prod_{k=1}^{(w_1+w_2)/2+1} k \leq C.$$

Similar analysis can be applied and the conclusions follow.

D.2.12. *Proof of Lemma D.*2.10 *(on Page 65)* When $w_1 = 0$, we know Lemma D.2.10 holds by Lemma D.2.5. Recall that we write $A_2(f) = A_{2,1}(f)A_{2,2}(f) - 1$ in Section D.2.7. Thus when $w_1 + w_2 \geq 1$,

$$\Delta_2^{w_1}\Delta_4^{w_2}(A_2, f) = \Delta_2^{w_1}\Delta_4^{w_2}(A_{2,1}A_{2,2}, f).$$

By the Leibniz rule in Lemma D.2.1,

$$\begin{aligned}
&\Delta_2^{w_1}\Delta_4^{w_2}(A_{2,1}A_{2,2}, f) \\
&= \sum_{k_1=0}^{w_1}\sum_{k_2=0}^{w_2} \binom{w_1}{k_1}\binom{w_2}{k_2} \Delta_2^{k_1}\Delta_4^{k_2}(A_{2,1}, f)\Delta_2^{w_1-k_1}\Delta_4^{w_2-k_2}(A_{2,2}, f+2k_1+4k_2). \qquad \text{(D.56)}
\end{aligned}$$

Following the proof of Lemma D.2.8, we have when $k_1 + k_2 \geq 1$,

$$\Delta_2^{k_1}\Delta_4^{k_2}(A_{2,1}, f) = (k_1+k_2)!O\left(C^{k_1+k_2} \prod_{s=1}^{k_1+k_2} \frac{1}{f+2s}\right),$$

and when $w_1 + w_2 - k_1 - k_2 \geq 1$,

$$\Delta_2^{w_1-k_1}\Delta_4^{w_2-k_2}(A_{2,2}, f+2k_1+4k_2) = (w_1+w_2-k_1-k_2)!O\left(C^{w_1+w_2-k_1-k_2} \prod_{s=k_1+k_2+1}^{w_1+w_2} \frac{1}{f+2s}\right).$$

Therefore,

$$\text{(D.56)} = w_1!w_2! \sum_{k_1=0}^{w_1}\sum_{k_2=0}^{w_2} \binom{k_1+k_2}{k_1}\binom{w_1+w_2-k_1-k_2}{w_1-k_1} O\left(C^{w_1+w_2} \prod_{s=1}^{w_1+w_2} \frac{1}{f+2s}\right).$$

By the ChuVandermonde identity,

$$\begin{aligned}
\sum_{k_1=0}^{w_1}\sum_{k_2=0}^{w_2} \binom{k_1+k_2}{k_1}\binom{w_1+w_2-k_1-k_2}{w_1-k_1} &= \sum_{m=0}^{w_1+w_2}\sum_{s_1=0}^{w_1} \binom{m}{s_1}\binom{w_1+w_2-m}{w_1-s_1} \\
&= (w_1+w_2+1)\binom{w_1+w_2}{w_1}.
\end{aligned}$$

Then $\Delta_2^{w_1}\Delta_4^{w_2}(A_2,f)=(w_1+w_2+1)!O\{C^{w_1+w_2}\prod_{s=1}^{w_1+w_2}(f+2s)^{-1}\}$.

D.2.13. *Proof of Lemma D.2.11 (on Page 65)* By the definition of $D_{2,4}(f)$, when $k_1+k_2\geq 1$,

$$\Delta_6^{k_2}\Delta_4^{k_1}(D_{2,4},f)=x^2\Delta_6^{k_2}\Delta_4^{k_1}(A_{1,1}A_{1,2},f),$$

where recall that we define $A_{1,t}=1/(f+2t)$ for integers t. By the Leibniz rule in Lemma D.2.1,

$$\Delta_6^{k_2}\Delta_4^{k_1}(A_{1,1}A_{1,2},f)=\sum_{s_2=0}^{k_2}\sum_{s_1=0}^{k_1}\binom{k_1}{s_1}\binom{k_2}{s_2}\Delta_6^{s_2}\Delta_4^{s_1}(A_{1,1},f)\Delta_6^{k_2-s_2}\Delta_4^{k_1-s_1}(A_{1,2},f+4s_1+6k_2)$$

Following the proof of Lemma D.2.8 in Section D.2.10, we similarly have

$$\Delta_6^{s_2}\Delta_4^{s_1}(A_{1,1},f)=(s_1+s_2)!O\left(C^{s_1+s_2}\prod_{k=1}^{s_1+s_2+1}\frac{1}{f+2k}\right).$$

Then following the proof of Lemma D.2.10 in Section D.2.12, we obtain Lemma D.2.11. The analysis will be very similar and thus the details are skipped.

