ESAIM: PS 23 (2019) 672–696
https://doi.org/10.1051/ps/2019020

ESAIM: Probability and Statistics
www.esaim-ps.org

A TEST FOR BLOCK CIRCULAR SYMMETRIC COVARIANCE STRUCTURE WITH DIVERGENT DIMENSION*

Junshan Xie** and Gaoming Sun

Abstract. The paper considers the likelihood ratio (LR) test on the block circular symmetric covariance structure of a multivariate Gaussian population with divergent dimension. When the sample size n, the dimension of each block p and the number of blocks u satisfy $pu < n-1$ and $p = p(n) \to \infty$ as $n \to \infty$, the asymptotic distribution and the moderate deviation principle of the logarithmic LR test statistic under the null hypothesis are established. Some numerical simulations indicate that the proposed LR test method performs well in the divergent-dimensional block circular symmetric covariance structure test.

Mathematics Subject Classification. 62H15, 62E20.

Received August 11, 2018. Accepted September 2, 2019.

1. Introduction

The block circular symmetric (BCS) covariance structure is firstly introduced by Olkin [28] on studying a special signal processing model. Consider a point source (or satellite) A situated at the geocenter of a regular polygon of u sides, from which a signal will be transmitted to u vertices $T_1, T_2, \ldots, T_u$. The signal received at T_i is denoted by a p-dimensional random vector $\boldsymbol{\xi}_i = (\xi_{i1}, \ldots, \xi_{ip})', i = 1, \ldots, u$. Assume that the vertices $T_1, T_2, \cdots, T_u$ are exchangeable, that is, the covariance matrices of $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$ only depend on the number of vertices separating between receivers T_i and T_j. Thus

$$\mathrm{Var}(\boldsymbol{\xi}_i, \boldsymbol{\xi}_{i+j}) = \boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}_{u-j}, \quad i = 1, 2, \ldots, u, \quad j = 0, 1, \ldots, u-1,$$

where $\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_{u-1}$ are all $p \times p$ positive definite symmetric matrices. Here $\boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}_{u-j}$ comes from the fact that the "distance" from two vertices T_i and T_{i+j}, located on a circle with u vertices, is same to that of T_i and

*This work is supported by National Natural Science Foundation of China (Grant No. 11401169) and Key Research Project in Higher Education of Henan Province (Grant No. 20A110001).

Keywords and phrases: Likelihood ratio test, block circular symmetric model, asymptotic normality, moderate deviation principle.

School of Mathematics and Statistics, Henan University, Kaifeng 475000, PR China.

** Corresponding author: xjsnwpu@163.com

Article published by EDP Sciences

T_{i+u-j}. Denote $\boldsymbol{X}' = (\boldsymbol{\xi}_1', \dots, \boldsymbol{\xi}_u')$. Then $\boldsymbol{X} \in \mathbb{R}^{pu}$ has the BCS covariance matrix

$$\boldsymbol{\Sigma}_{\mathrm{BCS}} := \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \cdots & \boldsymbol{\Sigma}_{u-1} \\ \boldsymbol{\Sigma}_{u-1} & \boldsymbol{\Sigma}_0 & \cdots & \boldsymbol{\Sigma}_{u-2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \cdots & \boldsymbol{\Sigma}_0 \end{pmatrix}, \tag{1.1}$$

where $\boldsymbol{\Sigma}_r = \boldsymbol{\Sigma}_{u-r},\ r = 0, 1, \dots, u$. Let $\boldsymbol{I}_r$ be the r-order identity matrix and $\boldsymbol{W}_r = \begin{pmatrix} \mathbf{0} & \boldsymbol{I}_{u-r} \\ \boldsymbol{I}_r & \mathbf{0} \end{pmatrix}$. Then we can rewrite (1.1) as

$$\boldsymbol{\Sigma}_{\mathrm{BCS}} = \sum_{r=0}^{u-1} \boldsymbol{W}_r \otimes \boldsymbol{\Sigma}_r,$$

where $\otimes$ is the Kronecker product. It is of interest to mention that the covariance matrices are different for the two cases that the number of vertices u is even or odd. For example, when $u = 4$ and $u = 5$, we can respectively get different structures of $\boldsymbol{\Sigma}_{\mathrm{BCS}}$ as

$$\begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 \end{pmatrix}.$$

The BCS covariance structure model is a generalization of the circular symmetric covariance structure model considered by Olkin and Press [29]. Evidently, when $p = 1$, $\boldsymbol{\Sigma}_k, k = 1, \dots, u-1$ are all real-valued constants, $\boldsymbol{\Sigma}_{\mathrm{BCS}}$ is a u-dimensional circular symmetric covariance structure. There are some results on the parameter estimation and hypothesis test on the circular symmetric covariance structure, such as Olkin and Press [29], Nagar *et al.* [25, 26], Marques and Coelho [23] and Yi and Xie [35]. On the other hand, if $p \geq 2$ and $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_{u-1}$, we can get the block compound symmetric covariance structure, which is also a popular object in statistics, one is refered to Rao [30, 31], Arnold [2], Levia [17], Roy and Levia [32] and Makoto *et al.* [22]. Furthermore, when $p = 1$ and $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_{u-1} = c \in \mathbb{R}$, we will get a u-dimensional intraclass correlation structure covariance matrix, which is also well studied in the literature includes Wilks [34], Srivastava [33] and Kato *et al.* [16].

For the BCS covariance structure model, Olkin [28] firstly studied the likelihood ratio (LR) test on the BCS covariance structure of multivariate Gaussian population under the assumption of both p and u are fixed. Coelho [8] further got the near-exact distribution of the LR statistics by the eigenblock and eigenmatrix decomposition method, Liang *et al.* [19] considered the LR test on the the BCS covariance matrix versus other special structure covariance matrix. Liang *et al.* [18, 20] also studied the parameter estimation of the BCS covariance matrix. In addition, the spectral properties of BCS covariance is also an interesting topic, some results had been obtained by Basilevsky [4], Nahtman and Rosen [27] and Liang *et al.* [20].

To the best of our knowledge, the existing results on the BCS covariance structure are all based on the assumption of both p and u are fixed. In modern statistical data case, we frequently encounter the data, including financial data, consumer data, multimedia data and signal data, whose dimension and the sample size are all very large, it is usually called by the divergent-dimensional problem in the literature, and the traditional multivariate statistical theories and methods, established under the assumption that the dimension is fixed, will no longer work efficiently in this case(see Bai and Saranadasa [3]). So it is an interesting work to find some effective methods to deal with the divergent-dimensional statistical problems.

In the paper, we will mainly consider the hypothesis test of the BCS covariance structure of a multivariate Gaussian population with divergent dimension. Assume that a random vector $\boldsymbol{X} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2', \ldots, \boldsymbol{\xi}_u')'$ is subject to multivariate Gaussian distribution $\boldsymbol{N}_{pu}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be a random sample of size n from the population $\boldsymbol{X}$. When the dimension of each block p is divergent along with the sample size n tends to infinity, this is, $p = p(n) \to \infty$ as $n \to \infty$, we will consider the hypothesis test

$$\boldsymbol{H}_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{BCS}} \quad \textbf{vs} \quad \boldsymbol{H}_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_{\mathrm{BCS}}, \tag{1.2}$$

where $\boldsymbol{\Sigma}_{\mathrm{BCS}}$ defined in (1.1) and $\boldsymbol{\Sigma}_i, i = 1, \ldots, u-1$ are all unspecified positive definite matrices.

The LR test method will be adopted in the paper. Firstly, we will establish the moments of the LR statistic under the null hypothesis. Secondly, using the continuity theorem of the moment generating function and the asymptotic expansion method of the Gamma function developed by Jiang *et al.* [14, 15], we will get the asymptotic distribution of the logarithmic LR statistic. At last, in order to investigate the convergence properties of the logarithmic LR statistic, we will also study the moderate deviation principle (MDP) by a similar arguments in Jiang and Wang [13].

Now, we will give some examples on testing the divergent-dimensional BCS covariance structure.

Example 1.1. Stationary reciprocal processes defined on a finite interval of the integer line is a special class of Markov random fields. They can be used to describe signals which naturally live in a finite region of the time or space line (see Carli *et al.* [5, 6]). A p-dimensional stationary periodic reciprocal process on a finite interval $[1, u]$ is an ordered collection of zero-mean random p-variate vector $\boldsymbol{y} := \{\boldsymbol{y}(k), k = 1, 2, \ldots, u\}$, which will be denote as a column vector with pu-dimensional components, the auto-covariance $E[\boldsymbol{y}(k)\boldsymbol{y}(j)']$ depend on the difference between k, j and it is periodic of period u, namely,

$$E[\boldsymbol{y}(k)\boldsymbol{y}(j)'] = \boldsymbol{\Sigma}_{k-j} = \boldsymbol{\Sigma}_{u-(k-j)}.$$

Then the covariance matrix $\boldsymbol{\Sigma}$ of the process $\boldsymbol{y}$ restricted to $[1, u]$ has a BCS covariance structure defined in (1.1). Lindquist and Picci [21] proposed an image compression method by virtue of the p-dimensional stationary periodic reciprocal process. In their method, it is assumed that the image can be regarded as a $p \times u$ matrix of pixels where the columns form a p-dimensional reciprocal process, which can be extended to a periodic process with period u outside the interval $[0, u]$. If one want to test the rationality of their model assumption under the case of p is large, then the test of the divergent-dimensional BCS covariance structure will be considered.

Example 1.2. In a public health problem studied by Hartley and Naik [12], the disease incidence rates of (relatively homogeneous) observation stations or sectors placed around the city center are usually assumed to be circularly correlated. Suppose there are u cities and each city has p observation stations. Let $Y_{ij}(i = 1, 2, \ldots, p, j = 1, 2, \ldots, u)$ denotes disease incidence rate in the ith observation station of the jth city. Write $\boldsymbol{y}' = (\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_u')$ where $\boldsymbol{y}_j = (Y_{1j}, Y_{2j}, \ldots, Y_{pj})'$ for each $1 \leq j \leq u$. The covariance matrix of $\boldsymbol{y}$ has a BCS structure defined in (1.1) when the u cities are exchangeable. If the number of the observation stations p is large, then the rationality test problem of the model can be achieved by the testing of the divergent-dimensional BCS covariance structure.

Example 1.3. The similar BCS covariance structure can also get from the data analysis problem in Gotway and Cressie [23], who studied the soil-water infiltration ability of u towns. The soil-water infiltration data are collected from p locations of each town. Denote $Y_{ij}(i = 1, 2, \ldots, p, j = 1, 2, \ldots, u)$ to be the soil-water-infiltration data of p locations contained by u towns. Write $\boldsymbol{y}' = (\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_u')$ where $\boldsymbol{y}_j = (Y_{1j}, Y_{2j}, \ldots, Y_{pj})'$ for each $1 \leq j \leq u$. As the location varies across the field, the ability of water to infiltrate soil will vary spatially so that the data of the p locations of a town are circularly correlated. And we also assume that all the towns are exchangeable by the prior knowledge. Thus, $\boldsymbol{y}$ also has a BCS covariance structure defined in (1.1), and the corresponding divergent-dimensional BCS covariance structure is appropriate when the number of the locations is large.

