Electronic Journal of Statistics
Vol. 10 (2016) 2973–3010
ISSN: 1935-7524
DOI: 10.1214/16-EJS1199

Testing the sphericity of a covariance matrix when the dimension is much larger than the sample size

Zeng Li and Jianfeng Yao*

Department of Statistics and Actuarial Science
The University of Hong Kong
e-mail: u3001205@hku.hk; jeffyao@hku.hk

Abstract: This paper focuses on the prominent sphericity test when the dimension p is much lager than sample size n. The classical likelihood ratio test(LRT) is no longer applicable when $p \gg n$. Therefore a Quasi-LRT is proposed and its asymptotic distribution of the test statistic under the null when $p/n \to \infty, n \to \infty$ is well established in this paper. We also re-examine the well-known John's invariant test for sphericity in this ultra-dimensional setting. An amazing result from the paper states that John's test statistic has exactly the same limiting distribution under the ultra-dimensional setting with under other high-dimensional settings known in the literature. Therefore, John's test has been found to possess the powerful *dimension-proof* property, which keeps exactly the same limiting distribution under the null with any (n, p)-asymptotic, i.e. $p/n \to [0, \infty]$, $n \to \infty$. Nevertheless, the asymptotic distribution of both test statistics under the alternative hypothesis with a general population covariance matrix is also derived and incorporates the null distributions as special cases. The power functions are presented and proven to converge to 1 as $p/n \to \infty,\ n \to \infty,\ n^3/p = O(1)$. All asymptotic results are derived for general population with finite fourth order moment. Numerical experiments are implemented to illustrate the finite sample performance of the results.

MSC 2010 subject classifications: Primary 62H15, 62H10; secondary 62F03.
Keywords and phrases: Sphericity test, large dimension, ultra-dimension, John's test, Quasi-likelihood ratio test.

Received August 2015.

Contents

*Jianfeng Yao's research is partially supported by the RGC GRF grant 17305814

1. Introduction

High dimensional data with dimension p of same scale with or even larger than the number of observations n has applausive statistical applications in biology and finance recently. In particular, practical needs for testing gene-wise independence in genomic studies have inspired a wide range of discussions regarding test of structures of the covariance matrix.

In this paper, we consider the prominent sphericity test when the dimension p is much larger than the sample size n. Let $X = (X_1, X_2, \cdots, X_n)$ be a $p \times n$ data matrix with n independent and identically distributed p-dimensional random vectors $\{X_i\}_{1 \le i \le n}$ with covariance $\Sigma = Var(X_i)$. Our interest is to test

$$H_0 : \Sigma = \sigma^2 I_p \quad \text{vs.} \quad H_1 : \Sigma \neq \sigma^2 I_p, \tag{1.1}$$

where σ^2 is an unknown positive constant. Among traditional tests are the likelihood ratio test(LRT) and John's invariant test.

Consider first the LRT with test statistic in [1]

$$-2\log L_n = -2\log\left(\frac{(l_1\cdots l_p)^{1/p}}{\frac{1}{p}(l_1+\cdots+l_p)}\right)^{\frac{pn}{2}} = n\log\left(\frac{\bar{l}^p}{\prod_{i=1}^p l_i}\right), \tag{1.2}$$

where $\{l_i\}_{1\le i\le p}$ are the eigenvalues of p-dimensional sample covariance matrix $S_n = \frac{1}{n}\sum_{i=1}^n X_i X_i' = \frac{1}{n}XX'$, $X = (X_1, \cdots, X_n)$. If we let $n \to \infty$ while keeping p fixed, classics asymptotic theory indicates that under the null hypothesis and assuming the population is normal,

$$-2\log L_n \xrightarrow{d} \chi^2_{\frac{1}{2}p(p+1)-1},$$

the chi-square distribution is further refined by the Box-Bartlett correction. However, this χ^2-convergence becomes slow when the dimension p increases so that the LRT (and its Box-Bartlett correction) is seriously biased when the dimension-to-sample size ratio p/n is not small enough.

[16] made bias correction to the traditional LRT test under the regime where both $p, n \to \infty$, $p/n \to c \in (0, 1)$. They derived that when $X = \{x_{ij}\}_{\substack{1\le i\le p\\ 1\le j\le n}}$ with

i.i.d entries satisfying $\mathbb{E}(x_{ij}) = 0$, $\mathbb{E}|x_{ij}|^2 = 1$, $\nu_4 := \mathbb{E}|x_{ij}|^4 < \infty$, and under H_0,

$$-\tfrac{2}{n}\log L_n + (p-n)\log(1-\tfrac{p}{n}) - p \xrightarrow{d} N\left(-\tfrac{1}{2}\log(1-c) + \tfrac{\nu_4-3}{2}c, -2\log(1-c)-2c\right). \tag{1.3}$$

Notice that here the scale parameter σ^2 in H_0 has been taken to be $\sigma^2 = 1$ as the LRT statistic is invariant under scaling. Extensive simulation study in [16] shows that this test is well adapted to high dimensions and has a very reasonable size and power for a wide range of dimension-sample size combinations (p, n). The LRT however requires that $p \le n$ because when $p > n$, $n - p$ of the sample eigenvalues $\{l_i\}$ are null so that the likelihood ratio L_n is identically null. In this paper, we introduce a quasi-LRT statistic which can be seen as a natural extension of the LRT statistic to the situation where $p > n$. The quasi-LRT test statistic is defined as

$$\mathcal{L}_n = \frac{p}{n}\log\frac{\left(\frac{1}{n}\sum_{i=1}^n \eta_i\right)^n}{\prod_{i=1}^n \eta_i}, \tag{1.4}$$

where $\{\eta_i\}_{1\le i\le n}$ are eigenvalues of n-dimensional matrix $\frac{1}{p}X'X$. The main idea is that the companion matrix $X'X$ has exactly the same n non-null eigenvalues with the sample covariance matrix XX'(up to some scaling). Therefore, the quasi-LRT test statistic removes all the null eigenvalues in the original LRT test statistic and we find that under the so-called ultra-dimensional asymptotic $p \gg n$, that is $p/n \to \infty$ and $n \to \infty$,

$$\mathcal{L}_n - \frac{n}{2} - \frac{n^2}{6p} - \frac{\nu_4 - 2}{2} \xrightarrow{d} N(0,1).$$

Based on this asymptotic result, a quasi-LRT test can be conducted to test sphericity to compensate for the inapplicability of the traditional LRT in the ultra-dimension setting.

Next we consider John's invariant test for sphericity. [7, 8] studied the problem for normal populations and proposed the testing statistic

$$U = \frac{1}{p}tr\left[\left(\frac{S_n}{(1/p)tr(S_n)} - I_p\right)^2\right] = \frac{p^{-1}\sum_{i=1}^p (l_i - \bar{l})^2}{\bar{l}^2}, \tag{1.5}$$

where $\bar{l} = \frac{1}{p}\sum_{i=1}^p l_i$. It has been proved that, as $n \to \infty$ while p remain fixed, the limiting distribution of U under H_0 is

$$nU - p \xrightarrow{d} \frac{2}{p}\chi^2_{p(p+1)/2-1} - p.$$

Contrary to the LRT, it has been noticed for a while that John's test does not suffer from high dimensions and this χ^2 limit is quite accurate even when the ratio p/n is not small. [9] studied the (n, p)-consistency of this test statistic under normality assumptions. They proved that, when $n, p \to \infty$, $\lim_{n\to\infty} p/n \to c \in (0, +\infty)$,

$$nU - p \xrightarrow{d} N(1,4). \tag{1.6}$$

Meanwhile, when $p \to \infty$,

$$\frac{2}{p}\chi^2_{p(p+1)/2-1} - p \xrightarrow{d} N(1,4).$$

In other words, [9] extended the classical n-asymptotic theory (where p is fixed) to the high-dimensional case where p goes to infinity proportionally with n. Meanwhile, the robustness of John's test is explained in this proportional high-dimensional scheme.