D.2.14. *Proof of Lemma D.2.12 (on Page 66)* Note that we can write $D_{2,6}(f)=x^3\prod_{k=1}^3 A_{1,k}(f)-1$. By the Leibniz rule in Lemma D.2.1,

$$\begin{aligned}\Delta_4^{k_1}\Delta_6^{k_2}(D_{2,6},f)=&\sum_{s_1=0}^{k_2}\sum_{s_2=0}^{s_1}\binom{k_1}{s_1}\binom{s_1}{s_2}\sum_{t_1=0}^{k_1}\sum_{t_2=0}^{t_1}\binom{k_2}{t_1}\binom{t_1}{t_2}x^3\times\Delta_4^{t_2}\Delta_6^{s_2}(A_{3,1},f)\\&\times\Delta_4^{t_1-t_2}\Delta_6^{s_1-s_2}(A_{3,2},f+6s_2+4t_2)\Delta_4^{k_1-t_1}\Delta_6^{k_2-s_1}(A_{3,3},f+6s_1+4t_1).\end{aligned}$$

Following the proof of Lemma D.2.8 in Section D.2.10, we similarly have that for integers $t+s\geq 1$, and $l\in\{1,2,3\}$,

$$\Delta_4^t\Delta_6^s(A_{3,l})=(t+s)!O\left(C^{t+s}\prod_{m=1}^{t+s+1}\frac{1}{f+2m}\right).$$

By $x=\chi_f^2(\alpha)=O(f)$,

$$\begin{aligned}\Delta_4^{k_1}\Delta_6^{k_2}(D_{2,6},f)=&\sum_{s_1=0}^{k_2}\sum_{s_2=0}^{s_1}\sum_{t_1=0}^{k_1}\sum_{t_2=0}^{t_1}\binom{k_1}{s_1}\binom{s_1}{s_2}\binom{k_2}{t_1}\binom{t_1}{t_2}(t_2+s_2)!(t_1+s_1-t_2-s_2)!\\&\times(k_1+k_2-t_1-s_1)\times O\left(\prod_{m=1}^{k_1+k_2}\frac{1}{f+2m}\right).\end{aligned}$$

Similarly to the proof of Lemma D.2.10 in Section D.2.12, by the ChuVandermonde identity, we obtain

$$\begin{aligned}\Delta_4^{k_1}\Delta_6^{k_2}(D_{2,6},f)=&\sum_{s_1=0}^{k_2}\sum_{t_1=0}^{k_1}k_1!k_2!\binom{k_1+k_2-s_1-t_1}{k_1-s_1}\binom{s_1+t_1}{s_1}(s_1+t_1+1)\\=&(k_1+k_2+2)!\times O\left(\prod_{m=1}^{k_1+k_2}\frac{1}{f+2m}\right),\end{aligned}$$

where we use $s_1+s_2+1\leq k_1+k_2+1$ in the second equation.

D.2.15. *Proof of Lemma D.2.13 (on Page 66)* We prove Lemma D.2.13 similarly to the proof of Lemma D.2.9 in Section D.2.11 by discussing k_1+k_2 is odd and even, respectively.

(1) When k_1+k_2 is odd, similarly to the analysis of (D.55), we assume without loss of generality that $k_2 \geq k_1$, and obtain

$$\begin{aligned}
&\frac{(k_1+k_2+2)!}{k_1!k_2!} O\left(f^{-(k_1+k_2+1)/2} \prod_{t=1}^{k_1+k_2} \frac{1}{f+2t}\right) \\
&\leq \frac{2^{-(k_1+k_2-1)/2}}{k_1!k_2!}(k_1+k_2+2)(k_1+k_2+1) \prod_{t=1}^{(k_1+k_2+1)/2} t. \qquad \text{(D.57)}
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{(k_1+k_2+2)(k_1+k_2+1)}{k_1!k_2!} \prod_{t=1}^{(k_1+k_2+1)/2} t \\
&\leq C\left(\frac{e^2}{k_1k_2}\frac{k_1+k_2+3}{2e}\right)^{k_1}\left(\frac{e^2}{k_2^2}\frac{k_1+k_2+3}{2e}\right)^{(k_2-k_1)/2}(k_1+k_2+3)^{5/2} \\
&\leq C\left(\frac{2e}{k_1}\right)^{k_1}\left(\frac{2e}{k_2}\right)^{(k_2-k_1-5)/2}\left\{\frac{2e(k_1+k_2+3)}{k_2}\right\}^{5/2}. \qquad \text{(D.58)}
\end{aligned}$$