For other spatial statistics model, if there is a spatial circular layout on one factor and another factor satisfies the property of exchangeability, and the dimension of the circularly correlated factor is large, then it is quite necessary to test whether the model has the divergent-dimensional BCS covariance structure.

The organization of the paper is as follows. In the next section, we will give the arbitrary order moments of the LR statistic. The asymptotic normality of the LR statistic with divergent dimension is given in Section 3. Section 4 states the MDP result of the LR statistic. In Section 5, we will consider the performance of the proposed method by numerical simulations. Section 6 summarizes the main results of the paper. At last, some technical proofs are listed in the appendix.

Throughout the paper, we assume that p depends on n and simply write p for brevity of notation. $|\boldsymbol{A}|$ denotes the determinant value of matrix $\boldsymbol{A}$, $\boldsymbol{B} > 0$ means the matrix $\boldsymbol{B}$ is positive definite. Meanwhile, we will use $\xrightarrow{d}$ to denote the convergence in distribution. The notation $a_n = o(b_n)$ stands for $a_n/b_n \to 0$ as $n \to \infty$, while $a_n = O(b_n)$ stands for the sequence $\{a_n/b_n, n = 1, 2, \dots\}$ is bounded. In addition, $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$.

2. Moments of the LR statistic

Recall that $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_n$ is a random sample from the population $\boldsymbol{X} \sim \boldsymbol{N}_{pu}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Write $\boldsymbol{X}_i = (\boldsymbol{X}_{i1}^{'}, \boldsymbol{X}_{i2}^{'}, \dots, \boldsymbol{X}_{iu}^{'})^{'}$, where $\boldsymbol{X}_{ij}$ is p-dimensional random vector for each (i, j) with $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, u$.

Write

$$\bar{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i, \quad \boldsymbol{S} = \sum_{i=1}^{n}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^{'}.$$

Denote Ω to be the entire parameter space of $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and Ω_0 to be the subset of Ω restricted by $\boldsymbol{H}_0$ in (1.2). Write $L(\omega) = \prod_{i=1}^{n} P\{\boldsymbol{X} = \boldsymbol{X}_i\}$ to be the likelihood functions. By the arguments in Olkin [28], we can obtain the LR statistic by

$$\frac{\max_{\omega\in\Omega_0} L(\omega)}{\max_{\omega\in\Omega} L(\omega)} = \left[\frac{2^{2p(u-m-1)}|\boldsymbol{V}|}{\prod_{j=1}^{u}|\boldsymbol{V}_j|}\right]^{\frac{n}{2}} := \Lambda^{\frac{n}{2}}, \tag{2.1}$$

where

$$\boldsymbol{V} = (\boldsymbol{V}_{ij})_{u\times u} = (\boldsymbol{B}\otimes \boldsymbol{I}_p)\boldsymbol{S}(\boldsymbol{B}'\otimes \boldsymbol{I}_p), \tag{2.2}$$

$$\boldsymbol{B} = (\beta_{ij})_{u\times u}, \quad i, j = 1, \dots, u,$$

$$\beta_{ij} = u^{-\frac{1}{2}}\Big\{\cos\left[2\pi u^{-1}(i-1)(j-1)\right] + \sin[2\pi u^{-1}(i-1)(j-1)]\Big\},$$

$\boldsymbol{V}_j, j = 1, \dots, u$ are independent random variables related to the diagonal entries of the random matrix $\boldsymbol{V}$ (see Rem. 2.2 below), and satisfy

$$\boldsymbol{V}_j = \boldsymbol{V}_{u-j+2}, \quad j = 2, \dots, u.$$

In order to get the asymptotic distribution of the test statistic Λ, we will first give a result on the moments of Λ as follows.

Theorem 2.1. *Let Λ be defined in (2.1). Under $\boldsymbol{H}_0$ in (1.2), for any $h > -\frac{n-pu}{2}$, we have*

$$E[\Lambda^h] = 2^{ph(u-v)} \frac{\Gamma_{pu}(\frac{n+2h-1}{2})}{\Gamma_{pu}(\frac{n-1}{2})} \left[\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2h-1}{2})}\right]^v \left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2h-1)}\right]^{\frac{u-v}{2}},$$

where

$$v = \begin{cases} 1, & if\ u\ is\ odd \\ 2, & if\ u\ is\ even \end{cases} \tag{2.3}$$

and $\Gamma_q(x)$ is the q-variate Gamma function (one can refer the definition of multivariate Gamma function to Thm. 2.1.12 in Murihead [24]).

Remark 2.2. Theorem 2.1 used the fact that $\boldsymbol{V}_j,\ j = 2, \ldots, u$ are different when u is even or odd, in particular, when $u = 2m + 1$ is odd, m is a positive integer, $\boldsymbol{V}_j, j = 1, \ldots, m + 1$ are defined by

$$(\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_{m+1}) = (\boldsymbol{V}_{11}, \boldsymbol{V}_{22} + \boldsymbol{V}_{uu}, \boldsymbol{V}_{33} + \boldsymbol{V}_{u-1,u-1} \ldots, \boldsymbol{V}_{m+1,m+1} + \boldsymbol{V}_{u-(m-1),u-(m-1)}),$$

when $u = 2m$ is even, the definitions of $\boldsymbol{V}_j, j = 1, \ldots, m + 1$ are

$$(\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_m, \boldsymbol{V}_{m+1}) = (\boldsymbol{V}_{11}, \boldsymbol{V}_{22} + \boldsymbol{V}_{uu}, \ldots, \boldsymbol{V}_{mm} + \boldsymbol{V}_{u-(m-2),u-(m-2)}, \boldsymbol{V}_{m+1,m+1}).$$

3. Asymptotic normality

When both the dimension of each block p and the length of each block u are all fixed, Olkin [28] get an asymptotic expansion for the logarithmic LR statistic Λ as

$$P\{-\rho(n-1)\log\Lambda \leq x\} = P\{\chi_f^2 \leq x\} + O(n^{-2}), \tag{3.1}$$

where

$$f = \frac{2p^2u(u-1) + p(p+1)(u-v)}{4},$$

$$\rho = 1 - \frac{4p^2u(u-1)(2pu+2p+u) + p(2p^2+3p-1)(u-v)}{48nf}$$

and v is defined in (2.3).

However, the performance of the Chi-square approximation in (3.1) will become poor when p is large. Now we can get the limiting distribution of LR statistic Λ under the assumption of $p \to \infty$.

Theorem 3.1. *Assume that $p = p(n)$ is a series of positive integers depending on n such that $pu < n - 1$ for all $n > 5$ and $p \to \infty$ as $n \to \infty$. Let Λ be defined in (2.1). Then under $\boldsymbol{H}_0$, we have*

$$\frac{\log\Lambda - \mu_{n,v}}{\sigma_{n,v}} \xrightarrow{d} \boldsymbol{N}(0,1),$$

where

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

Remark 3.2. It follows by Theorem 3.1 and the Taylor expansion formula that

$$\begin{aligned}\sigma_{n,v}^2 &= 2\Big[-v\sum_{k=1}^{\infty}\frac{p^k}{k(n-1)^k}-2(u-v)\sum_{k=1}^{\infty}\frac{p^k}{k(2n-2)^k}+\sum_{k=1}^{\infty}\frac{p^k u^k}{k(n-1)^k}\Big]\\ &= 2\sum_{k=1}^{\infty}\frac{p^k}{k(n-1)^k}\left(u^k-\frac{u-v}{2^{k-1}}-v\right).\end{aligned}$$

Recall the definition of v in (2.3). By the fact $u > 1$, we have

$$u^k-\frac{u-v}{2^{k-1}}-v>0$$

for all $k > 1$. Thus, we can see that $\sigma_{n,v}^2 > 0$ is well defined.

In Theorem 3.1, we do not make any special assumptions on the parameter u except that $p \to \infty$ and $pu < n-1$. For example, we can take u to be a suitable fixed integer or $u = u(n) \to \infty$ with some suitable rate. By the definition of v in (2.3), the asymptotic distribution in Theorem 3.1 depends on whether u is odd or even. In fact, when $u = u(n) \to \infty$, the dependence on the odevity of u will be vanished. This can seen from the following result, which can be deduced by using Taylor expansion method on Theorem 3.1.

Corollary 3.3. *Let $p = p(n)$, $u = u(n)$ are two series of positive integers depending on n such that $pu < n-1$ for all $n > 5$, and $p = p(n) \to \infty$, $u = u(n) \to \infty$ as $n \to \infty$. If $\frac{p^2}{n-1} \to 0$, then under $\boldsymbol{H}_0$, we have*

$$\frac{\log\Lambda-\tilde{\mu}_n}{\tilde{\sigma}_n}\xrightarrow{d}\boldsymbol{N}(0,1),$$

where

$$\begin{aligned}\tilde{\mu}_n &= u\left[2(n-1)-p-\frac{1}{2}\right]\log\left(1-\frac{p}{2n-2}\right)-(n-pu-\frac{3}{2})\log\left(1-\frac{pu}{n-1}\right),\\ \tilde{\sigma}_n^2 &= 4u\log\left(1-\frac{p}{2n-2}\right)-2\log\left(1-\frac{pu}{n-1}\right).\end{aligned}$$

4. Moderate deviation principle

In order to deeply investigate the convergence properties of the LR statistic under the null hypothesis, we will further consider the MDP of $\log\Lambda$. By Theorem 3.1, we can easy to see that for any $x > 0$,

$$\lim_{a\to\infty}\lim_{n\to\infty}\frac{1}{a^2}\log P\Big(\Big|\frac{\log\Lambda-\mu_{n,v}}{\sigma_{n,v}}\Big|\ge ax\Big)=-\frac{x^2}{2}.$$

Then a natural question is whether for any sequence $\{a_n\}$ with $\lim_{n\to\infty} a_n = \infty$, it also holds that

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log P\Big(\Big|\frac{\log\Lambda - \mu_{n,v}}{\sigma_{n,v}}\Big| \geq a_n x\Big) = -\frac{x^2}{2}. \tag{4.1}$$

In fact, (4.1) extends the conventional local asymptotic analysis for $\log\Lambda$, focusing on σ_n neighborhoods, *i.e.*, $P(|\log\Lambda - \mu_{n,v}| \geq \sigma_{n,v}x)$, to the moderate deviation region, focusing on $a_n\sigma_n$ neighborhoods, *i.e.*, $P(|\log\Lambda - \mu_{n,v}| \geq a_n\sigma_{n,v}x)$. In general, (4.1) is called the moderate deviation estimation or more generally MDP, which is introduced systematically by Dembo and Zeitouni [9].

If we use $\log\Lambda$ to construct the rejection region of the test problem (1.2), then the decay rate of type I error can be determined by (4.1). In particular, if given type I error, (4.1) can be used to estimate the minimum sample sizes. Thus, from the viewpoint of the statistical cost of experiments, the MDP is also meaningful for the test problem.