[16] further relaxed the normality restriction and proved that, if $\{x_{ij}\}$ are i.i.d. with $\mathbb{E}x_{ij} = 0$, $\mathbb{E}|x_{ij}|^2 = 1$, $\nu_4 \triangleq \mathbb{E}|x_{ij}|^4 < \infty$, then when $n, p \to \infty$, $\lim_{n\to\infty} p/n \to c \in (0, +\infty)$,

$$nU - p \xrightarrow{d} N(\nu_4 - 2, 4). \tag{1.7}$$

Since $\nu_4 = 3$ for normal distribution, it shows that the existing results confirm with each other. In this paper, we extend the above result one step further, i.e. consider the asymptotic behavior of the John's test statistic under the ultra-dimensional $p \gg n$ setting. We find that this test statistic possesses a remarkable *dimension-proof* property, which shows that under the (n,p)-asymptotic, the limit in (1.7) still holds when $\lim_{n\to\infty} p/n = \infty$. This *dimension-proof* property of John's test makes it a very competitive candidate for sphericity testing regardless of p, n.

Related methods have also been proposed in the literature for the high dimensional sphericity test. Noteworthy work include [12] where a test statistic based on the logarithm of the norm of sample correlation matrix under (n,p)-asymptotic has been well studied. Yet multivariate normality assumption has been assumed in this paper. Similarly in [6], a novel test statistic utilizing the ratio of the fourth and second arithmetic means of the sample covariance matrix is developed under the $p/n \to c$, (n,p)-asymptotic with normality restriction. [13] considered the ratio of arithmetic means of the eigenvalues of sample covariance matrix in the normal case when $n = O(p^\delta), \delta > 0$, $n, p \to \infty$ and [15] further proved the robustness of this test statistic against non-normality assumption irrespective of either $n/p \to 0$ or $n/p \to \infty$. However, their results are only applicable under some specified factorized settings, which makes it less general than John's test. [5] developed a high-dimensional test based on the John's test, however this test is very time-consuming (See Section 2.4). [19] considered the multivariate-sign-based covariance matrices to construct robust test for sphericity and significantly enhanced test performance when the non-normality is severe, particularly for heavy tailed distributions. In their paper the asymptotic distributions of the test statistic when $p = O(n^2)$ is derived. [14] studied a quasi-likelihood ratio test under the $n = O(p^\delta)$, $0 < \delta < 1$, $n, p \to \infty$ asymptotic in the normal case, while in this paper, the normality assumption is released and results are discussed under a wider range of (n,p)-asymptotic. These tests are compared in the simulation studies of the paper in Section 2.4.

The rest of the paper is organized as follows. Section 2 discusses the asymptotic behavior of the John's test statistic and the quasi-LRT test statistic under the ultra-dimensional setting. The main theorem and its proof is presented in this section. Section 2.4 compared the empirical sizes and powers of John's test, quasi-LRT test and other methods under various scenarios. Section 3 presented theoretical results for power of John's test and quasi-LRT test and testified these results with simulations. Section 4 generalize the asymptotic results to the case when population mean is unknown and the unbiased sample covariance matrix is adopted. Section 5 concludes. Section A displays some technique lemmas and related proofs.

2. New tests and their asymptotic distributions

2.1. Preliminary knowledge

For any $n\times n$ Hermitian matrix M with real eigenvalues $\lambda_1, \cdots, \lambda_n$, the empirical spectral distribution (ESD for short) of M is defined by $F^M = n^{-1}\sum_{j=1}^n \delta_{\lambda_j}$, where δ_a denotes the Dirac mass at a. The Stieltjes transform of any distribution G is defined as

$$m_G(z) = \int \frac{1}{x-z} dG(x),\ \ \Im(z) > 0,$$

where $\Im(z)$ stands for the imaginary part of z.

Consider the re-normalized sample covariance matrix $A = \sqrt{\frac{p}{n}}\left(\frac{1}{p}X'X - I_n\right)$, where $X = (x_{ij})_{p\times n}$ and $x_{ij}, i = 1, \cdots, p,\ j = 1, \cdots, n$ are i.i.d. real random variables with mean zero and variance one, I_n is the identity matrix of order n. It's known that under the ultra-dimensional setting [3], with probability one, the ESD of matrix A, F^A converges to the semicircle law F with density

$$F'(x) = \begin{cases} \dfrac{1}{2\pi}\sqrt{4-x^2}, & \text{if } |x| \le 2, \\ 0, & \text{if } |x| > 2. \end{cases}$$

We denote the Stieltjes transform of the semicircle law F by $m(z)$. Let $\mathscr{S}$ denote any open region on the complex plane including $[-2, 2]$, the support of F and $\mathscr{M}$ be the set of functions which are analytic on $\mathscr{S}$. For any $f \in \mathscr{M}$, denote

$$G_n(f) \triangleq n\int_{-\infty}^{+\infty} f(x)d\left(F^A(x) - F(x)\right) - \frac{n}{2\pi i}\oint_{|m|=\rho} f\left(-m-m^{-1}\right)\chi_n(m)\frac{1-m^2}{m^2}\mathrm{d}m, \tag{2.1}$$

where

$$\chi_n(m) \triangleq \frac{-\mathcal{B} + \sqrt{\mathcal{B}^2 - 4\mathcal{A}\mathcal{C}}}{2\mathcal{A}},\ \mathcal{A} = m - \sqrt{\frac{n}{p}}(1+m^2),$$

$$\mathcal{B} = m^2 - 1 - \frac{n}{p}m(1+2m^2),\ \mathcal{C} = \frac{m^3}{n}\left(\frac{m^2}{1-m^2} + \nu_4 - 2\right) - \sqrt{\frac{n}{p}}m^4,$$

$\nu_4 = \mathbb{E}X_{11}^4$ and $\sqrt{\mathcal{B}^2 - 4\mathcal{AC}}$ is a complex number whose imaginary part has same sign as that of $\mathcal{B}$. The integral's contour is taken as $|m| = \rho$ with $\rho < 1$. [4] gives a calibration in advance for the mean correction term in (2.1), where only $\mathcal{C}$ is replaced with

$$\mathcal{C}^{\text{Calib}} = \frac{m^3}{n}\left[\nu_4 - 2 + \frac{m^2}{1-m^2} - 2(\nu_4 - 1)m\sqrt{\frac{n}{p}}\right] - \sqrt{\frac{n}{p}}m^4$$

while others remain the same.