When $k_2-k_1 \geq 5$, we can see that (D.58) is bounded. When $k_2-k_1 \leq 4$, we have

$$\text{(D.58)} \leq C\left(\frac{k_2}{k_1}\right)^{(5-k_2+k_1)/2}\left(\frac{2e}{k_1}\right)^{(k_1+k_2-5)/2}\left(\frac{k_1+k_2+3}{k_2}\right)^{5/2},$$

which suggests that (D.58) is bounded. In summary, we know (D.58) is bounded, and therefore (D.57) $= O\{2^{-(k_1+k_2-1)/2}\}$.

(2) When k_1+k_2 is even, similar analysis can be applied, and then Lemma D.2.13 is proved.

D.2.16. *Proof of Lemma C.*3.1 *(on Page 41)* We prove Lemma C.3.1 based on (C.31). In each testing problem, we have $|\tau_{1,k}+\upsilon_{1,k}|/|\eta\xi_{1,k}| = o(1)$; see Sections C.3.1–C.3.6. Then under the conditions of Lemma C.3.1, we can apply Lemma D.1.2 and obtain for $1 \leq k \leq K_1$,

$$\begin{aligned}
&\log\Gamma\big\{\eta\xi_{1,k}(1-2it)+\tau_{1,k}+\upsilon_{1,k}\big\} \\
&= \left\{\eta\xi_{1,k}(1-2it)+\tau_{1,k}+\upsilon_{1,k}-\frac{1}{2}\right\}\log\big\{\eta\xi_{1,k}(1-2it)\big\} - \eta\xi_{1,k}(1-2it)+\log\sqrt{2\pi} \\
&\quad + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(\tau_{1,k}+\upsilon_{1,k})}{l(l+1)}\Big\{\eta\xi_{1,k}(1-2it)\Big\}^{-l} + O\Big(|\tau_{1,k}+\upsilon_{1,k}|^{L+1}/|\eta\xi_{1,k}|^{L}\Big).
\end{aligned}$$

Applying similar expansion to $\log\Gamma(\eta\xi_{1,k}+\tau_{1,k}+\upsilon_{1,k})$, we obtain

$$\begin{aligned}
&\log\Gamma\big\{\eta\xi_{1,k}(1-2it)+\tau_{1,k}+\upsilon_{1,k}\big\} - \log\Gamma\big(\eta\xi_{1,k}+\tau_{1,k}+\upsilon_{1,k}\big) \\
&= \left(\eta\xi_{1,k}+\tau_{1,k}+\upsilon_{1,k}-\frac{1}{2}\right)\log(1-2it) - 2it\eta\xi_{1,k}\log\big\{\eta\xi_{1,k}(1-2it)\big\} + 2it\eta\xi_{1,k} \\
&\quad + \sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(\tau_{1,k}+\upsilon_{1,k})}{l(l+1)(\eta\xi_{1,k})^l}\Big\{(1-2it)^{-l}-1\Big\} + O\Big(|\tau_{1,k}+\upsilon_{1,k}|^{L+1}/|\eta\xi_{1,k}|^{L}\Big).
\end{aligned}$$

Similarly, for $1 \le j \le K_2$, we have

$$\begin{aligned}
&\log\Gamma\big\{\eta\xi_{2,j}(1-2it)+\tau_{2,j}+\upsilon_{2,j}\big\}-\log\Gamma\big(\eta\xi_{2,j}+\tau_{2,j}+\upsilon_{2,j}\big)\\
&=\left(\eta\xi_{2,j}+\tau_{2,j}+\upsilon_{2,j}-\frac{1}{2}\right)\log(1-2it)-2it\eta\xi_{2,j}\log\big\{\eta\xi_{2,j}(1-2it)\big\}+2it\eta\xi_{2,j}\\
&\quad+\sum_{l=1}^{L-1}\frac{(-1)^{l+1}B_{l+1}(\tau_{2,j}+\upsilon_{2,j})}{l(l+1)(\eta\xi_{2,j})^l}\Big\{(1-2it)^{-l}-1\Big\}+O\Big(|\tau_{2,j}+\upsilon_{2,j}|^{L+1}/|\eta\xi_{2,j}|^L\Big).
\end{aligned}$$