Define

$$Z_n = \frac{1}{a_n}\frac{\log\Lambda - \mu_{n,v}}{\sigma_{n,v}}, \tag{4.2}$$

where a_n is the moderate deviation scale, *i.e.*, it is a sequence of positive numbers satisfying that $\lim_{n\to\infty} a_n = \infty$. Under the assumption of

$$p = p(n) \to \infty, \quad pu < n-1 \ \text{ and } \ \frac{pu}{n-1} \to y \in [0,1] \quad \text{as} \quad n\to\infty, \tag{4.3}$$

we can get the following result by the similar arguments in Jiang and Wang [13].

Theorem 4.1. *Under the assumption in (4.3), let a_n be a sequence of positive numbers tends to infinity and satisfy*

(*i*) *if* $\lim_{n\to\infty} \frac{pu}{n-1} = 1$, *then* $\limsup_{n\to\infty} \frac{a_n}{\sigma_{n,v}} = 0$,

(*ii*) *if* $\lim_{n\to\infty} \frac{pu}{n-1} = y \in (0,1]$, *then* $\lim_{n\to\infty} \frac{a_n}{n} = 0$,

(*iii*) *if* $\lim_{n\to\infty} \frac{pu}{n-1} = 0$, *then* $\lim_{n\to\infty} \frac{a_n}{pu} = 0$.

Then under $\boldsymbol{H}_0$, the statistic Z_n defined in (4.2) satisfies the moderate deviation principle with speed a_n^2 and good rate function $I(x) = \frac{x^2}{2}$, i.e., for any fixed $x \geq 0$, we have

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log P(|Z_n| \geq x) = -\frac{x^2}{2}.$$

Remark 4.2. If we take $\{|(\log\Lambda - \mu_{n,v})/a_n\sigma_{n,v}| \geq c\}$ to be the rejection region for testing (1.2), where c is a constant. Then the type I error can be read as

$$\alpha_n = P\Big(\Big|\frac{\log\Lambda - \mu_{n,v}}{a_n\sigma_{n,v}}\Big| \geq c|\boldsymbol{H}_0\Big).$$

According to Theorem 4.1, for any fixed $c \geq 0$, we can get that $\lim_{n\to\infty} \frac{1}{a_n^2}\log\alpha_n = -\frac{c^2}{2}$, which implies the decay rate of the type I error is

$$\alpha_n = \exp\left(-\frac{c^2a_n^2}{2}\right)(1+o(1))$$

as $n\to\infty$.

Remark 4.3. For the moderate deviation result in Theorem 4.1, the rate of a_n should be controlled by $\sigma_{n,v}$ or n, and the rate of $\sigma_{n,v}$ is related to the convergence of $\frac{pu}{n-1}$, thus the assumption in (4.3) is needed.

5. Numerical simulations

In this section, we will first investigate the behavior of our LR test method on dealing with divergent-dimensional BCS test by numerical simulations. The proposed divergent-dimensional likelihood ratio (DLR) method in Theorem 3.1 and the traditional Chi-square approximation method (TCA) in (3.1) will be compared under the BCS test with divergent dimension. At last, some simulations will also conducted to verify the MDP result of the proposed LR test statistics in Theorem 4.1.

Now begin with the comparison of the empirical sizes of the TCA and our DLR method. Select $\boldsymbol{\mu} = \mathbf{0}_{pu\times 1}$ and

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix}_{p\times p},$$

and for $u = 2m$ or $u = 2m+1$ for $m \in \mathbb{N}$,

$$\boldsymbol{\Sigma}_i = \frac{1}{2^i}\boldsymbol{\Sigma}_0, \quad i = 1, 2, \ldots, m.$$

Let the simple size $n = 100$ and all the simulation result are based on 10 000 independent repetition calculations. In order to get the comprehensive performance of the two methods, we will set the five different significance levels $\alpha = 0.01, 0.05, 0.5, 0.95$ and 0.99 respectively.

Choosing different values of u and p, we can get the empirical sizes of the TCA method and the DLR method in Tables 1 and 2 respectively. We can see that when the product dimension pu of block dimension p and point u is small (less than 30 for example), the empirical sizes of the TCA method always close to the given significance level α, and it's performance is much better than our DLR method. However, when pu becoming large, the empirical sizes of the TCA method will no longer match the given significant level well, and the sizes of our DLR method are still very close to the given significance levels, then we can say that the proposed DLR method is better than the traditional TCA method when dealing with the divergent-dimensional BCS test.

We also give the histograms of the two test statistics in (3.1) and Theorem 3.1 in Figures 1 and 2. We will choose $n = 100$ and $pu = 6, 18, 30, 60, 90$ respectively. In Figure 1, we can see that when pu is small, the histograms of $-\rho(n-1)\log\Lambda$ in (3.1) match the χ^2 density curve very well, however, when pu is large, the histogram of $-\rho(n-1)\log\Lambda$ match the χ^2 density curve badly. As for the histogram of the test statistic $(\log\Lambda - \mu_{n,v})/\sigma_{n,v}$ in Theorem 3.1 and the standard normal density curve, Figure 2 reveals that the histogram of $(\log\Lambda - \mu_{n,v})/\sigma_{n,v}$ matches the standard normal density curve very well when pu is large, which also reveals that our DLR method is better than the TCA method in the divergent-dimensional test.

By the simulations above, we can see that the proposed DLR method in Theorem 3.1 performs well in the divergent-dimensional test.

At last, we will conduct some numerical simulations to verify the moderate deviation result in Theorem 4.1. We will take the moderate deviation scale $a_n = \sqrt{n}$ and define

$$P(x) = \frac{1}{T}\sum_{k=1}^{T} I_{\{|\log\Lambda^{(k)} - \mu_{n,v}| \geq a_n\sigma_{n,v}x\}},$$

$$Q(x) = \exp\left(-\frac{x^2 a_n^2}{2}\right)$$

TABLE 1. Empirical sizes of TCA method.

pu	p	u	The significance level α				
			0.01	0.05	0.50	0.95	0.99
4	2	2	0.0099	0.0503	0.5028	0.9547	0.9915
6	2	3	0.0113	0.0524	0.4968	0.9510	0.9907
	3	2	0.0010	0.0521	0.5101	0.9537	0.9896
8	2	4	0.0098	0.0487	0.4986	0.9515	0.9905
	4	2	0.0106	0.0533	0.5043	0.9509	0.9899
10	2	5	0.0108	0.0503	0.5001	0.9478	0.9896
	5	2	0.0102	0.0517	0.5078	0.9547	0.9918
12	3	4	0.0101	0.0494	0.4935	0.9524	0.9886
	4	3	0.0088	0.0479	0.5013	0.9518	0.9898
15	3	5	0.0102	0.0527	0.5003	0.9478	0.9893
	5	3	0.0103	0.0546	0.5125	0.9526	0.9905
18	3	6	0.0101	0.0493	0.5056	0.9500	0.9897
	6	3	0.0125	0.0593	0.5183	0.9539	0.9925
	4	5	0.0108	0.0546	0.5085	0.9497	0.9890
20	5	4	0.0117	0.0514	0.5103	0.9545	0.9923
	10	2	0.0118	0.0549	0.5137	0.9514	0.9906
	5	6	0.0105	0.0548	0.5216	0.9551	0.9908
30	10	3	0.0176	0.0743	0.5795	0.9668	0.9944
	15	2	0.0123	0.0598	0.5433	0.9595	0.9928
	5	8	0.0152	0.0696	0.5444	0.9607	0.9919
40	10	4	0.0228	0.0897	0.6097	0.9716	0.9941
	20	2	0.0188	0.0801	0.5833	0.9646	0.9926
	5	10	0.0215	0.0911	0.6135	0.9711	0.9946
50	10	5	0.0479	0.1564	0.7224	0.9842	0.9977
	25	2	0.0298	0.1170	0.6566	0.9797	0.9959
	5	12	0.0561	0.1706	0.7435	0.9987	0.9984
60	10	6	0.1103	0.2874	0.8472	0.9952	0.9998
	20	3	0.1905	0.4105	0.9110	0.9981	0.9997
	5	14	0.1969	0.4167	0.9092	0.9984	0.9998
70	10	7	0.3656	0.6201	0.9685	0.9997	1
	14	5	0.4223	0.6758	0.9748	1	1
	5	16	0.6767	0.8562	0.9948	1	1
80	10	8	0.8588	0.9584	0.9997	1	1
	20	4	0.9067	0.9740	0.9995	1	1
	6	15	0.9988	0.9997	1	1	1
90	10	9	0.9998	1	1	1	1
	30	3	1	1	1	1	1

for all $x \geq 0$, where T is the independent running times of the numerical simulation and $\Lambda^{(k)}(k = 1, \ldots, T)$ is the sample value of the statistics Λ in the kth independent simulation. Under different choices of the parameters p and u satisfy $\frac{pu}{n-1} < 1$, for a fixed large n, the proximity of $P(x)$ and $Q(x)$ will be observed. If we take $T = 10\,000$, $n = 100$ and choose $pu = 9$, 27 and 81, then the curves of $P(x)$ and $Q(x)$ for $x \in [0, 1]$ are plotted in the left, middle and right panel of Figure 3, respectively.

By Figure 3, we can see that the red solid line (stands for $Q(x)$) and the green dashed line (stands for $P(x)$) are always very close and both rapidly tend to zero when x increases, this confirms the moderate deviation result in Theorem 4.1. By the way, since $P(x)$ shows the empirical value of the exponential order decay rate of

TABLE 2. Empirical sizes of DLR method.

pu	p	u	The significance level α				
			0.01	0.05	0.50	0.95	0.99
4	2	2	0.0239	0.0492	0.4833	0.9522	0.9905
6	2	3	0.0205	0.0527	0.5123	0.9504	0.9897
	3	2	0.0178	0.0465	0.4984	0.9501	0.9899
8	2	4	0.0135	0.0480	0.5103	0.9537	0.9908
	4	2	0.0153	0.0475	0.5004	0.9461	0.9878
10	2	5	0.0141	0.0554	0.5081	0.9516	0.9909
	5	2	0.0134	0.0451	0.4950	0.9459	0.9896
12	3	4	0.0122	0.0494	0.5125	0.9496	0.9917
	4	3	0.0102	0.0479	0.5026	0.9561	0.9913
15	3	5	0.0118	0.0557	0.5083	0.9518	0.9909
	5	3	0.0108	0.0544	0.5071	0.9497	0.9905
18	3	6	0.0114	0.0514	0.5056	0.9492	0.9893
	6	3	0.0125	0.0544	0.5083	0.9527	0.9914
	4	5	0.0129	0.0554	0.5157	0.9529	0.9915
20	5	4	0.0111	0.0487	0.5133	0.9514	0.9906
	10	2	0.0107	0.0484	0.5061	0.9513	0.9917
	5	6	0.0109	0.0508	0.5112	0.9492	0.9889
30	10	3	0.0103	0.0542	0.5041	0.9510	0.9910
	15	2	0.0088	0.0498	0.5058	0.9509	0.9904
	5	8	0.0107	0.0573	0.5031	0.9531	0.9905
40	10	4	0.0110	0.0526	0.5009	0.9506	0.9899
	20	2	0.0123	0.0546	0.5116	0.9540	0.9925
	5	10	0.0114	0.0507	0.4959	0.9521	0.9890
50	10	5	0.0121	0.0557	0.5100	0.9528	0.9895
	25	2	0.0118	0.0501	0.5052	0.9510	0.9909
	5	12	0.0117	0.0535	0.5046	0.9481	0.9893
60	10	6	0.0104	0.0520	0.5039	0.9493	0.9992
	20	3	0.0114	0.0531	0.5032	0.9525	0.9910
	5	14	0.0110	0.0523	0.5080	0.9526	0.9917
70	10	7	0.0113	0.0513	0.5032	0.9514	0.9897
	14	5	0.0111	0.0573	0.4997	0.9496	0.9881
	5	16	0.0110	0.0518	0.5008	0.9517	0.9892
80	10	8	0.0121	0.0503	0.4977	0.9485	0.9881
	20	4	0.0110	0.0538	0.5015	0.9521	0.9902
	6	15	0.0120	0.0535	0.5019	0.9514	0.9904
90	10	9	0.0118	0.0513	0.5039	0.9546	0.9912
	30	3	0.0116	0.0569	0.5063	0.9520	0.9886

the deviation probability and there is only slightly difference between $P(x)$ under different values of pu, then the difference between the three graphics is very small.