The central limit theorem (CLT) of linear functions of eigenvalues of the renormalized sample covariance matrix A when the dimension p is much larger than the sample size n derived by [4] is stated as follows.

Theorem 2.1. *Suppose that*

(a) $\mathbf{X} = (x_{ij})_{p\times n}$ *where* $\{x_{ij} : i = 1, \cdots, p;\ j = 1, \cdots, n\}$ *are i.i.d. real random variables with* $\mathbb{E}X_{11} = 0$, $\mathbb{E}X_{11}^2 = 1$ *and* $\nu_4 = \mathbb{E}X_{11}^4 < \infty$.
(b) $n/p \to 0$ *as* $n \to \infty$.

Then, for any $f_1, \cdots, f_k \in \mathscr{M}$, *the finite dimensional random vector* $(G_n(f_1), \cdots, G_n(f_k))$ *converges weakly to a Gaussian vector* $(Y(f_1), \cdots, Y(f_k))$ *with mean function* $\mathbb{E}Y(f) = 0$ *and covariance function*

$$\begin{aligned} cov\,(Y(f_1), Y(f_2)) &= (\nu_4 - 3)\Phi_1(f_1)\Phi_1(f_2) + 2\sum_{k=1}^{\infty} k\Phi_k(f_1)\Phi_k(f_2) \qquad (2.2)\\ &= \frac{1}{4\pi^2}\int_{-2}^{2}\int_{-2}^{2} f_1'(x)f_2'(y)H(x,y)\,\mathrm{d}x\,\mathrm{d}y \end{aligned}$$

where

$$\Phi_k(f) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos\theta)e^{ik\theta}\,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos\theta)\cos k\theta\,\mathrm{d}\theta,$$

$$H(x,y) = (\nu_4 - 3)\sqrt{4-x^2}\sqrt{4-y^2} + 2\log\left(\frac{4 - xy + \sqrt{(4-x^2)(4-y^2)}}{4 - xy - \sqrt{(4-x^2)(4-y^2)}}\right).$$

The proofs of the main theorems in this paper are based on two lemmas derived from this CLT. Notice that the limiting covariance functions in (2.2) has been first established in [2] for Wigner matrices.

Lemma 2.1. *Let* $\{\lambda_i,\ 1 \le i \le n\}$ *be eigenvalues of the matrix* $A = \sqrt{\frac{p}{n}}\left(\frac{1}{p}X'X - I_n\right)$, *where* X *satisfies the assumptions in Theorem 2.2, then as* $p/n \to \infty$, $n \to \infty$,

$$\begin{pmatrix} \sum_{i=1}^n \lambda_i^2 - n - (\nu_4 - 2) \\ \sum_{i=1}^n \lambda_i \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & \nu_4 - 1 \end{pmatrix}\right).$$

Lemma 2.2. *Let $\{\lambda_i,\ 1\le i\le n\}$ be eigenvalues of matrix $A=\sqrt{\frac{p}{n}}\left(\frac{1}{p}X'X-I_n\right)$, where X satisfies the assumptions in Theorem 2.3, then as $p/n\to\infty$, $n\to\infty$,*

$$\left(\begin{array}{c}\sum_{i=1}^n\lambda_i\\ \sqrt{\frac{p}{n}}\sum_{i=1}^n\log\left(1+\lambda_i\sqrt{\frac{n}{p}}\right)+\frac{1}{2}\sqrt{\frac{n^3}{p}}+\frac{n^2}{6p}\sqrt{\frac{n}{p}}+\frac{\nu_4-2}{2}\sqrt{\frac{n}{p}}\end{array}\right)=\xi_n+o_p(1),$$

where

$$\xi_n\sim N\left(\left(\begin{array}{c}0\\0\end{array}\right),\ \left(\begin{array}{cc}\nu_4-1 & (\nu_4-1)\left(1+\frac{n}{p}\right)\\ (\nu_4-1)\left(1+\frac{n}{p}\right) & \nu_4-1+\frac{n}{p}(2\nu_4-1)\end{array}\right)\right).$$

The proofs of these two lemma are postponed to Appendix A.

2.2. John's test

Consider John's test statistic U defined in (1.5) based on eigenvalues of the p-dimensional sample covariance matrix $S_n=\frac{1}{n}XX'$. Here we assume that the $X_j's$ in X have representation $X_j=\Sigma^{1/2}Z_j$, where $\{Z_1,\cdots,Z_n\}=\{z_{ij}\}_{1\le i\le p,1\le j\le n}$ is a $p\times n$ matrix with i.i.d. entries z_{ij} satisfying $\mathbb{E}(z_{ij})=0$, $\mathbb{E}(z_{ij}^2)=1$. It can be seen that, under the null hypothesis H_0, the John's test statistic is independent from the scale parameter σ^2. Therefore, we assume w.l.o.g. $\sigma^2=1$ when we derive the null distribution of the test statistic. In other words, under H_0, we assume in the rest of this paper that sample vectors $\{x_{ij}\}_{1\le i\le p,1\le j\le n}$ satisfy $\mathbb{E}(x_{ij})=0$, $\mathbb{E}(x_{ij}^2)=1$, $\mathbb{E}(|x_{ij}|^4)=\nu_4<+\infty$. The first main result of this paper is the following.

Theorem 2.2. *Assume $X=\{x_{ij}\}_{p\times n}$, $x_{ij}=z_{ij}$ which are i.i.d. satisfying $\mathbb{E}(z_{ij})=0$, $\mathbb{E}(z_{ij}^2)=1$, $\mathbb{E}|z_{ij}|^4=\nu_4<\infty$, then when $p/n\to\infty$, $n\to\infty$,*

$$nU-p\xrightarrow{d}N(\nu_4-2,4).\tag{2.3}$$

Similarly with this theorem, [16] shows that if $\{x_{ij}\}$ are i.i.d. with $\mathbb{E}x_{ij}=0$, $\mathbb{E}|x_{ij}|^2=1$, $\nu_4\triangleq\mathbb{E}|x_{ij}|^4<\infty$, then when $n,p\to\infty$, $\lim_{n\to\infty}p/n\to c\in(0,+\infty)$,

$$nU-p\xrightarrow{d}N(\nu_4-2,4).$$

It indicates that as long as $X=\{x_{ij}\}_{p\times n}$ are i.i.d with zero mean, unit variance and finite fourth order moment, John's test statistic $nU-p$ has a consistent limiting distribution $N(\nu_4-2,4)$, regardless of normality, under any (n,p)-asymptotic, $n/p\to[0,\infty)$. Therefore, the powerful *dimension-proof* property assigns John's test top priority when little information about the data is known before implementing sphericity test.

The proof of Theorem 2.2 is based on Lemma 2.1.