Then by the form of $\varphi(t)$ in (C.31), we calculate

$$\begin{aligned}
\text{(C.31)}&=2it\eta\left(\sum_{k=1}^{K_1}\xi_{1,k}\log\xi_{1,k}-\sum_{j=1}^{K_2}\xi_{2,j}\log\xi_{2,j}\right)\\
&\quad+\left\{\sum_{k=1}^{K_1}(\xi_{1,k}+\tau_{1,k}+\upsilon_{1,k}-1/2)-\sum_{j=1}^{K_2}(\xi_{2,j}+\tau_{2,j}+\upsilon_{2,j}-1/2)\right\}\log(1-2it)\\
&\quad-2it\eta\left(\sum_{k=1}^{K_1}\xi_{1,k}\log\xi_{1,k}-\sum_{j=1}^{K_2}\xi_{2,j}\log\xi_{2,j}\right)-2it\eta(\log\eta-1)\left(\sum_{k=1}^{K_1}\xi_{1,k}-\sum_{j=1}^{K_2}\xi_{2,j}\right)\\
&\quad+\sum_{l=1}^{L-1}\varsigma_l\Big\{(1-2it)^{-l}-1\Big\}+O\left(\sum_{k=1}^{K_1}\frac{|\tau_{1,k}+\upsilon_{1,k}|^{L+1}}{|\eta\xi_{1,k}|^L}+\sum_{j=1}^{K_2}\frac{|\tau_{2,j}+\upsilon_{2,j}|^{L+1}}{|\eta\xi_{2,j}|^L}\right).
\end{aligned}$$

By the facts that $\tau_{1,k}=\eta\xi_{1,k}$, $\tau_{2,j}=\eta\xi_{2,j}$, and $\sum_{k=1}^{K_1}\xi_{1,k}=\sum_{j=1}^{K_2}\xi_{2,k}$, Lemma C.3.1 is proved.

D.3. *Lemmas for Theorems* 2.3, *A*.3, *&* *A*.6

D.3.1. *Proof of Lemma B.*3.2 *(on Page 29)* By (B.30) on Page 28,

$$\log\psi_1(s)=-\frac{pnti}{2}\log\frac{2e}{n}-\frac{pn(1-ti)}{2}\log(1-ti)+\log\frac{\Gamma_p\{(n-1)/2-nti/2\}}{\Gamma_p\{(n-1)/2\}}+\mu_n ti,$$

where $t=s/(n\sigma_n)$. We next examine $\log\psi_1(s)$ by the following Lemma D.3.1.

LEMMA D.3.1. *Let* $\{p=p_n;n\ge1\}$, $\{m=m_n;n\ge1\}$, $\{t_n;n\ge1\}$, *and* $\{s_n;n\ge1\}$ *satisfy that (i)* $p_n\to\infty$ *and* $p_n=o(n)$*; (ii) there exists* $\epsilon\in(0,1)$ *such that* $\epsilon\le m_n/n\le\epsilon^{-1}$*; (iii)* $t=t_n=O(ns/p)$*; (iv)* $s=s_n=o(\min\{(n/p)^{1/2},f^{1/6}\})$. *Then as* $n\to\infty$*,*

$$\begin{aligned}
&\log\frac{\Gamma_p\left(\frac{m-1}{2}+ti\right)}{\Gamma_p\left(\frac{m-1}{2}\right)}\\
&=\beta_{m,1}ti-\beta_{m,2}t^2+\beta_{m,3}(ti)+O\left(\frac{p^2t}{m^2}\right)+\left(\frac{1}{p}+\frac{p}{m}\right)O\left(\frac{p^2t^2}{m^2}\right)+O\Big(\frac{p^2t^3}{m^3}\Big),
\end{aligned}$$

where

$$\begin{aligned}
&\beta_{m,1}=-\left\{2p+\left(m-p-\frac{3}{2}\right)\log\left(1-\frac{p}{m-1}\right)\right\};\quad \beta_{m,2}=-\left\{\frac{p}{m-1}+\log\left(1-\frac{p}{m-1}\right)\right\};\\
&\beta_{m,3}(ti)=p\left\{\left(\frac{m-1}{2}+ti\right)\log\left(\frac{m-1}{2}+ti\right)-\frac{m-1}{2}\log\frac{m-1}{2}\right\}.
\end{aligned}$$

Proof. Please see Section D.3.2 on Page 75. □

By (B.7) and $f = \Theta(p^2)$, we know $t = s/(n\sigma_n) = O(s/p)$. Thus we can apply Lemma D.3.1 and expand

$$\log \frac{\Gamma_p\{(n-1)/2 - nti/2\}}{\Gamma_p\{(n-1)/2\}} = -\frac{n\beta_{n,1}ti}{2} - \frac{\beta_{n,2}n^2t^2}{4} + \beta_{n,3}\left(-\frac{nti}{2}\right) + O\left(\frac{p^2t}{n}\right) + \left(\frac{1}{p} + \frac{p}{n}\right) O\left(p^2t^2\right) + O\big(p^2t^3\big).$$

We next use the following Lemma D.3.2 to evaluate $\beta_{n,3}(-nti/2)$.