6. Conclusion

The paper considers the LR test on the block circular symmetric covariance structure of a multivariate Gaussian population with divergent dimension. We firstly give the arbitrary order moments of the LR statistic under the null hypothesis. Then, with the help of the asymptotic expansion of high-order Gamma function, the asymptotical normality of the logarithmic LR statistic is proved under the assumption of $p \to \infty$ and $pu < n - 1$ along with $n \to \infty$. At last, the MDP of the logarithmic LR statistic is also established. The

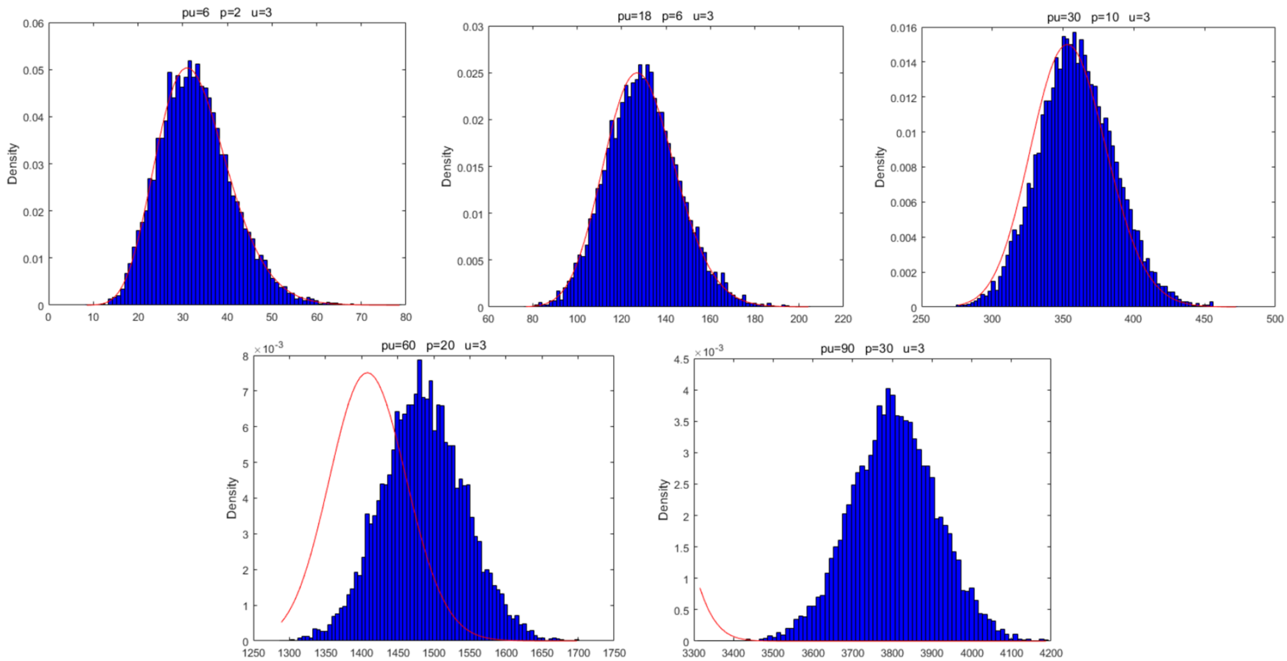

FIGURE 1. Comparison between the histogram of $-\rho(n-1)\log\Lambda$ in (3.1) and the χ^2 distribution density curve.

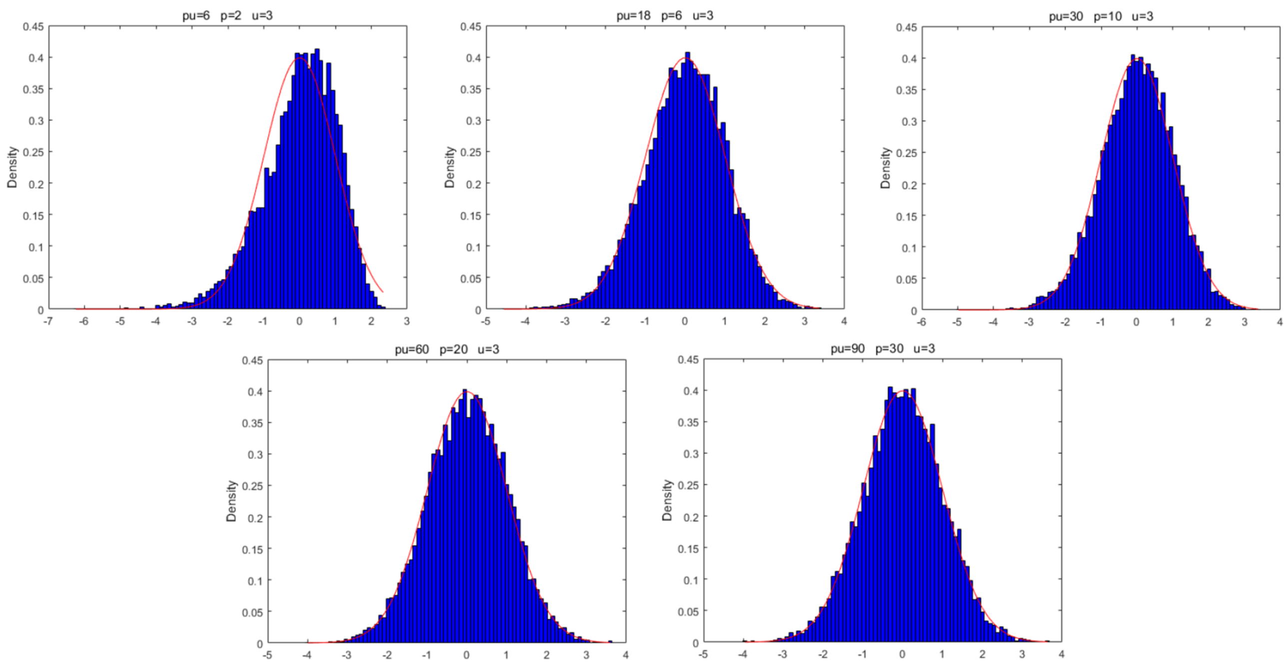

FIGURE 2. Comparison between the histograms of $(\log\Lambda-\mu_{n,v})/\sigma_{n,v}$ and the standard normal density curve.

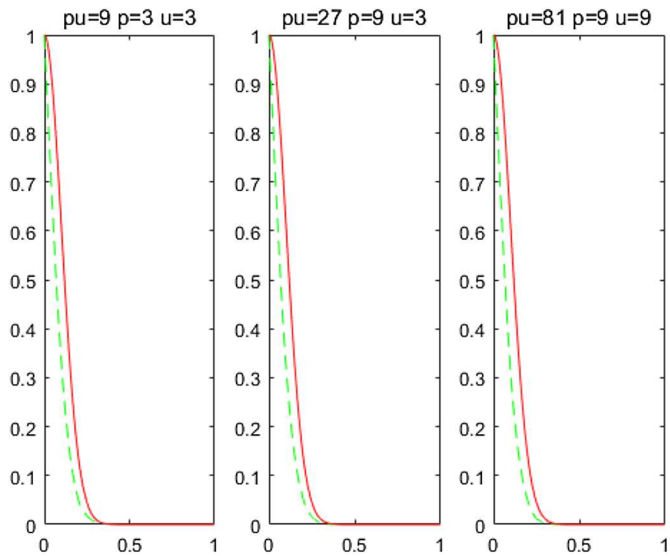

FIGURE 3. The proximity of $P(x)$ and $Q(x)$ for $x \in [0, 1]$.

numerical simulations reveal that the proposed divergent-dimensional LR method performs well, and much better than traditional Chi-square approximation method in the BCS covariance structure test with divergent dimension.

The proposed divergent-dimensional LR test method is based on the Gaussian population, and we only consider the asymptotical distribution of the logarithmic LR test under the null hypothesis. Since the asymptotical distribution of the LR test statistic related to the specific structure of the covariance matrix $\mathbf{\Sigma}$, then we cannot get the power function unless the alternative covariance matrix is given. At the same time, our LR test method is based on the assumption $pu < n - 1$, how to deal with the corresponding test under the assumption of $pu > n - 1$ with $p \to \infty$ and establish the corresponding test method for general non-Gaussian population are also meaningful issues.

Appendix A. Technical proofs

The proofs of Theorems 2.1, 3.1 and 4.1 will be listed in this section. We will first establish the moments of the LR test statistic in Theorem 2.1. With the help of the continuity theorem of the moment generating function and the asymptotic expansion formula of the multivariate Gamma function, we can get the asymptotic distribution and moderate deviation of the LR test statistic in Theorems 3.1 and 4.1, respectively. Some ideas come from the results of Jiang and Yang [14], Jiang and Qi [15] and Jiang and Wang [13]. The major difference between the following proofs and those in the literature is that the actual calculations are more involved.