Proof. Denote the eigenvalues of $p\times p$ matrix $S_n=\frac{1}{n}XX'$ in descending order by $l_i(1\le i\le p)$, and the eigenvalues of $n\times n$ matrix $A=\sqrt{\dfrac{p}{n}}\left(\dfrac{1}{p}X'X-I_n\right)$

by $\lambda_i (1 \le i \le n)$. Since $p > n$, S_n has $p - n$ zero eigenvalues and the remaining n non-zero eigenvalues $l_i (1 \le i \le n)$ are related with $\lambda_i (1 \le i \le n)$ eigenvalues of A as

$$\sqrt{\frac{p}{n}}\lambda_i + \frac{p}{n} = l_i, \quad 1 \le i \le n.$$

We have, for John's test statistic

$$\begin{aligned} U &= \left(\frac{1}{p} \sum_{i=1}^{n} \frac{p^2}{n^2} \left(\sqrt{\frac{n}{p}} \lambda_i + 1 \right)^2 \right) \Bigg/ \left(\frac{1}{p} \sum_{i=1}^{n} \frac{p}{n} \left(\sqrt{\frac{n}{p}} \lambda_i + 1 \right) \right)^2 - 1 \\ &= \frac{\sum_{i=1}^{n} \lambda_i^2 + 2\sqrt{\frac{p}{n}} \sum_{i=1}^{n} \lambda_i + p}{\left(\sqrt{\frac{1}{p}} \sum_{i=1}^{n} \lambda_i + \sqrt{n} \right)^2} - 1, \end{aligned}$$

Define the function $G(u, v) = \dfrac{u + 2v\sqrt{\frac{p}{n}} + p}{(\sqrt{\frac{1}{p}}v + \sqrt{n})^2} - 1$, then John's test statistic can be written as

$$U = G\left(u = \sum_{i=1}^{n} \lambda_i^2, v = \sum_{i=1}^{n} \lambda_i \right).$$

According to Lemma 2.1, when $p/n \to \infty$, $n \to \infty$,

$$\begin{pmatrix} \sum_{i=1}^{n} \lambda_i^2 - n - (\nu_4 - 2) \\ \sum_{i=1}^{n} \lambda_i \end{pmatrix} \xrightarrow{d} N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & \nu_4 - 1 \end{pmatrix} \right).$$

Then by the Delta Method,

$$n \left(U - G(u, v)|_{u = n + \nu_4 - 2, v = 0} \right) = \xi_n + o_p(1)$$

where

$$\xi_n \sim N \left(0,\ n^2 \nabla G \begin{pmatrix} 4 & 0 \\ 0 & \nu_4 - 1 \end{pmatrix} \nabla G' \right),$$

and $\nabla G = \left(\dfrac{\partial U}{\partial u}, \dfrac{\partial U}{\partial \nu} \right) \Bigg|_{u = n + \nu_4 - 2, v = 0}$ is the corresponding gradient vector.

We have, for $(u, v) = (n + \nu_4 - 2, 0)$,

$$G = \frac{p}{n} + \frac{\nu_4 - 2}{n},$$

and

$$\nabla G \begin{pmatrix} 4 & 0 \\ 0 & \nu_4 - 1 \end{pmatrix} \nabla G' = \frac{4}{n^2} + \frac{4(\nu_4 - 1)}{np} \left(1 + \frac{\nu_4 - 2}{n} \right)^2.$$

The conclusion thus follows. $\square$

2.3. Quasi-likelihood ratio test

Consider the Quasi-LRT statistic $\mathcal{L}_n$ in (1.4) based on the eigenvalues of n-dimensional matrix $\frac{1}{p}X'X$, which are also proportional to the non-null eigenvalues of p-dimensional sample covariance matrix $\frac{1}{n}XX'$. Similarly with John's test statistic, it can be seen that, under the null hypothesis H_0, the $\mathcal{L}_n$ statistic is independent of the scale parameter σ^2. Therefore, we again assume w.l.o.g. $\sigma^2 = 1$ when we derive the null distribution of the test statistic. The second main result of this paper is the following theorem.

Theorem 2.3. *Assume $X = \{x_{ij}\}_{p\times n}$, $x_{ij} = z_{ij}$ which are i.i.d. satisfying $\mathbb{E}(z_{ij}) = 0$, $\mathbb{E}(z_{ij}^2) = 1$, $\mathbb{E}|z_{ij}|^4 = \nu_4 < \infty$, then when $p/n \to \infty$, $n \to \infty$,*

$$\mathcal{L}_n - \frac{n}{2} - \frac{n^2}{6p} - \frac{\nu_4 - 2}{2} \xrightarrow{d} N(0,1). \tag{2.4}$$

Recall the classic LRT when H_0 holds and p is fixed while $n \to \infty$, if the population is Gaussian, the test statistic

$$-2\log L_n = n \log\left(\frac{\bar{l}^p}{\prod_{i=1}^p l_i}\right) \xrightarrow{d} \chi^2_{\frac{1}{2}p(p+1)-1},$$

where $\{l_i\}_{1\le i\le p}$ are the eigenvalues of p-dimensional sample covariance matrix $\frac{1}{n}XX'$. Here we notice that $n/p \to \infty$.

By interchanging the role of n and p, which is feasible under H_0, it can be seen that when n fixed and $p/n \to \infty$, the test statistic

$$-2\log L_p = p \log\left(\frac{\bar{l}^n}{\prod_{i=1}^n l_i}\right) \xrightarrow{d} \chi^2_{\frac{1}{2}n(n+1)-1},$$

$\{l_i\}_{1\le i\le n}$ are the eigenvalues of n-dimensional sample covariance matrix $\frac{1}{p}X'X$. Note that $(-2\log L_p)/n$ coincides with our Quasi-LRT statistic $\mathcal{L}_n$. Heuristically, if next we let $n \to \infty$, then

$$\frac{\chi^2_{\frac{1}{2}n(n+1)-1}}{n} - \frac{n+1}{2} \xrightarrow{d} N(0,1),$$

which is nothing but (2.4) applied to the normal case ($\nu_4 = 3$) with fixed n and $p \to \infty$. Therefore, the classical LRT can be thought of as a particular "finite-dimensional" instance of the general limit of (2.4) for the Quasi-LRT, that is, Theorem 2.3 covers a wide range of "large p, small n" situations.

The proof of Theorem 2.3 is based on lemma 2.2.

Proof. Denote the eigenvalues of $n \times n$ matrix $\frac{1}{p}X'X$ in descending order by $\eta_i (1 \le i \le n)$, and eigenvalues of $n \times n$ matrix $A = \sqrt{\frac{p}{n}}\left(\frac{1}{p}X'X - I_n\right)$ by $\lambda_i (1 \le i \le n)$. These eigenvalues are related as

$$\sqrt{\frac{n}{p}}\lambda_i + 1 = \eta_i, \quad 1 \le i \le n.$$

We have, for the Quasi-LRT test statistic

$$\begin{aligned}&\mathcal{L}_n-\frac{n}{2}-\frac{n^2}{6p}-\frac{\nu_4-2}{2}\\&=\frac{p}{n}\log\left[\left(\frac{1}{n}\sum_{i=1}^{n}\eta_i\right)^n\Big/\prod_{i=1}^{n}\eta_i\right]-\frac{n}{2}-\frac{n^2}{6p}-\frac{\nu_4-2}{2}\\&=p\log\left(1+\sqrt{\frac{n}{p}}\left(\frac{1}{n}\sum_{i=1}^{n}\lambda_i\right)\right)-\frac{p}{n}\sum_{i=1}^{n}\log\left(1+\sqrt{\frac{n}{p}}\lambda_i\right)-\frac{n}{2}-\frac{n^2}{6p}-\frac{\nu_4-2}{2}\end{aligned}$$