LEMMA D.3.2. *When* $p = p_n \to \infty$, $p = o(n)$, *and* $t = t_n = O(s/p)$ *with* $s = s_n = o(\min\{(n/p)^{1/2}, f^{1/6}\})$,

$$\beta_{n,3}\bigg(-\frac{nti}{2}\bigg) = -\frac{pnti}{2}\log\frac{n}{2} + \frac{pn(1-ti)}{2}\log(1-ti) + \frac{pti}{2} + O\bigg(pt^2 + \frac{pt}{n}\bigg).$$

Proof. Please see Section D.3.3 on Page 76. □

It follows that

$$\log\psi_1(s) = -\frac{\{p(n-1) + n\beta_{n,1}\}ti}{2} - \frac{\beta_{n,2}n^2t^2}{4} + \mu_n ti + O\bigg(\frac{p^2t}{n}\bigg) + \left(\frac{1}{p} + \frac{p}{n}\right) O\left(p^2t^2\right) + O\big(p^2t^3\big).$$

Since $\sigma_n^2 = \beta_{n,2}/2$, $\mu_n = \{p(n-1) + n\beta_{n,1}\}/2$, and $t = s/(n\sigma_n)$,

$$\log\psi_1(s) = -\frac{s^2}{2} + O\bigg(\frac{ps}{n}\bigg) + \left(\frac{1}{p} + \frac{p}{n}\right) O(s^2) + O\bigg(\frac{s^3}{p}\bigg),$$

where we use $t = O(s/p)$. As $\log\psi_0(s) = -s^2/2$, (B.31) is proved.

D.3.2. *Proof of Lemma D.*3.1 *(on Page 74)* By the property of the multivariate gamma function; see, e.g., Theorem 2.1.12 in Muirhead (2009),

$$\log\frac{\Gamma_p\left(\frac{m-1}{2} + ti\right)}{\Gamma_p\left(\frac{m-1}{2}\right)} = \sum_{j=1}^{p}\log\frac{\Gamma\left(\frac{m-j}{2} + ti\right)}{\Gamma\left(\frac{m-j}{2}\right)}. \tag{D.59}$$

Then by Lemma D.1.3 on Page 53,

$$\log\frac{\Gamma\left(\frac{m-j}{2} + ti\right)}{\Gamma\left(\frac{m-j}{2}\right)} = \sum_{j=1}^{p}\Bigg[\left(\frac{m-j}{2} + ti\right)\log\left(\frac{m-j}{2} + ti\right) - \left(\frac{m-j}{2}\right)\log\left(\frac{m-j}{2}\right) - ti - \frac{ti}{m-j} + O\bigg\{\frac{t+t^2}{(m-j)^2}\bigg\}\Bigg], \tag{D.60}$$

as $m \to \infty$ uniformly for all $1 \le j \le p$. Note that $t/(m-j) = t/m + (t/m) \times \{j/(m-j)\}$, and then

$$\sum_{j=1}^{p}\frac{ti}{m-j} = \frac{pti}{m} + O\left(\frac{p^2}{m^2}\right)ti. \tag{D.61}$$

By (D.60) and (D.61), we obtain as $m \to \infty$,

$$\text{(D.59)} = \sum_{j=1}^{p}\Bigg[\left(\frac{m-j}{2} + ti\right)\log\left(\frac{m-j}{2} + ti\right) - \left(\frac{m-j}{2}\right)\log\left(\frac{m-j}{2}\right)\Bigg] - \frac{(m+1)pti}{m} + O\left(\frac{p^2}{m^2}t + \frac{p}{m^2}t^2\right). \tag{D.62}$$

For $1 \le j \le p$, define

$$g_j(z) = \left(\frac{m-j}{2} + z\right)\log\left(\frac{m-j}{2} + z\right) - \left(\frac{m-1}{2} + z\right)\log\left(\frac{m-1}{2} + z\right),$$

where the real part of $z > -(m-p)/2$. It follows that the "$\sum_{j=1}^{p}$" term in the first row of (D.62) is equal to

$$p\left\{\left(\frac{m-1}{2} + ti\right)\log\left(\frac{m-1}{2} + ti\right) - \frac{m-1}{2}\log\frac{m-1}{2}\right\} + \sum_{j=1}^{p}\{g_j(ti) - g_j(0)\}. \tag{D.63}$$

To evaluate (D.63), we use the following Lemma D.3.3.