Proof of Theorem 2.1. By the definitions of Λ and $\boldsymbol{V}$ in (2.1) and (2.2) respectively, we have

$$\Lambda = \frac{2^{2p(u-m-1)}|(\boldsymbol{B} \otimes \boldsymbol{I}_p)||\boldsymbol{S}||(\boldsymbol{B}' \otimes \boldsymbol{I}_p)|}{\prod_{j=1}^{u} |\boldsymbol{V}_j|}. \tag{A.1}$$

For the numerator of (A.1), we have

$$|(\boldsymbol{B} \otimes \boldsymbol{I}_p)||\boldsymbol{S}||(\boldsymbol{B}' \otimes \boldsymbol{I}_p)| = |(\boldsymbol{B} \otimes \boldsymbol{I}_p)||\mathbf{\Sigma}_{\mathrm{BCS}}||(\boldsymbol{B}' \otimes \boldsymbol{I}_p)| \cdot \frac{|\boldsymbol{S}|}{|\mathbf{\Sigma}_{\mathrm{BCS}}|}. \tag{A.2}$$

By Olkin [28], we can rewrite

$$(\boldsymbol{B} \otimes \boldsymbol{I}_p)\boldsymbol{\Sigma}_{\mathrm{BCS}}(\boldsymbol{B}' \otimes \boldsymbol{I}_p) := \mathrm{Diag}(\boldsymbol{\Psi}_1, \dots, \boldsymbol{\Psi}_u),$$

where $\boldsymbol{\Psi}_j, j = 1, \dots, u$ are all $p \times p$ positive definite matrices satisfy

$$\boldsymbol{\Psi}_j = \boldsymbol{\Psi}_{u-j+2}, \quad j = 2, \dots, u.$$

Then we can obtain that

$$|(\boldsymbol{B} \otimes \boldsymbol{I}_p)||\boldsymbol{\Sigma}_{\mathrm{BCS}}||(\boldsymbol{B}' \otimes \boldsymbol{I}_p)| = \prod_{j=1}^{u} |\boldsymbol{\Psi}_j|. \tag{A.3}$$

Denote $|\boldsymbol{V}_j|/|\boldsymbol{\Psi}_j| = H_j$. We can see that H_j is related to the sample covariance matrix $\boldsymbol{S}$ with parameters n, p, u, that is, we can write $H_j = H_j(\boldsymbol{S})$. Inserting (A.2) and (A.3) into (A.1), we can see that

$$E[\Lambda^h] = 2^{2hp(u-m-1)} \cdot E\bigg(\prod_{j=1}^{u} H_j{}^{-1} \cdot \frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}_{\mathrm{BCS}}|}\bigg)^h. \tag{A.4}$$

Since $\boldsymbol{S}$ obeys the Wishart distribution, that is, $\boldsymbol{S} \sim \boldsymbol{W}_{pu}(n-1, \boldsymbol{\Sigma}_{\mathrm{BCS}})$, and the density function of $\boldsymbol{S}$ is

$$f_{n-1,p,u}(\boldsymbol{S}) = \frac{1}{2^{\frac{(n-1)pu}{2}} \Gamma_{pu}(\frac{n-1}{2}) |\boldsymbol{\Sigma}_{\mathrm{BCS}}|^{\frac{n-1}{2}}} |\boldsymbol{S}|^{\frac{n-pu-2}{2}} etr(-\frac{1}{2}\boldsymbol{\Sigma}_{\mathrm{BCS}}{}^{-1}\boldsymbol{S}), \quad \boldsymbol{S} > 0.$$

Then it follows by (A.4) that

$$\begin{aligned}
E[\Lambda^h] &= \int_{\boldsymbol{S}>0} 2^{2hp(u-m-1)} \cdot \prod_{j=1}^{u} H_j{}^{-h} \cdot \frac{|\boldsymbol{S}|^h}{|\boldsymbol{\Sigma}_{\mathrm{BCS}}|^h} \cdot f_{n-1,p,u}(\boldsymbol{S})\mathrm{d}\boldsymbol{S} \\
&= \int_{\boldsymbol{S}>0} 2^{2hp(u-m-1)} \cdot \prod_{j=1}^{u} H_j{}^{-h} \cdot \frac{|\boldsymbol{S}|^h}{|\boldsymbol{\Sigma}_{\mathrm{BCS}}|^h} \cdot \frac{|\boldsymbol{S}|^{\frac{n-pu-2}{2}} etr(-\frac{1}{2}\boldsymbol{\Sigma}_{\mathrm{BCS}}^{-1}\boldsymbol{S})}{2^{\frac{(n-1)pu}{2}} \Gamma_{pu}(\frac{n-1}{2}) |\boldsymbol{\Sigma}_{\mathrm{BCS}}|^{\frac{n-1}{2}}} \mathrm{d}\boldsymbol{S} \\
&= 2^{hp(3u-2m-2)} \frac{\Gamma_{pu}(\frac{n+2h-1}{2})}{\Gamma_{pu}(\frac{N-1}{2})} \int_{\boldsymbol{S}>0} \bigg(\prod_{j=1}^{u} H_j\bigg)^{-h} \frac{|\boldsymbol{S}|^{\frac{n+2h-pu-2}{2}} etr(-\frac{1}{2}\boldsymbol{\Sigma}_{\mathrm{BCS}}^{-1}\boldsymbol{S})}{2^{\frac{(n+2h-1)pu}{2}} \Gamma_{pu}(\frac{n+2h-1}{2}) |\boldsymbol{\Sigma}_{\mathrm{BCS}}|^{\frac{n+2h-1}{2}}} \mathrm{d}\boldsymbol{S} \\
&= 2^{hp(3u-2m-2)} \frac{\Gamma_{pu}(\frac{n+2h-1}{2})}{\Gamma_{pu}(\frac{n-1}{2})} \int_{\tilde{\boldsymbol{S}}>0} \bigg(\prod_{j=1}^{u} \tilde{H}_j\bigg)^{-h} f_{n+2h-1,p,u}(\tilde{\boldsymbol{S}})\mathrm{d}\tilde{\boldsymbol{S}},
\end{aligned}$$

where $h > -\frac{n-pu}{2}$. Here, $f_{n+2h-1,p,u}(\tilde{\boldsymbol{S}})$ is also a density function of sample covariance matrix $\tilde{\boldsymbol{S}}$ with parameters $n+2h-1$ and $\boldsymbol{\Sigma}_{\mathrm{BCS}}$, and it holds that $\mathrm{d}\tilde{\boldsymbol{S}} = \mathrm{d}\boldsymbol{S}$. $\tilde{H}_j$ is defined by $\tilde{H}_j = |\tilde{\boldsymbol{V}}_j|/|\boldsymbol{\Psi}_j|$, $\tilde{\boldsymbol{V}}_j$ share the same structure with $\boldsymbol{V}_j$, the only difference is $\tilde{\boldsymbol{V}}_j$ is generated by $\tilde{\boldsymbol{S}}$ with parameters $n+2h-1$ and $\boldsymbol{\Sigma}_{\mathrm{BCS}}$, while $\boldsymbol{V}_j$ is generated by $\boldsymbol{S}$ with parameters n and $\boldsymbol{\Sigma}_{\mathrm{BCS}}$.

Under $\boldsymbol{H}_0$, Olkin [28] proved that when $u = 2m+1$ is odd,

$$\tilde{\boldsymbol{V}}_1 \sim \boldsymbol{W}_p(n+2h-1, \boldsymbol{\Psi}_1), \quad \tilde{\boldsymbol{V}}_j \sim \boldsymbol{W}_p(2(n+2h-1), \boldsymbol{\Psi}_j), \quad j = 2, \dots, m+1,$$

when $u = 2m$ is even,

$$\tilde{\boldsymbol{V}}_1 \sim \boldsymbol{W}_p(n+2h-1, \boldsymbol{\Psi}_1), \quad \tilde{\boldsymbol{V}}_j \sim \boldsymbol{W}_p(2(n+2h-1), \boldsymbol{\Psi}_1), \quad j = 2, \ldots, m,$$

and

$$\tilde{\boldsymbol{V}}_{m+1} \sim \boldsymbol{W}_p(n+2h-1, \boldsymbol{\Psi}_{m+1}),$$

By the probability properties of Wishart distribution (see *e.g.* Thm. 3.2.15 in Muirhead [24]), we can get that when $u = 2m+1$ is even,

$$\tilde{H}_1 \sim \prod_{r=1}^{p} \chi^2_{n+2h-r}, \ \tilde{H}_j \sim \prod_{r=1}^{p} \chi^2_{2(n+2h-1)-r+1}, \ j = 2, \ldots, m, \tag{A.5}$$

and when $u = 2m$ is odd,

$$\tilde{H}_1 \sim \prod_{r=1}^{p} \chi^2_{n+2h-r}, \ \tilde{H}_j \sim \prod_{r=1}^{p} \chi^2_{2(n+2h-1)-r+1}, \ j = 2, \ldots, m, \ \tilde{H}_{m+1} \sim \prod_{r=1}^{p} \chi^2_{n+2h-r}, \tag{A.6}$$

where χ^2_{n+2h-r} and $\chi^2_{2(n+2h-1)-r+1}(r = 1, \ldots, p)$ are independent χ^2 random variables with degrees of freedom $n + 2h - r$ and $2(n + 2h - 1) - r + 1$, respectively.

Thus, we can write

$$E[\Lambda^h] = 2^{hp(3u-2m-2)} \frac{\Gamma_{pu}(\frac{n+2h-1}{2})}{\Gamma_{pu}(\frac{n-1}{2})} E\left[\prod_{j=1}^{u} \tilde{H}_j^{-h}\right]. \tag{A.7}$$

When $u = 2m+1$ is odd, we have

$$E\left[\prod_{j=1}^{u} \tilde{H}_j^{-h}\right] = E[\tilde{H}_1^{-h}]\big\{E[\tilde{H}_2^{-2h}]\big\}^m, \tag{A.8}$$

and when $u = 2m$ is even,

$$E\left[\prod_{j=1}^{u} \tilde{H}_j^{-h}\right] = \big\{E[\tilde{H}_1^{-h}]\big\}^2\big\{E[\tilde{H}_2^{-2h}]\big\}^{m-1}. \tag{A.9}$$

As the distributions of $\tilde{H}_j, j = 1, \ldots, u$ are given in (A.5) and (A.6), we have (see page 269 in Anderson [1])

$$E[\tilde{H}_1^{-h}] = 2^{-ph} \frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2h-1}{2})} \tag{A.10}$$

and

$$E[\tilde{H}_2^{-2h}] = 2^{-2ph} \frac{\Gamma_p(n-1)}{\Gamma_p(n+2h-1)}. \tag{A.11}$$

Combining (A.8)–(A.11), we know when $u = 2m+1$ is odd,

$$E\left[\prod_{j=1}^{u} \tilde{H}_j^{-h}\right] = 2^{-(2m+1)ph} \frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2h-1}{2})} \left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2h-1)}\right]^m, \tag{A.12}$$

and when $u = 2m$ is even,

$$E\left[\prod_{j=1}^{u} \tilde{H}_j^{-h}\right] = 2^{-2mph} \left[\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2h-1}{2})}\right]^2 \left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2h-1)}\right]^{m-1}. \tag{A.13}$$

Putting (A.12) and (A.13) into (A.7), we can reach to the fact that when $u = 2m+1$ is odd,

$$E[\Lambda^h] = 2^{2phm} \frac{\Gamma_{pu}(\frac{n+2h-1}{2})}{\Gamma_{pu}(\frac{N-1}{2})} \frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2h-1}{2})} \left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2h-1)}\right]^m,$$

and when $u = 2m$ is even,

$$E[\Lambda^h] = 2^{2ph(m-1)} \frac{\Gamma_{pu}(\frac{n+2h-1}{2})}{\Gamma_{pu}(\frac{n-1}{2})} \left[\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2h-1}{2})}\right]^2 \left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2h-1)}\right]^{m-1},$$

which imply Theorem 2.1 immediately. □

In the sequel, we will introduce a useful lemma under a slightly stronger assumption of

$$p = p(n) \to \infty \quad \text{and} \quad \frac{pu}{n-1} \to y \in (0,1] \quad as \quad n \to \infty. \tag{A.14}$$

Lemma A.1. *Assume that $p = p(n)$ is a series of positive integers depending on n such that $pu < n-1$ for all $n > 5$ and the assumption (A.14) holds. Let Λ be defined as in (2.1), then under $\boldsymbol{H}_0$, we have*

$$\frac{\log \Lambda - \mu_{n,v}}{\sigma_{n,v}} \xrightarrow{d} \boldsymbol{N}(0,1),$$

where

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

and v is defined in (2.3).