Define the function

$$G(u,v)=p\log\left(1+\sqrt{\frac{n}{p}}\left(\frac{1}{n}u\right)\right)-\sqrt{\frac{p}{n}}v,$$

then the Quasi-LRT test statistic can be written as

$$\mathcal{L}_n-\frac{n}{2}-\frac{n^2}{6p}-\frac{\nu_4-2}{2}=$$

$$G\left(u=\sum_{i=1}^{n}\lambda_i,v=\sqrt{\frac{p}{n}}\sum_{i=1}^{n}\log\left(1+\lambda_i\sqrt{\frac{n}{p}}\right)+\frac{1}{2}\sqrt{\frac{n^3}{p}}+\frac{n^2}{6p}\sqrt{\frac{n}{p}}+\frac{\nu_4-2}{2}\sqrt{\frac{n}{p}}\right).$$

According to Lemma 2.2, when $p/n\to\infty$, $n\to\infty$,

$$\begin{pmatrix}\sum_{i=1}^{n}\lambda_i\\ \sqrt{\frac{p}{n}}\sum_{i=1}^{n}\log\left(1+\lambda_i\sqrt{\frac{n}{p}}\right)+\frac{1}{2}\sqrt{\frac{n^3}{p}}+\frac{n^2}{6p}\sqrt{\frac{n}{p}}+\frac{\nu_4-2}{2}\sqrt{\frac{n}{p}}\end{pmatrix}=\xi_n+o_p(1),$$

where

$$\xi_n\sim N\left(\begin{pmatrix}0\\0\end{pmatrix},\ \begin{pmatrix}\nu_4-1 & (\nu_4-1)\left(1+\frac{n}{p}\right)\\ (\nu_4-1)\left(1+\frac{n}{p}\right) & \nu_4-1+\frac{n}{p}(2\nu_4-1)\end{pmatrix}\right).$$

Then by the Delta Method,

$$\mathcal{L}_n-\frac{n}{2}-\frac{n^2}{6p}-\frac{\nu_4-2}{2}-G(u,v)|_{u=0,v=0}\xrightarrow{d}$$

$$N\left(0,\ \nabla G\begin{pmatrix}\nu_4-1 & (\nu_4-1)\left(1+\frac{n}{p}\right)\\ (\nu_4-1)\left(1+\frac{n}{p}\right) & \nu_4-1+\frac{n}{p}(2\nu_4-1)\end{pmatrix}\nabla G'\right),$$

where $\nabla G=\left(\frac{\partial U}{\partial u},\frac{\partial U}{\partial \nu}\right)\Big|_{u=0,v=0}$ is the corresponding gradient vector.

We have, for $(u,v)=(0,0)$, $G=0$ and

$$\nabla G\begin{pmatrix}\nu_4-1 & (\nu_4-1)\left(1+\frac{n}{p}\right)\\ (\nu_4-1)\left(1+\frac{n}{p}\right) & \nu_4-1+\frac{n}{p}(2\nu_4-1)\end{pmatrix}\nabla G'=1.$$

Therefore, when $p/n \to \infty$, $n \to \infty$,

$$\mathcal{L}_n - \frac{n}{2} - \frac{n^2}{6p} - \frac{\nu_4 - 2}{2} \xrightarrow{d} N(0,1). \qquad \square$$

Remark 2.1. *Both Theorem 2.2 and 2.3 involve the parameter $\nu_4 = \mathbb{E}(|x_{ij}|^4)$. In practice, this parameter is unknown and needs to be estimated. Notice that in our setting and under the null, the whole array of data $\{x_{ij}\}$ are i.i.d. random variables. As a consequence, we use the fouth-order sample moment as an estimator, i.e.*

$$\hat{\nu}_4 = \frac{1}{np} \sum_{i=1}^{p} \sum_{j=1}^{n} x_{ij}^4.$$

By the law of large numbers, we have $\hat{\nu}_4 = \nu_4 + o(1)$ almost surely, therefore substituting $\hat{\nu}_4$ for ν_4 in both Theorem 2.2 and 2.3 will not change the limit distribution. The performance of both test statistics with this estimated fourth moment $\hat{\nu}_4$ is checked in the following simulation studies.

2.4. Simulation studies

In order to further explore the finite sample behavior of John's sphericity test when dimension p is significantly larger than the sample size n, Monte Carlo simulations are implemented in this session to evaluate the size and power of John's Sphericity Test. Test statistic proposed by [5] is also considered for comparison.

In the simulation, without loss of generality, we conduct the sphericity test with $\sigma^2 = 1$. To find the empirical sizes of these two tests, we consider two different scenarios to generate sample data:

(1) $\{X_j\},\ 1 \le j \le n$ i.i.d p-dimensional random vector generated from multivariate normal population $N(0, I_p)$, $\mathbb{E}x_{ij}^4 = \nu_4 = 3$;
(2) $\{x_{ij},\ 1 \le i \le p,\ 1 \le j \le n\}$ i.i.d follow $Gamma(4,2) - 2$ distribution, then $\mathbb{E}x_{ij} = 0$, $\mathbb{E}x_{ij}^2 = 1$, $\mathbb{E}x_{ij}^4 = \nu_4 = 4.5$.

We set sample size $n = 64$, dimension $p = 320, 640, 960, 1280, 1600, 2400, 3200$ in order to understand the effect of an increasing dimension. The nominal test level is $\alpha = 0.05$. For each pair of (p, n), 10000 replications are used to get the empirical size.

For John's test, we reject H_0 if $nU - p$ exceeds the 5% upper quantile of $N(\nu_4 - 2, 4)$ distribution. For Quasi-LRT test, we reject H_0 if $\mathcal{L}_n - \frac{n}{2} - \frac{n^2}{6p} - \frac{\nu_4 - 2}{2}$ exceeds the 5% upper quantile of $N(0,1)$ distribution.

As for the test in [5], the test statistic is defined as follows:

$$U_n = p\left(\frac{T_{2,n}}{T_{1,n}^2}\right) - 1,$$

where

$$T_{1,n} = \frac{1}{n}\sum_{i=1}^{n} X_i' X_i - \frac{1}{P_n^2}\sum_{i\neq j} X_i' X_j,$$

$$T_{2,n} = \frac{1}{P_n^2}\sum_{i\neq j}(X_i'X_j)^2 - \frac{2}{P_n^3}\sum_{i,j,k}^{*} X_i'X_jX_j'X_k + \frac{1}{P_n^4}\sum_{i,j,k,l}^{*} X_i'X_jX_k'X_l,$$

where $P_n^r = n!/(n-r)!$, $\sum^*$ denotes summation over mutually different indices. Then we reject H_0 if nU_n exceeds the 5% upper quantile of $N(0,4)$ distribution.

For the test in [15](Sri for short), the test statistic is defined as follows:

$$W_n = \frac{n}{2}\cdot\left[\frac{c_n\cdot\frac{1}{p}\left[trS^2 - \frac{1}{n}(trS)^2\right]}{\left(\frac{1}{p}trS\right)^2} - 1\right]$$

where $S = \frac{1}{n}XX'$, $c_n = \frac{n^2}{(n-1)(n+2)}$. According to the limiting distribution of W_n, we reject H_0 if W_n exceeds the 5% upper quantile of $N(0,1)$ distribution. As for empirical powers, we generate sample data from two alternatives:

- **Power 1:** Σ is diagonal with half of its diagonal elements 0.5 and half 1. This power scenario is denoted by Power 1;
- **Power 2:** Σ is diagonal with 1/4 of its diagonal elements 0.5 and 3/4 equal to 1. This power scenario is denoted by Power 2.