LEMMA D.3.3. *Let $p = p_m$ such that $1 \le p < m$, $p \to \infty$ and $p/m \to 0$ as $m \to \infty$. When $t = t_m = O(ms/p)$ with $s = s_m = o(\min\{(m/p)^{1/2}, p^{1/3}\})$, we have that, as $m \to \infty$,*

$$\sum_{j=1}^{p}\{g_j(ti) - g_j(0)\} = \nu_{1,m}ti - \frac{\nu_{2,m}^2}{2}t^2 + O\left(\frac{p^2t}{m^2}\right) + \left(\frac{1}{p} + \frac{p}{m}\right)O\left(\frac{p^2t^2}{m^2}\right) + O\Big(\frac{p^2t^3}{m^3}\Big),$$

where

$$\begin{aligned}\nu_{1,m} =& \left(p - m + \frac{3}{2}\right)\log\left(1 - \frac{p}{m-1}\right) - \frac{m-1}{m}p, \\ \nu_{2,m}^2 =& -2\left\{\frac{p}{m-1} + \log\left(1 - \frac{p}{m-1}\right)\right\}.\end{aligned} \tag{D.64}$$

Proof. Please see Section D.3.4 on Page 77. □

Then by Lemma D.3.3,

$$\begin{aligned}\text{(D.63)} =& \, p\left\{\left(\frac{m-1}{2} + ti\right)\log\left(\frac{m-1}{2} + ti\right) - \frac{m-1}{2}\log\frac{m-1}{2}\right\} \\ &+ \nu_{1,m}ti - \frac{\nu_{2,m}^2}{2}t^2 + O\left(\frac{p^2t}{m^2}\right) + \left(\frac{1}{p} + \frac{p}{m}\right)O\left(\frac{p^2t^2}{m^2}\right) + O\Big(\frac{p^2t^3}{m^3}\Big).\end{aligned}$$

In summary, Lemma D.3.1 can be proved by noticing

$$\begin{aligned}&\beta_{m,1} = \nu_{1,m} - \frac{(m+1)p}{m}, \qquad \beta_{m,2} = \nu_{2,m}^2/2 \\ &\beta_{m,3}(ti) = p\left\{\left(\frac{m-1}{2} + ti\right)\log\left(\frac{m-1}{2} + ti\right) - \frac{m-1}{2}\log\frac{m-1}{2}\right\}.\end{aligned}$$

D.3.3. *Proof of Lemma D.*3.2 *(on Page 75)* By Taylor's series,

$$\begin{aligned}p^{-1}\beta_{n,3}(-nti/2) =& -\frac{nti}{2}\log\frac{n}{2} - \frac{nti}{2}\log\left(1 - ti - \frac{1}{n}\right) + \frac{n-1}{2}\log\left(1 - ti - \frac{ti}{n-1}\right) \\ =& -\frac{nti}{2}\log\frac{n}{2} - \frac{nti}{2}\log(1-ti) + \frac{nti}{2n(1-ti)} + O\Big(\frac{nt}{n^2}\Big) \\ &+ \frac{n-1}{2}\log(1-ti) - \frac{n-1}{2}\frac{ti}{(n-1)(1-ti)} + \frac{n-1}{2}O\Big(\frac{t^2}{n^2}\Big) \\ =& -\frac{nti}{2}\log\frac{n}{2} + \frac{n(1-ti)-1}{2}\log(1-ti) + O\Big(\frac{t+t^2}{n}\Big).\end{aligned}$$

It follows that

$$\beta_{n,3}(-nti/2) = -\frac{pnti}{2}\log\frac{n}{2} + \frac{pn(1-ti)}{2}\log(1-ti) + \frac{pti}{2} + O\Big(pt^2 + \frac{pt}{n}\Big).$$

D.3.4. *Proof of Lemma D.*3.3 *(on Page 76)* The first-order derivatives of $g_j(z)$ is

$$g_j^{(1)}(z) = \log\left(\frac{m-j}{2}+z\right) - \log\left(\frac{m-1}{2}+z\right),$$

and for $l \geq 2$, the l-th order derivatives of $g_j(z)$ is

$$\begin{aligned} g_j^{(l)}(z) &= (-1)^{l-2}(l-2)!\left\{\left(\frac{m-j}{2}+z\right)^{-(l-1)} - \left(\frac{m-1}{2}+z\right)^{-(l-1)}\right\} \\ &= (-1)^{l-2}(l-2)!\left(\frac{m-1}{2}+z\right)^{-(l-1)}\sum_{v=1}^{l-1}\binom{l-1}{v}\left(\frac{j-1}{m-j+2z}\right)^v. \end{aligned}$$