Before the proof of Lemma A.1, we will first introduce the asymptotic expansion formula of the multivariate Gamma function, it can be seen from Lemma 5.4 of Jiang and Yang [14].

Lemma A.2. *Let $n > p = p_n$ and $\gamma_n = [-\log(1-\frac{p}{n})]^{\frac{1}{2}}$. Assume $\frac{p}{n} \to y \in (0,1]$ and $s = s_n = O\left(\frac{1}{\gamma_n}\right)$ and $t = t_n = O\left(\frac{1}{\gamma_n}\right)$ as $n \to \infty$, Then, as $n \to \infty$, we have*

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

Proof of Lemma A.1. Let $\boldsymbol{Y} \sim \boldsymbol{N}(0,1)$, then the moment generating function $Ee^{s\boldsymbol{Y}} = e^{\frac{s^2}{2}}$. By the continuity theorem of the moment generating function stated in Theorem 9.5 in Gut [11], we know that $\frac{\log\Lambda - \mu_{n,v}}{\sigma_{n,v}} \xrightarrow{d} \boldsymbol{N}(0,1)$ can be proved by showing that there exists $\delta > 0$ such that

$$E\exp\left(\frac{\log\Lambda - \mu_{n,v}}{\sigma_{n,v}} s\right) \to e^{\frac{s^2}{2}}$$

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

By the assumption $\frac{pu}{n-1} \to y \in (0,1]$ in (A.14), we can see from the expression of $\sigma_{n,v}^2$ that

$$\sigma_{n,v}^2 \to 2[v\log\left(1-\frac{y}{u}\right) + 2(u-v)\log\left(1-\frac{y}{2u}\right) - \log(1-y)] > 0 \tag{A.15}$$

as $n \to \infty$ for $y \in (0,1)$, and $\sigma_{n,v}^2 \to +\infty$ as $n \to \infty$ for $y = 1$. Thus we can define

$$\delta_0 := \inf\{\sigma_{n,v} : n > 5\} > 0.$$

Fix $|s| < \frac{\delta_0}{2}$ and set $t = t_n = s/\sigma_{n,v}$, then $\{t_n : n > 5\}$ is bounded and $|t_n| < \frac{1}{2}$ for all $n > 5$. By Theorem 2.1, we have

$$\begin{aligned} Ee^{t\log\Lambda} &= E\Lambda^t \\ &= 2^{tp(u-v)} \frac{\Gamma_{pu}(\frac{n+2t-1}{2})}{\Gamma_{pu}(\frac{n-1}{2})} \left[\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n+2t-1}{2})}\right]^v \left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2t-1)}\right]^{\frac{u-v}{2}} \end{aligned} \tag{A.16}$$

for all $n > 5$.

Set $\alpha_n = [-\log\left(1-\frac{pu}{n-1}\right)]^{\frac{1}{2}}$ for $n-1 > pu > 4$. When $n \to \infty$, we have

$$\begin{aligned} t^2\alpha_n^2 &= \frac{s^2}{\sigma_{n,v}^2}\left[-\log\left(1-\frac{pu}{n-1}\right)\right] \\ &\to \begin{cases} \dfrac{s^2}{2} \cdot \dfrac{\log(1-y)}{\log(1-y) - v\log\left(1-\frac{y}{u}\right) - 2(u-v)\log\left(1-\frac{y}{2u}\right)} & y \in (0,1), \\ \dfrac{s^2}{2} & y = 1, \end{cases} \end{aligned}$$

and then $t = O\left(\frac{1}{\alpha_n}\right)$ as $n \to \infty$.

By Lemma A.2 and the assumption that $\frac{pu}{n-1} \to y \in (0,1]$, we can get that

$$\begin{aligned}
&\log \frac{\Gamma_{pu}\left(\frac{n-1}{2}+t\right)}{\Gamma_{pu}(\frac{n-1}{2})} \\
&= tpu\left(\log\frac{n-1}{2}-1\right)+\gamma_n^2\left[t^2-t\left(pu-n+\frac{3}{2}\right)\right]+o(1) \\
&= tpu\left(\log\frac{n-1}{2}-1\right)-\log\left(1-\frac{pu}{n-1}\right)\left[t^2-t\left(pu-n+\frac{3}{2}\right)\right]+o(1) \\
&= -tpu+tpu\log\frac{n-1}{2}+t\left(pu-n+\frac{3}{2}\right)\log\left(1-\frac{pu}{n-1}\right)-t^2\log\left(1-\frac{pu}{n-1}\right)+o(1).
\end{aligned} \tag{A.17}$$

Similarly, set $\beta_n = [-\log\left(1-\frac{p}{n-1}\right)]^{\frac{1}{2}}$ and $\gamma_n = [-\log\left(1-\frac{p}{2(n-1)}\right)]^{\frac{1}{2}}$ for $n-1 > pu > 4$. We also have $t = O\left(\frac{1}{\beta_n}\right)$ and $t = O\left(\frac{1}{\gamma_n}\right)$ as $n \to \infty$. Then, by Lemma A.2 again, we have

$$\begin{aligned}
&\log\left[\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n-1}{2}+t)}\right]^v = -v\log\frac{\Gamma_p(\frac{n-1}{2}+t)}{\Gamma_p(\frac{n-1}{2})} \\
&= tpv-tpv\log\frac{n-1}{2}+tv\left(n-p-\frac{3}{2}\right)\log\left(1-\frac{p}{n-1}\right)+t^2v\log\left(1-\frac{p}{n-1}\right)+o(1)
\end{aligned} \tag{A.18}$$

and

$$\begin{aligned}
&\log\left[\frac{\Gamma_p(n-1)}{\Gamma_p(n+2t-1)}\right]^{\frac{u-v}{2}} = -\frac{u-v}{2}\log\frac{\Gamma_p(\frac{2n-2}{2}+2t)}{\Gamma_p(\frac{2n-2}{2})} \\
&= tp(u-v)-tp(u-v)\log n-1-t(u-v)\left[2(n-1)-p-\frac{1}{2}\right]\log\left(1-\frac{p}{2n-2}\right) \\
&\quad +2(u-v)t^2\log\left(1-\frac{p}{n-1}\right)+o(1).
\end{aligned} \tag{A.19}$$

Putting (A.17)–(A.19) into (A.16), we can see

$$\begin{aligned}
&\log Ee^{t\log\Lambda} \\
&= \left\{v\left(n-p-\frac{3}{2}\right)\log\left(1-\frac{p}{n-1}\right)+(u-v)\left[2(n-1)-p-\frac{1}{2}\right]\log\left(1-\frac{p}{2n-2}\right)\right. \\
&\quad \left.-\left(n-pu-\frac{3}{2}\right)\log\left(1-\frac{pu}{n-1}\right)\right\}t+\left[v\log\left(1-\frac{p}{n-1}\right)\right. \\
&\quad \left.+2(u-v)\log\left(1-\frac{p}{2n-2}\right)-\log\left(1-\frac{pu}{n-1}\right)\right]t^2+o(1) \\
&= \mu_{n,v}t+\frac{\sigma_{n,v}^2t^2}{2}+o(1),
\end{aligned}$$

where

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

Thus, we can reach to

$$\begin{aligned}\log Ee^{t\log\Lambda} &= \mu_{n,v}t+\frac{\sigma_{n,v}^2t^2}{2}+o(1)\\ &= \frac{s^2}{2}+\frac{\mu_{n,v}}{\sigma_{n,v}}s+o(1),\end{aligned} \tag{A.20}$$

that is,

$$E\exp\left(\frac{\log\Lambda-\mu_{n,v}}{\sigma_{n,v}}s\right)\to e^{\frac{s^2}{2}}$$

as $n\to\infty$ for all $|s|<\frac{\delta_0}{2}$. Then the proof of the lemma is completed. □

Now we come to the proof of Theorem 3.1. In the following proofs, a result form Proposition 5.1 of Jiang and Qi [15] will be introduced.

Lemma A.3. *Let $\{p=p_n\in\mathbb{N};n\geq 1\}$, $\{m=m_n\in\mathbb{N};n\geq 1\}$, $\{t_n\in\mathbb{R};n\geq 1\}$, satisfy that (i) $p_n\to\infty$, $p_n=o(n)$; (ii) there exists $\epsilon\in(0,1)$ such that $\epsilon\leq\frac{m_n}{n}\leq\epsilon^{-1}$ for large n; (iii) $t=t_n=O\left(\frac{n}{p}\right)$. Then, when $n\to\infty$, we have*

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

where

$$\begin{aligned}\alpha_n &= -\left[2p+\left(m-p-\frac{3}{2}\right)\log\left(1-\frac{p}{m-1}\right)\right],\\ \beta_n &= -\left[\frac{p}{m-1}+\log\left(1-\frac{p}{m-1}\right)\right],\\ \gamma_n(t) &= p\left[(\frac{m-1}{2}+t)\log\left(\frac{m-1}{2}+t\right)-\frac{m-1}{2}\log\frac{m-1}{2}\right].\end{aligned}$$

Proof of Theorem 3.1. Similar to the arguments in Jiang and Qi [15], we will adopt the subsequence strategy to complete the proof. By Bolzano-Weierstrass theorem, we know that a real sequence converges if and only if every subsequence contains a convergent further subsequence, and this subsequence principle also holds for random variables convergence in distribution (see Sect. 5.7 of Gut [11] or Lem. 8.2.2 (iii) of Chow and

Teicher [7]). Thus, the statment

$$H_n := \frac{\log \Lambda - \mu_{n,v}}{\sigma_{n,v}} \xrightarrow{d} \boldsymbol{N}(0,1), \tag{A.21}$$

will be proved by showing that for any subsequence $\{H_{n_k}\}$, there is a further subsequence $\{H_{n_{k_j}}\}$ such that $H_{n_{k_j}} \xrightarrow{d} \boldsymbol{N}(0,1)$ as $j \to \infty$.

Since $\frac{p_n u_n}{N} \in [0,1]$ for all n, then for any subsequence $\{n_k\}$, we can select a further subsequence $\{n_{k_j}\}$ such that $\frac{p_{n_{k_j}} u_{n_{k_j}}}{n_{k_j}} \to y \in [0,1]$ as $j \to \infty$. For the sake of simplicity and without loss of generality, we only need to show (A.21) under the assumption that $\frac{p_n u_n}{n} \to y \in [0,1]$ as $n \to \infty$. Note that when $\frac{p_n u_n}{n} \to y \in (0,1]$ as $n \to \infty$, (A.21) has been proved in Lemma A.1. Thus, we only need to prove (A.21) under the assumption of $\frac{p_n u_n}{n} \to 0$ as $n \to \infty$.