Table 1 to 3 report the empirical sizes and powers of four tests for both Gaussian and Non-Gaussian data. Table 4 and 5 show the test results with estimated fourth moment $\hat{\nu}_4$.

TABLE 1
Test size for both Gaussian (Scenario 1) and Non-Gaussian Data (Scenario 2)

	Size (Gaussian)				Size (Non-Gaussian)			
(p,n)	Sri	Chen	John	QLRT	Sri	Chen	John	QLRT
(320,64)	0.0480	0.0539	0.0492	0.0998	0.1828	0.0584	0.0566	0.1084
(640,64)	0.0504	0.0538	0.0515	0.0668	0.1875	0.0594	0.0598	0.0735
(960,64)	0.0532	0.0581	0.0544	0.0620	0.1869	0.0580	0.0551	0.0631
(1280,64)	0.0519	0.0603	0.0530	0.0568	0.1856	0.0570	0.0517	0.0605
(1600,64)	0.0529	0.0571	0.0539	0.0593	0.1811	0.0555	0.0536	0.0580
(2400,64)	0.0493	0.0536	0.0501	0.0506	0.1790	0.0581	0.0533	0.0564
(3200,64)	0.0472	0.0538	0.0481	0.0503	0.1757	0.0518	0.0503	0.0522

It can be seen from the above results that both John's test and QLRT perform well with respect to sizes and powers. Estimated fourth moment $\hat{\nu}_4$ does not cause negative effect on the performance of these two test statistics. Empirical powers under Power 1 are in general higher than under Power 2 because of more significant difference between H_0 and H_1. John's test performs slightly better than Chen's method. In all tested scenarios, the QLRT dominates the other two tests in term of power even though the difference is quite marginal. Srivastava's

TABLE 2
Power 1 for both Gaussian and Non-Gaussian Data

	Power1 (Gaussian)				Power1 (Non-Gaussian)			
(p, n)	Sri	Chen	John	QLRT	Sri	Chen	John	QLRT
(320,64)	0.9571	0.9532	0.9580	0.9777	0.9909	0.9476	0.9538	0.9701
(640,64)	0.9595	0.9542	0.9602	0.9638	0.9927	0.9566	0.9603	0.9653
(960,64)	0.9598	0.9569	0.9604	0.9647	0.9923	0.9524	0.9589	0.9608
(1280,64)	0.9609	0.9569	0.9615	0.9656	0.9927	0.9529	0.9599	0.9620
(1600,64)	0.9583	0.9539	0.9588	0.9627	0.9925	0.9557	0.9622	0.9642
(2400,64)	0.9588	0.9542	0.9591	0.9615	0.9910	0.9497	0.9567	0.9577
(3200,64)	0.9617	0.9576	0.9624	0.9625	0.9909	0.9529	0.9610	0.9611

TABLE 3
Power 2 for both Gaussian and Non-Gaussian Data

	Power2 (Gaussian)				Power2 (Non-Gaussian)			
(p, n)	Sri	Chen	John	QLRT	Sri	Chen	John	QLRT
(320,64)	0.6155	0.6117	0.6194	0.7352	0.8374	0.6044	0.6196	0.7299
(640,64)	0.6089	0.6065	0.6128	0.6562	0.8379	0.6051	0.6201	0.6601
(960,64)	0.6201	0.6144	0.6231	0.6482	0.8394	0.6121	0.6298	0.6502
(1280,64)	0.6076	0.6043	0.6129	0.6256	0.8483	0.6133	0.6206	0.6416
(1600,64)	0.6194	0.6146	0.6231	0.6378	0.8433	0.6143	0.6330	0.6407
(2400,64)	0.6171	0.6099	0.6210	0.6291	0.8425	0.6110	0.6261	0.6304
(3200,64)	0.6212	0.6190	0.6251	0.6301	0.8413	0.6143	0.6266	0.6319

TABLE 4
Test Size and Power with $\hat{\nu}_4$ For Gaussian Data (Scenario 1)

	Size		Power1		Power2	
(p, n)	John	QLRT	John	QLRT	John	QLRT
(320,64)	0.0491	0.1051	0.9875	0.9927	0.7062	0.8066
(640,64)	0.0507	0.0664	0.9882	0.9911	0.7070	0.7429
(960,64)	0.0470	0.0581	0.9892	0.9903	0.7060	0.7299
(1280,64)	0.0526	0.0574	0.9890	0.9889	0.7106	0.7218
(1600,64)	0.0516	0.0587	0.9881	0.9882	0.7136	0.7244
(2400,64)	0.0485	0.0520	0.9882	0.9882	0.7120	0.7198
(3200,64)	0.0507	0.0525	0.9888	0.9886	0.7129	0.7168

TABLE 5
Test Size and Power with $\hat{\nu}_4$ For Non-Gaussian Data (Scenario 2)

	Size		Power1		Power2	
(p, n)	John	QLRT	John	QLRT	John	QLRT
(320,64)	0.0541	0.1083	0.9916	0.9952	0.7613	0.8434
(640,64)	0.0596	0.0727	0.9940	0.9941	0.7569	0.7819
(960,64)	0.0543	0.0637	0.9948	0.9954	0.7578	0.7738
(1280,64)	0.0511	0.0593	0.9934	0.9948	0.7569	0.7685
(1600,64)	0.0537	0.0576	0.9945	0.9946	0.7601	0.7685
(2400,64)	0.0573	0.0600	0.9940	0.9939	0.7566	0.7597
(3200,64)	0.0543	0.0562	0.9946	0.9940	0.7651	0.7682

test performs slightly below John's test in the Gaussian case and still suffers from non-normality with non-negligible bias. Furthermore, we have recorded the execution time of these two tests within different scenarios and we find that Chen's method is more time-consuming due to more complicated computations.

3. Power of the tests

In this section we study the asymptotic power of the two tests. To begin with, some preliminary knowledge is introduced as follows.