By Taylor's expansion, $g_j(ti) - g_j(0) = \sum_{l=1}^{\infty} g_j^{(l)}(0)z^l/l!$. In particular,

$$g_j^{(1)}(0) = \log(m-j) - \log(m-1), \qquad g_j^{(2)}(0) = \frac{2}{m-j} - \frac{2}{m-1}.$$

When $z = ti$, $t = t_m = O(ms/p)$, and $l \geq 3$, as $j - 1/(m-j+2z) = O(p/m) = o(1)$,

$$g_j^{(l)}(0)z^l/l! = O\Big(\frac{1}{m^{l-1}}\frac{p}{m}t^l\Big) = O\Big(\frac{p}{m^l}\Big)t^l.$$

As $t/m = O(s/p) = o(1)$,

$$\sum_{j=1}^{p}\{g_j(ti) - g_j(0)\} = \sum_{j=1}^{p} g_j^{(1)}(0)ti - \frac{1}{2}\sum_{j=1}^{p} g_j^{(2)}(0)t^2 + O\Big(\frac{p^2t^3}{m^3}\Big).$$

By Lemma A.2 in Jiang and Qi (2015),

$$\sum_{j=1}^{p} g_j^{(1)}(0) = \nu_{1,m} + O(\nu_{2,m}^2), \qquad \sum_{j=1}^{p} g_j^{(2)}(0) = \nu_{2,m}^2\left\{1 + O\left(\frac{1}{p}+\frac{p}{m}\right)\right\},$$

where $\nu_{1,m}$ and $\nu_{2,m}^2$ are defined in (D.64). In summary,

$$\sum_{j=1}^{p}\{g_j(ti) - g_j(0)\} = \nu_{1,m}ti - \frac{\nu_{2,m}^2}{2}t^2 + O(\nu_{2,m}^2)t + \nu_{2,m}^2 O\left(\frac{1}{p}+\frac{p}{n}\right)t^2 + O\Big(\frac{p^2t^3}{m^3}\Big).$$

Then Lemma D.3.3 follows by $\nu_{2,m}^2 = O(p^2/m^2)$.

D.3.5. *Proof of Lemma C.*4.1 *(on Page 51)* By Taylor's series, we have (C.44). In addition, for (C.45), note that we can write

$$p^{-1}\varrho_l(t) = \frac{l-1}{2}\log\left(1+\frac{lt}{l-1}\right) + \frac{lt}{2}\log\left(\frac{l-1}{2}+\frac{lt}{2}\right).$$

By Taylor's series $\log x = \log a + \sum_{l=1}^{L-1} (-1)^{l-1} l^{-1} (x/a-1)^l + O\{(x/a-1)^L\}$, we obtain

$$\begin{aligned}\frac{\varrho_l(t)}{p} &= \frac{l}{2}\log\left(1+t+\frac{t}{l-1}\right) - \frac{1}{2}\log\left(1+\frac{lt}{l-1}\right) + \frac{lt}{2}\log\left\{\frac{l(1+t)}{2} - \frac{1}{2}\right\} \\ &= \frac{l}{2}\log(1+t) + \frac{lt}{2(l-1)(1+t)} - \frac{lt}{2(l-1)} + \frac{lt}{2}\log\frac{l(1+t)}{2} - \frac{t}{2(1+t)} + O\left(\frac{t}{l}+t^2\right) \\ &= \frac{l(1+t)}{2}\log(1+t) + \frac{lt}{2}\log\frac{l}{2} - \frac{t}{2} + O\left(\frac{t}{l}+t^2\right).\end{aligned}$$

Then by $n = \sum_{j=1}^{k} n_j$, we have

$$-\varrho_n(t) + \sum_{j=1}^{k} \varrho_{n_j}(t) = \left(1-k-n\log n + \sum_{j=1}^{k} n_j \log n_j\right)\frac{tp}{2} + O\left(\frac{pt}{n}+pt^2\right).$$

REFERENCES

Abramowitz, M. and I. A. Stegun (1970). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (9th ed.), Volume 55. US Government printing office.

Anastasiou, A. and G. Reinert (2018). Bounds for the asymptotic distribution of the likelihood ratio. *arXiv preprint arXiv:1806.03666*.

Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis*. Wiley Series in Probability and Statistics. Wiley.

Bai, Z., 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.

Bai, Z., D. Jiang, J.-f. Yao, and S. Zheng (2013). Testing linear hypotheses in high-dimensional regressions. *Statistics 47*(6), 1207–1223.