By the continuity theorem of the moment generating function again, it suffices to show that

$$E \exp\Big(\frac{\log \Lambda - \mu_{n,v}}{\sigma_{n,v}} s\Big) \to e^{\frac{s^2}{2}} \tag{A.22}$$

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

$$\log E e^{t \log \Lambda} = \mu_{n,v} t + \frac{\sigma_{n,v}^2 t^2}{2} + o(1)$$

as $n \to \infty$ with $|\sigma_{n,v} t| \le 1$, where $t = t_n = \frac{s}{\sigma_{n,v}}$.

Rewrite (A.16) as

$$\begin{aligned} &\log E e^{t \log \Lambda} = \log E \Lambda^t \\ &= tp(u-v)\log 2 + \log \frac{\Gamma_{pu}(\frac{n-1}{2}+t)}{\Gamma_{pu}(\frac{n-1}{2})} + v \log \frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n-1}{2}+t)} + \frac{u-v}{2} \log \frac{\Gamma_p(\frac{2n-2}{2})}{\Gamma_p(\frac{2n-2}{2}+2t)} \end{aligned} \tag{A.23}$$

for all $n > 5$.

By the assumption of $\frac{pu}{n-1} \to 0$, it is easy to see that $\sigma_{n,v}^2 \sim \frac{p^2 u^2}{(n-1)^2}$, which means $\frac{tpu}{n-1}$ is bounded, then we can write $t = O\left(\frac{n}{pu}\right)$ and $t = O\left(\frac{n}{p}\right)$. By Lemma A.3, we have

$$\begin{aligned} \log \frac{\Gamma_{pu}(\frac{n-1}{2}+t)}{\Gamma_{pu}(\frac{n-1}{2})} &= -2tpu - t\left(n - pu - \frac{3}{2}\right)\log\left(1 - \frac{pu}{n-1}\right) - \left[\frac{pu}{n-1} + \log\left(1 - \frac{pu}{n-1}\right)\right] t^2 \\ &\quad + pu\left[\left(\frac{n-1}{2}+t\right)\log\left(\frac{n-1}{2}+t\right) - \frac{n-1}{2}\log\frac{n-1}{2}\right] + o(1), \end{aligned} \tag{A.24}$$

$$\begin{aligned} v \log \frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n-1}{2}+t)} &= 2tpv + v\left(n - p - \frac{3}{2}\right) t \log\left(1 - \frac{p}{n-1}\right) + v\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right] t^2 \\ &\quad - vp\left[\left(\frac{n-1}{2}+t\right)\log\left(\frac{n-1}{2}+t\right) - \frac{n-1}{2}\log\frac{n-1}{2}\right] + o(1) \end{aligned} \tag{A.25}$$

and

$$\begin{aligned}
&\frac{u-v}{2}\log\frac{\Gamma_p(\frac{2n-2}{2})}{\Gamma_p(\frac{2n-2}{2}+2t)}\\
&=2tp(u-v)+t(u-v)\left[2(n-1)-p-\frac{1}{2}\right]\log\left(1-\frac{p}{2n-2}\right)\\
&\quad+2(u-v)\left[\frac{p}{2n-2}+\log\left(1-\frac{p}{2n-2}\right)\right]t^2\\
&\quad-\frac{p(u-v)}{2}\left[(n+2t-1)\log(n+2t-1)-(n-1)\log(n-1)\right]+o(1)\\
&=2tp(u-v)+t(u-v)\left[2(n-1)-p-\frac{1}{2}\right]\log\left(1-\frac{p}{2n-2}\right)\\
&\quad-tp(u-v)\log 2+2(u-v)\left[\frac{p}{2n-2}+\log\left(1-\frac{p}{2n-2}\right)\right]t^2\\
&\quad-p(u-v)\left[\left(\frac{n-1}{2}+t\right)\log\left(\frac{n-1}{2}+t\right)-\frac{n}{2}\log\frac{n-1}{2}\right]+o(1).
\end{aligned}\tag{A.26}$$

Combining the relations (A.23)–(A.26), we can reach to (A.20) again and then (A.22) is proved. This completes the proof of Theorem 3.1. □

At last, we will give the proof of Theorem 4.1. Two results appeared in Jiang and Wang [13] are needed. The results are similar to Lemmas A.2 and A.3, but they can't be replaced by them.

Lemma A.4. *Let λ_n, $n\geq 1$ be a sequence of positive numbers satisfying $\lambda_n\to\infty$, $\frac{\lambda_n}{n}\to 0$ as $n\to\infty$. Moreover, we assume that $p\to\infty$, $\frac{p}{n}\to y\in(0,1]$ as $n\to\infty$. Then, for any $\mu\in\mathbb{R}$, as $n\to\infty$, we have*

$$\begin{aligned}
\log\frac{\Gamma_p(n+\mu\lambda_n)}{\Gamma_p(n)}&=\sum_{i=1}^{p}\left[\log\left(n-\frac{i-1}{2}\right)-\frac{1}{2n+1-i}\right]\lambda_n\mu\\
&\quad+\sum_{i=1}^{p}\frac{\lambda_n^2\mu^2}{2n+1-i}+\max\left\{O\left(\frac{1}{n}\right),O\left(\frac{\lambda_n^3}{n^2}\right)\right\}.
\end{aligned}$$

Lemma A.5. *Let $n>p$, $\frac{p}{n}\to y\in(0,1]$ and $r_n=[-\log\left(1-\frac{p}{n}\right)]^{\frac{1}{2}}$. For any $t,s=o(1)$ as $n\to\infty$, we have*

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

as $n\to\infty$.

Proof of Theorem 4.1. According to the Gärtner-Eillis theorem (see Dembo and Zeitouni [9]), we only need to show

$$\Psi_n(\lambda)=\frac{1}{a_n^2}\log E\left[\exp\left(\lambda a_n\frac{\log\Lambda-\mu_{n,v}}{\sigma_{n,v}}\right)\right]\to\frac{\lambda^2}{2}\tag{A.27}$$

as $n\to\infty$ for any fixed $\lambda\in\mathbb{R}$.

Let $\lambda_n = \lambda a_n/\sigma_{n,v}$, then (A.27) can be rewritten as

$$\Psi_n(\lambda) = \frac{1}{a_n^2}\left(\log E\Lambda^{\lambda_n} - \lambda_n \mu_{n,v}\right) \to \frac{\lambda^2}{2}. \tag{A.28}$$

When $\frac{pu}{n-1} \to 1$ as $n \to \infty$, according to the assumption $a_n/\sigma_{n,v} \to 0$, which means $\lambda_n \to 0$ as $n \to \infty$. By Lemma A.1 and Lemma A.5, we can easy to see that

$$\Psi_n(\lambda) = \frac{1}{a_n^2}\Big(\lambda_n\mu_{n,v} + \frac{\sigma_{n,v}^2\lambda_n^2}{2} - \lambda_n\mu_{n,v}\Big) + o(1) = \frac{\lambda^2}{2} + o(1)$$

as $n \to \infty$. This implies (A.28) immediately.

When $\frac{pu}{n-1} \to y \in (0,1)$ as $n \to \infty$, it follows by (A.16) that

$$\begin{aligned}
&\log E\Lambda^{\lambda_n} \\
&= \lambda_n p(u-v)\log 2 + \log\frac{\Gamma_{pu}(\frac{n-1}{2}+\lambda_n)}{\Gamma_{pu}(\frac{n}{2})} + v\log\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n-1}{2}+\lambda_n)} + \frac{u-v}{2}\log\frac{\Gamma_p(n-1)}{\Gamma_p(n-1+2\lambda_n)}
\end{aligned}$$

for all $n > 5$.

By Lemma A.4, we know

$$\begin{aligned}
\log\frac{\Gamma_{pu}(\frac{n-1}{2}+\lambda_n)}{\Gamma_{pu}(\frac{n-1}{2})} &= \sum_{i=1}^{pu}\Big[\log\left(\frac{n-1}{2}-\frac{i-1}{2}\right) - \frac{1}{n-i}\Big]\lambda_n \\
&\quad + \sum_{i=1}^{pu}\frac{\lambda_n^2}{n-i} + \max\left\{O\left(\frac{1}{n}\right), O\left(\frac{\lambda_n^3}{n^2}\right)\right\} \\
&= \lambda_n pu\log\frac{n-1}{2} + \lambda_n\sum_{i=1}^{pu}\Big[\log\left(1-\frac{i-1}{n-1}\right) - \frac{1}{n-i}\Big] \\
&\quad + \lambda_n^2\sum_{i=1}^{pu}\frac{1}{n-i} + \max\left\{O\left(\frac{1}{n}\right), O\left(\frac{\lambda_n^3}{n^2}\right)\right\}.
\end{aligned} \tag{A.29}$$

Similarly, we also have

$$\begin{aligned}
v\log\frac{\Gamma_p(\frac{n-1}{2})}{\Gamma_p(\frac{n-1}{2}+\lambda_n)} &= -\lambda_n v\sum_{i=1}^{p}\Big[\log\left(\frac{n-1}{2}-\frac{i-1}{2}\right) - \frac{1}{n-i}\Big] \\
&\quad - v\sum_{i=1}^{p}\frac{\lambda_n^2}{n-i} + \max\left\{O\left(\frac{1}{n}\right), O\left(\frac{\lambda_n^3}{n^2}\right)\right\} \\
&= -\lambda_n pv\log\frac{n-1}{2} - \lambda_{n-1}v\sum_{i=1}^{p}\Big[\log\left(1-\frac{i-1}{n-1}\right) - \frac{1}{n-i}\Big] \\
&\quad - \lambda_n^2 v\sum_{i=1}^{p}\frac{1}{n-i} + \max\left\{O\left(\frac{1}{n}\right), O\left(\frac{\lambda_n^3}{n^2}\right)\right\}
\end{aligned} \tag{A.30}$$

and

$$\begin{aligned}
&\frac{u-v}{2}\log\frac{\Gamma_p(n-1)}{\Gamma_p(n-1+2\lambda_n)} \qquad \text{(A.31)}\\
&= -\lambda_n(u-v)\sum_{i=1}^{p}\left\{\log\left(n-1-\frac{i-1}{2}\right)-\frac{1}{2(n-1)+1-i}\right\}\\
&\quad -2(u-v)\sum_{i=1}^{p}\frac{\lambda_n^2}{2(n-1)+1-i}+\max\left\{O\left(\frac{1}{n}\right),O\left(\frac{\lambda_n^3}{n^2}\right)\right\}\\
&= -\lambda_n p(u-v)\log(n-1)-\lambda_n(u-v)\sum_{i=1}^{p}\left[\log\left(1-\frac{i-1}{2n-2}\right)-\frac{1}{2(n-1)+1-i}\right]\\
&\quad -2\lambda_n^2(u-v)\sum_{i=1}^{p}\frac{1}{2(n-1)+1-i}+\max\left\{O\left(\frac{1}{n}\right),O\left(\frac{\lambda_n^3}{n^2}\right)\right\}.
\end{aligned}$$