3.1. Preliminary knowledge

Consider the re-normalized sample covariance matrix

$$\widetilde{A} = \sqrt{\frac{1}{n}}\left(\frac{1}{\sqrt{tr(\Sigma_p^2)}}Z'\Sigma_p Z - \frac{tr(\Sigma_p)}{\sqrt{tr(\Sigma_p^2)}}I_n\right),$$

where $Z = (z_{ij})_{p\times n}$ and $z_{ij}, i = 1, \cdots, p,\ j = 1, \cdots, n$ are i.i.d. real random variables with mean zero and variance one, I_n is the identity matrix of order n, Σ_p is a sequence of $p \times p$ non-negative definite matrices with bounded spectral norm. Assume the following limit exist,

(a) $\gamma = \lim_{p\to\infty} \frac{1}{p}\, tr(\Sigma_p)$,
(b) $\theta = \lim_{p\to\infty} \frac{1}{p}\, tr(\Sigma_p^2)$,
(c) $\omega = \lim_{p\to\infty} \frac{1}{p} \sum_{i=1}^{p} (\Sigma_{ii})^2$,

where Σ_{ii} denotes the i-th diagonal entry of Σ_p. It has been proven that, under the ultra-dimensional setting in [3], with probability one, the ESD of matrix $\widetilde{A}$, $F^{\widetilde{A}}$ converges to the semicircle law F with density

$$F'(x) = \begin{cases} \dfrac{1}{2\pi}\sqrt{4-x^2}, & \text{if } |x| \le 2, \\ 0, & \text{if } |x| > 2. \end{cases}$$

We denote the Stieltjes transform of the semicircle law F by $m(z)$. Let $\mathscr{S}$ denote any open region on the complex plane including $[-2, 2]$, the support of F and $\mathscr{M}$ be the set of functions which are analytic on $\mathscr{S}$. For any $f \in \mathscr{M}$, denote

$$G_n(f) \triangleq n \int_{-\infty}^{+\infty} f(x) d\left(F^{\widetilde{A}}(x) - F(x)\right) - \sqrt{\frac{n^3}{p}}\Phi_3(f)$$

where, for any positive integer k,

$$\Phi_k(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos(\theta))\cos(k\theta)\,\mathrm{d}\theta.$$

Limiting theory of the test statistics under the alternative H_1 is based on a new CLT for linear statistics of $\widetilde{A}$, provided in [10], as follows.

Theorem 3.1. *Suppose that*

(1) $Z = (z_{ij})_{p\times n}$ *where* $\{z_{ij} : i = 1, \cdots, p;\ j = 1, \cdots, n\}$ *are i.i.d. real random variables with* $\mathbb{E}z_{ij} = 0$, $\mathbb{E}z_{ij}^2 = 1$ *and* $\nu_4 = \mathbb{E}z_{ij}^4 < \infty$;
(2) (Σ_p) *is a sequence of* $p \times p$ *non-negative definite matrices with bounded spectral norm and the following limit exist,*

(a) $\gamma = \lim_{p\to\infty} \frac{1}{p}\, tr(\Sigma_p)$,

(b) $\theta = \lim_{p\to\infty} \frac{1}{p}\, tr(\Sigma_p^2)$,

(c) $\omega = \lim_{p\to\infty} \frac{1}{p} \sum_{i=1}^{p} (\Sigma_{ii})^2$;

(3) $p/n \to \infty$ as $n \to \infty$, $n^3/p = O(1)$.

Then, for any $f_1, \cdots, f_k \in \mathscr{M}$, the finite dimensional random vector $(G_n(f_1), \cdots, G_n(f_k))$ converges weakly to a Gaussian vector $(Y(f_1), \cdots, Y(f_k))$ with mean function

$$\mathbb{E}Y(f) = \frac{1}{4}\left(f(2) + f(-2)\right) - \frac{1}{2}\Phi_0(f) + \frac{\omega}{\theta}(\nu_4 - 3)\Phi_2(f),$$

and covariance function

$$\begin{aligned} cov\left(Y(f_1), Y(f_2)\right) &= \frac{\omega}{\theta}(\nu_4 - 3)\Phi_1(f_1)\Phi_1(f_2) + 2\sum_{k=1}^{\infty} k\Phi_k(f_1)\Phi_k(f_2) \\ &= \frac{1}{4\pi^2}\int_{-2}^{2}\int_{-2}^{2} f_1'(x) f_2'(y) H(x,y)\,\mathrm{d}x\,\mathrm{d}y, \end{aligned}$$

where

$$\Phi_k(f) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos\theta)e^{ik\theta}\,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(2\cos\theta)\cos k\theta\,\mathrm{d}\theta,$$

$$H(x,y) = \frac{\omega}{\theta}(\nu_4 - 3)\sqrt{4 - x^2}\sqrt{4 - y^2} + 2\log\left(\frac{4 - xy + \sqrt{(4 - x^2)(4 - y^2)}}{4 - xy - \sqrt{(4 - x^2)(4 - y^2)}}\right).$$

The proofs of Theorem 3.2 and 3.3 about the power of the two test statistics are based on two lemmas derived from this CLT.

Lemma 3.1. *Let $\{\widetilde{\lambda}_i,\ 1 \le i \le n\}$ be eigenvalues of matrix $\widetilde{A} = \sqrt{\frac{1}{n}}\left(\frac{1}{\sqrt{tr(\Sigma_p^2)}} Z'\Sigma_p Z - \frac{tr(\Sigma_p)}{\sqrt{tr(\Sigma_p^2)}} I_n\right)$, where Z, Σ_p satisfies the assumptions in Theorem 3.1, then*

$$\begin{pmatrix} \sum_{i=1}^{n} \widetilde{\lambda}_i^2 - n - \left(\frac{\omega}{\theta}(\nu_4 - 3) + 1\right) \\ \sum_{i=1}^{n} \widetilde{\lambda}_i \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & \frac{\omega}{\theta}(\nu_4 - 3) + 2 \end{pmatrix}\right)$$

as $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$,

Lemma 3.2. *Let $\{\widetilde{\lambda}_i,\ 1 \le i \le n\}$ be eigenvalues of matrix $\widetilde{A} = \sqrt{\frac{1}{n}}\left(\frac{1}{\sqrt{tr(\Sigma_p^2)}} Z'\Sigma_p Z - \frac{tr(\Sigma_p)}{\sqrt{tr(\Sigma_p^2)}} I_n\right)$, where Z, Σ_p satisfies the assumptions in Theorem 3.1, then*

$$\begin{pmatrix} \sum_{i=1}^{n} \widetilde{\lambda}_i \\ \sqrt{\frac{p}{n}}\sum_{i=1}^{n} \log\left(\gamma + \widetilde{\lambda}_i\sqrt{\frac{n\theta}{p}}\right) - \sqrt{pn}\log(\gamma) + \frac{\theta}{2\gamma^2}\sqrt{\frac{n^3}{p}} + \left(\left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{n^2}{p} + \frac{\theta + \omega(\nu_4 - 3)}{2\gamma^2}\right)\sqrt{\frac{n}{p}} \end{pmatrix}$$

$$= \xi_n + o_p(1),$$

where

$$\xi_n \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix} \frac{\omega}{\theta}(\nu_4-3)+2 & \left(\frac{\omega}{\theta}(\nu_4-3)+2\right)\left(\frac{\sqrt{\theta}}{\gamma}+\frac{\theta\sqrt{\theta}}{\gamma^3}\frac{n}{p}\right) \\ \left(\frac{\omega}{\theta}(\nu_4-3)+2\right)\left(\frac{\sqrt{\theta}}{\gamma}+\frac{\theta\sqrt{\theta}}{\gamma^3}\frac{n}{p}\right) & \frac{\left(\frac{\omega}{\theta}(\nu_4-3)+2\right)\theta}{\gamma^2}+\frac{\left(\frac{2\omega}{\theta}(\nu_4-3)+5\right)\theta^2 n}{\gamma^4 p}\end{pmatrix}\right)$$

as $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$.

The proofs of these two lemma are postponed to Appendix A.