Barndorff-Nielsen, O. and P. Hall (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika 75*(2), 374–378.

Boucheron, S. and P. Massart (2011). A high-dimensional Wilks phenomenon. *Probability theory and related fields 150*(3-4), 405–433.

Candès, E. J. and P. Sur (2020). The phase transition for the existence of the maximum likelihood estimate in high-dimensional logistic regression. *The Annals of Statistics*.

Chen, S. X. and H. Cui (2006). On bartlett correction of empirical likelihood in the presence of nuisance parameters. *Biometrika 93*(1), 215–220.

Chen, S. X., L. Peng, and Y.-L. Qin (2009). Effects of data dimension on empirical likelihood. *Biometrika 96*(3), 711–722.

Chen, Y., J. Huang, Y. Ning, K.-Y. Liang, and B. G. Lindsay (2018). A conditional composite likelihood ratio test with boundary constraints. *Biometrika 105*(1), 225–232.

Cleff, T. (2019). *Applied Statistics and Multivariate Data Analysis for Business and Economics: A Modern Approach Using SPSS, Stata, and Excel*. Springer.

Cordeiro, G. M. and F. Cribari-Neto (2014). *An introduction to Bartlett correction and bias reduction*. Springer.

DiCiccio, T., P. Hall, and J. Romano (1991). Empirical likelihood is Bartlett-correctable. *the Annals of Statistics 19*(2), 1053–1061.

Drton, M. and B. Williams (2011). Quantifying the failure of bootstrap likelihood ratio tests. *Biometrika 98*(4), 919–934.

Fan, J., H.-N. Hung, and W.-H. Wong (2000). Geometric understanding of likelihood ratio statistics. *Journal of the American Statistical Association 95*(451), 836–841.

Fan, J., C. Zhang, and J. Zhang (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of statistics 29*(1), 153–193.

Fan, J. and W. Zhang (2004). Generalised likelihood ratio tests for spectral density. *Biometrika 91*(1), 195–209.

He, X. and Q.-M. Shao (2000). On parameters of increasing dimensions. *Journal of Multivariate Analysis 73*(1), 120–135.

He, Y., T. Jiang, J. Wen, and G. Xu (2020). Likelihood ratio test in multivariate linear regression: from low to high dimension. *Statistica Sinica*.

Hjort, N. L., I. W. McKeague, and I. Van Keilegom (2009). Extending the scope of empirical likelihood. *The Annals of Statistics 37*(3), 1079–1111.

Jiang, T. and Y. Qi (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scandinavian Journal of Statistics 42*(4), 988–1009.

Jiang, T. 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.

Luke, Y. L. (1969). *Special functions and their approximations*, Volume 2. Academic press.

Muirhead, R. J. (2009). *Aspects of multivariate statistical theory*, Volume 197. John Wiley & Sons.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 90–120.

Pituch, K. A. and J. P. Stevens (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBMs SPSS*. Routledge.

Portnoy, S. (1985). Asymptotic behavior of M estimators of p regression parameters when p^2/n is large; II. Normal approximation. *The Annals of Statistics*, 1403–1417.

Portnoy, S. (1988). Asymptotic behavior of likelihood methods for exponential families when the number of parameters tends to infinity. *The Annals of Statistics*, 356–366.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1992). *Numerical recipes in Fortran 77: the art of scientific computing*. Cambridge university press.

Sur, P. and E. J. Candès (2019). A modern maximum-likelihood theory for high-dimensional logistic regression. *Proceedings of the National Academy of Sciences 116*(29), 14516–14525.

Sur, P., Y. Chen, and E. J. Candès (2019). The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square. *Probability Theory and Related Fields 175*(1-2), 487–558.

Tang, C. Y. and C. Leng (2010). Penalized high-dimensional empirical likelihood. *Biometrika 97*(4), 905–920.

Ushakov, N. G. (2011). *Selected topics in characteristic functions*. Walter de Gruyter.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, Volume 3. Cambridge university press.

Wang, L. (2011). GEE analysis of clustered binary data with diverging number of covariates. *The Annals of Statistics 39*(1), 389–417.

Whittaker, E. T. and G. N. Watson (1996). *A course of modern analysis*. Cambridge university press.

Xu, M., D. Zhang, and W. B. Wu (2019). Pearson's chi-squared statistics: approximation theory and beyond. *Biometrika 106*(3), 716–723.

Zheng, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Annales de l'IHP Probabilités et statistiques 48*(2), 444–476.

Zwillinger, D. (2002). *CRC standard mathematical tables and formulae* (31st ed.). CRC press.