Combining the relations (A.29)–(A.31), we can obtain that

$$\Psi_n(\lambda)=\frac{\lambda_n}{a_n^2}(A_{1n}+A_{2n}-\mu_{n,v})+\frac{\lambda_n^2}{a_n^2}A_{3n}+o(1), \qquad \text{(A.32)}$$

where

$$\begin{aligned}
A_{1n}&=\sum_{i=1}^{pu}\log\left(1-\frac{i-1}{n-1}\right)-v\sum_{i=1}^{p}\log\left(1-\frac{i-1}{n-1}\right)-(u-v)\sum_{i=1}^{p}\log\left(1-\frac{i-1}{2n-2}\right),\\
A_{2n}&=-\sum_{i=1}^{pu}\frac{1}{n-i}+v\sum_{i=1}^{p}\frac{1}{n-i}+(u-v)\sum_{i=1}^{p}\frac{1}{2(n-1)+1-i},\\
A_{3n}&=\sum_{i=1}^{pu}\frac{1}{n-i}-v\sum_{i=1}^{p}\frac{1}{n-i}-2(u-v)\sum_{i=1}^{p}\frac{1}{2(n-1)+1-i}.
\end{aligned}$$

By the simple fact

$$\int_0^p\log\left(1-\frac{x}{n-1}\right)\mathrm{d}x\le\sum_{i=1}^{p}\log\left(1-\frac{i-1}{n-1}\right)\le\int_0^p\log\left(1-\frac{x-1}{n-1}\right)\mathrm{d}x, \qquad \text{(A.33)}$$

we know

$$\begin{aligned}
A_{1n}\ &\le\ pu\log\left(1+\frac{1}{n-1}\right)-(n-pu)\log\left(1-\frac{pu}{n-1}\right)\\
&\quad +v(n-1-p)\log\left(1-\frac{p}{n-1}\right)+(u-v)[2(n-1)-p]\log\left(1-\frac{p}{2n-2}\right)\\
&:= B_{1n}
\end{aligned}$$

and

$$A_{1n}\ \ge\ -(n-1-pu)\log\left(1-\frac{pu}{n-1}\right)-vp\log\left(1+\frac{1}{n-1}\right)+v(n-p)\log\left(1-\frac{p}{n-1}\right)$$

$$
\begin{aligned}
&-p(u-v)\log\left(1+\frac{1}{2n-2}\right)+(u-v)\left[2(n-1)+1-p\right]\log\left(1-\frac{p}{2n-2}\right)\\
&:=B_{2n}.
\end{aligned}
$$

Thus we have

$$
\lim_{n\to\infty}\left(B_{1n}-\mu_{n,v}\right)=\frac{v}{2}\log\left(1-\frac{y}{u}\right)+\frac{u-v}{2}\log\left(1-\frac{y}{2u}\right)+\frac{3}{2}\log\left(1-y\right)+y
$$

and

$$
\lim_{n\to\infty}(B_{2n}-\mu_{n,v})=\frac{3v}{2}\log\left(1-\frac{y}{u}\right)+\frac{u-v}{2}\log\left(1-\frac{y}{2u}\right)-\frac{1}{2}\log\left(1-y\right)-\frac{u+v}{2u}y,
$$

then we can get $\{A_{1n}-\mu_{n,v}, n\geq 1\}$ is bounded.

By (A.33) again, using the facts $\frac{pu}{n-1}\to y\in(0,1)$ as $n\to\infty$, we can see by some elementary calculations that

$$
\lim_{n\to\infty}\sum_{i=1}^{p}\frac{1}{n-i}=-\log\left(1-\frac{y}{u}\right),\quad \lim_{n\to\infty}\sum_{i=1}^{p}\frac{1}{2(n-1)+1-i}=-\log\left(1-\frac{y}{2u}\right)
$$

and

$$
\lim_{n\to\infty}\sum_{i=1}^{pu}\frac{1}{n-i}=-\log\left(1-y\right).
$$

Then

$$
\lim_{n\to\infty}A_{2n}=-v\log\left(1-\frac{y}{u}\right)-(u-v)\log\left(1-\frac{y}{2u}\right)+\log\left(1-y\right)
$$

and

$$
\lim_{n\to\infty}A_{3n}=v\log\left(1-\frac{y}{u}\right)+2(u-v)\log\left(1-\frac{y}{2u}\right)-\log\left(1-y\right). \tag{A.34}
$$

Using the fact of $\lambda_n/a_n^2=\lambda/(a_n\sigma_{n,v})\to 0$, we have

$$
\frac{\lambda_n}{a_n^2}(A_{1n}+A_{2n}-\mu_{n,v})=o(1) \tag{A.35}
$$

as $n\to\infty$.

Recall the limitation of $\sigma_{n,v}^2$ in (A.15) for $\frac{pu}{n-1}\to y\in(0,1)$. Inserting (A.34) and (A.35) into (A.32), we can obtain that

$$
\Psi_n(\lambda)=\frac{\lambda_n^2}{a_n^2}A_{3n}+o(1)=\frac{\lambda^2}{\sigma_{n,v}^2}A_{3n}+o(1)=\frac{\lambda^2}{2}+o(1)
$$

as $n\to\infty$ for any fixed $\lambda\in\mathbb{R}$ and (A.28) is obtained.

When $\frac{pu}{n-1}\to 0$ as $n\to\infty$. By the definition of $\sigma_{n,v}$, we know $\sigma_{n,v}^2=O\left(\frac{pu}{n-1}\right)$. If $a_n=o(pu)$, then $\lambda_n\to\infty$ and $\frac{\lambda_n}{n}\to 0$. In fact, we can easily deduce that the same results as Lemma A.4 are also held under the case of $\frac{p}{n}\to 0$. Then we also can reach to (A.28) by the similar route.

To summarize, the above discussion shows that the proof of Theorem 4.1 is completed.

Acknowledgements. The authors would like to thank two anonymous referees for many valuable comments and suggestions that have led to improvements in the paper.

References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis. John Wiley & Sons Inc., New York (1984).

[2] S.F. Arnold, Application of the theory of products of problems to certain patterned matrics. *Ann. Stat.* **1** (1973) 682-699.

[3] Z.D. Bai and H. Saranadasa, Effect of high dimension: by an example of a two sample problem. *Stat. Sin.* **6** (1996) 311–329.

[4] A. Basilevsky, Applied Matrix Algebra in the Statistical Sciences. North-Holland, New York (1983).

[5] F.P. Carli, A. Ferrante, M. Pavon and G. Picci, A maximum entropy solution of the covariance extension problem for reciprocal processes. *IEEE Trans. Autom. Control*, **56** (2011) 1999–2012.

[6] F.P. Carli, A. Ferrante, M. Pavon and G. Picci, An efficient algorithm for maximum-entropy extension of block-circulant covariance matrices. *Linear Algebra Appl.* **439** (2011) 2309–2329.

[7] Y.S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingale, 3rd edn.* Springer-Verlag, New York (1997).

[8] C.A. Coelho, The eigenblock and eigenmatrix decomposition of a matrix: its usefulness in statistics-application to the likelihood ratio test for block-circularity. (2013), Preprint.

[9] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications. Springer, New York (2009).

[10] C.A. Gotway and N.A.C. Cressie, A spatial analysis of variance applied to soil-water infiltration. *Water Resour. Res.* **26** (1990) 2695–2703.

[11] A. Gut, Probability: A Graduate Course. Springer-Verlag, New York (2005).

[12] A.M. Hartley and D.N. Naik, Estimation of familial correlations under autoregressive circular covariance. *Commun. Stat. Theory Methods* **30** (2001) 1811–1828.

[13] H. Jiang and S.C. Wang, Moderate deviation principles for classical likelihood ratio tests of high-dimensional normal distributions. *J. Multivar. Anal.* **156** (2017) 57–69.

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

[15] T.F. Jiang and Y.C. Qi, Likelihood ratio tests for high-dimensional normal distributions. *Scand. J. Stat.* **42** (2015) 988–1009.

[16] N. Kato, T. Yamada and Y. Fujikoshi, High-dimensional asymptotic expansion of LR statistic for testing intraclass correlation structure and its error bound. *J. Multivar. Anal.* **101** (2010) 101–112.

[17] R. Leiva, Linear discrimination with equicorrelated training vectors. *J. Multivar. Anal.* **98** (2007) 384–409.

[18] Y.L. Liang, T. Von Rosen and D. Von Rosen, Block circular symmetry in multilevel models, in Research Report 2011, vol. 3, Department of Statistics, Stockholm University (2011).

[19] Y.L. Liang, T. Von Rosen and D. Von Rosen, On estimation in multilevel models with block circular symmetric covariance structures. *Acta et Commentationes Universitatis Tartuensis de Mathematica* **16** (2012) 1–14.

[20] Y.L. Liang, T. Von Rosen and D. Von Rosen, On estimation in hierarchical models with block circular covariance structures. *Ann. Inst. Stat. Math.* **67** (2015) 1–19.

[21] A. Lindquist and G. Picci, Modeling of stationary periodic time series by ARMA representations, in Optimization and Its Applications in Control and Data Sciences, edited by B. Goldengorin. *Springer Optimization and Its Applications*, Vol 115. Springer, Cham (2016).

[22] M. Makoto, K. Kazuyuki and S. Takashi, Likelihood ratio test statistic for block compound symmetry covariance structure and its asymptoic expansion. Technical Report No.15-03, Statistical Research Group, Hiroshima University, Japan (2015).

[23] F.J. Marques and C.A. Coelho, Obtaining the exact and near-exact distributions of the likelihood ratio statistic to test circular symmetry through the use of characteristic functions. *Comput. Stat.* **28** (2013) 2091–2115.

[24] R.J. Muirhead, Aspects of Multivariate Statistical Theory. John Wiley & Sons Inc., New York (1982).

[25] D.K. Nagar, J. Chen and A.K. Gupta, Distribution and percentage points of the likelihood ratio statistic for testing circular symmetry. *Comput. Stat. Data Anal.* **47** (2004) 79–89.

[26] D.K. Nagar, S.K. Jain and A.K. Gupta, On testing circular stationarity and related models. *J. Stat. Comput. Simul.* **29** (1988) 225–239.

[27] T. Nahtman and D.V. Rosen, Shift permutation invariance in linear random factor models. *Math. Methods Stat.* **17** (2008) 173–185.

[28] I. Olkin, Testing and estimation for structures which are circularly symmetric in blocks, in Multivariate Statistical Inference, edited by D.G. Kabe, R.P. Gupta. North-Holland, Amsterdam (1973) 183–195.

[29] I. Olkin and S.J. Press, Testing and estimation for a circular stationary model. *Ann. Math. Stat.* **40** (1969) 1358–1373.

[30] C.R. Rao, Familial correlations or the multivariate generalizations of the intraclass correlation. *Curr. Sci.* **14** (1945) 66–67.

[31] C.R. Rao, Discriminant functions for genetic differentiation and selection. *Sankhya* **12** (1953) 229–246.

[32] A. Roy and R. Leiva, Estimating and testing a structured covariance matrix for three-level multivariate data. *Commun. Stat. Theory Methods* **40** (2011) 1945–1963.

[33] M.S. Srivastava, Estimation of intraclass correlations in familial data. *Biometrika* **71** (1984) 177–185.
[34] S.S. Wilks, Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Ann. Math. Stat.* **17** (1946) 257–281.
[35] L.Q. Yi and J.S. Xie, A high-dimensional likelihood ratio test for circular symmetric covariance structure. *Commun. Stat. Theory Methods* **47** (2018) 1392–1402.