3.2. John's test

Suppose that an i.i.d. p-dimensional sample vectors $X_1, \cdots, X_n$ follow the multivariate distribution with covariance matrix Σ_p. To explore the power of John's test under the alternative hypothesis $H_1 : \Sigma_p \neq \sigma^2 I_p$, we assume that the $X_j's$ in X have representation $X_j = \Sigma_p^{1/2} Z_j$, so as $S = \frac{1}{n}\Sigma_p^{1/2} ZZ' \Sigma_p^{1/2}$, where $Z = \{Z_1, \cdots, Z_n\} = \{z_{ij}\}_{1\le i\le p, 1\le j\le n}$ is a $p \times n$ matrix with i.i.d. entries z_{ij} satisfying $\mathbb{E}(z_{ij}) = 0$, $\mathbb{E}(z_{ij}^2) = 1$ and $\mathbb{E}(|z_{ij}|^4) = \nu_4 < +\infty$. Then John's test statistic is

$$U = \frac{p^{-1}\sum_{i=1}^{p}(l_i - \bar{l})^2}{\bar{l}^2},$$

where $\{l_i,\ 1 \le i \le p\}$ are eigenvalues of the p-dimensional sample covariance matrix $S = \frac{1}{n}\Sigma_p^{1/2} ZZ' \Sigma_p^{1/2}$. The main result of the power of John's test is as follows.

Theorem 3.2. *Assume $X_1, \cdots, X_n$ are i.i.d. p-dimensional sample vectors follow multivariate distribution with covariance matrix Σ_p, $X = \Sigma_p^{1/2} Z$ where $Z = \{z_{ij}\}$ is a $p\times n$ matrix with i.i.d. entries z_{ij} satisfying $\mathbb{E}(z_{ij}) = 0$, $\mathbb{E}(z_{ij}^2) = 1$, $\mathbb{E}|z_{ij}|^4 = \nu_4 < \infty$, Σ_p is a sequence of $p \times p$ non-negative definite matrices with bounded spectral norm and the following limit exist,*

(a) $\gamma = \lim_{p\to\infty} \frac{1}{p}\, tr(\Sigma_p)$,
(b) $\theta = \lim_{p\to\infty} \frac{1}{p}\, tr(\Sigma_p^2)$,
(c) $\omega = \lim_{p\to\infty} \frac{1}{p}\sum_{i=1}^{p}(\Sigma_{ii})^2$,

then when $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$,

$$nU - p - \left(\frac{\theta}{\gamma^2} - 1\right) n \xrightarrow{d} N\left(\frac{\theta + \omega(\nu_4 - 3)}{\gamma^2},\ \frac{4\theta^2}{\gamma^4}\right). \tag{3.1}$$

Note that the theorem above reveals the limit distribution of John's test statistic under alternative hypothesis H_1. Nevertheless, if let $\Sigma_p = \sigma^2 I_p$, then $\gamma = \sigma^2$, $\theta = \omega = \sigma^4$, Theorem 3.2 reduces to Theorem 2.2, which states the null distribution of John's test statistic under H_0. With the two limit distributions of John's test statistic under H_0 and H_1, power of the test is derived as below.

Proposition 3.1. *With the same assumptions as in Theorem 3.2, when* $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$, *the power of John's test*

$$\beta_{John}(H_1) = 1 - \Phi\left(\frac{\gamma^2}{\theta}Z_\alpha + \frac{\gamma^2(\nu_4-2)-\theta-\omega(\nu_4-3)}{2\theta} + \frac{(\gamma^2-\theta)n}{2\theta}\right) \to 1,$$

where α is the nominal test level, Z_α, $\Phi(\cdot)$ are the alpha upper quantile and cdf of standard normal distribution respectively.

For John's test statistic U, under H_0,

$$nU - p \xrightarrow{d} N(\nu_4 - 2, 4),$$

under H_1,

$$nU - p - \left(\frac{\theta}{\gamma^2} - 1\right)n \xrightarrow{d} N\left(\frac{\theta+\omega(\nu_4-3)}{\gamma^2}, \frac{4\theta^2}{\gamma^4}\right),$$

$$\begin{aligned}
\beta_{\text{John}}(H_1) &= P\left(\frac{nU-p-(\nu_4-2)}{2} > Z_\alpha \middle| H_1\right) \\
&= P\left(\frac{nU-p-n\left(\frac{\theta}{\gamma^2}-1\right)-\frac{\theta+\omega(\nu_4-3)}{\gamma^2}}{\frac{2\theta}{\gamma^2}}\right. \\
&\qquad\qquad \left. > \frac{2Z_\alpha+(\nu_4-2)-n\left(\frac{\theta}{\gamma^2}-1\right)-\frac{\theta+\omega(\nu_4-3)}{\gamma^2}}{\frac{2\theta}{\gamma^2}}\right) \\
&= 1-\Phi\left(\frac{\gamma^2}{\theta}Z_\alpha + \frac{\gamma^2(\nu_4-2)-\theta-\omega(\nu_4-3)}{2\theta} + \frac{(\gamma^2-\theta)n}{2\theta}\right),
\end{aligned}$$

since $\gamma^2 \le \theta$, Proposition 3.1 follows.

The proof of Theorem 3.2 is based on Lemma 3.1.

Proof. Denote the eigenvalues of $p \times p$ matrix $S_n = \frac{1}{n}XX' = \frac{1}{n}Z\Sigma_p Z'$ in descending order by $\{l_i,\ 1 \le i \le p\}$, and eigenvalues of $n \times n$ matrix $\widetilde{A} = \sqrt{\frac{1}{n}}\left(\frac{1}{\sqrt{tr(\Sigma_p^2)}}Z'\Sigma_p Z - \frac{tr(\Sigma_p)}{\sqrt{tr(\Sigma_p^2)}}I_n\right)$ by $\{\widetilde{\lambda}_i,\ 1 \le i \le n\}$. Since $p > n$, S_n has $p-n$ zero eigenvalues and the remaining n non-zero eigenvalues l_i are related with $\widetilde{\lambda}_i$ as

$$\sqrt{\frac{1}{n}tr(\Sigma_p^2)}\widetilde{\lambda}_i + \frac{1}{n}tr(\Sigma_p) = l_i, \quad 1 \le i \le n.$$

We have, for John's test statistic

$$\begin{aligned}
U &= \left(\frac{1}{p}\sum_{i=1}^n\left(\sqrt{\frac{1}{n}tr(\Sigma_p^2)}\widetilde{\lambda}_i + \frac{1}{n}tr(\Sigma_p)\right)^2\right)\Bigg/\left(\frac{1}{p}\sum_{i=1}^n\left(\sqrt{\frac{1}{n}tr(\Sigma_p^2)}\widetilde{\lambda}_i + \frac{1}{n}tr(\Sigma_p)\right)\right)^2 - 1 \\
&= \frac{\theta\sum_{i=1}^n\widetilde{\lambda}_i^2 + 2\gamma\sqrt{\frac{p\theta}{n}}\sum_{i=1}^n\widetilde{\lambda}_i + p\gamma^2}{\left(\sqrt{\frac{\theta}{p}}\sum_{i=1}^n\widetilde{\lambda}_i + \sqrt{n}\gamma\right)^2} - 1.
\end{aligned}$$

Define function $G(u,v) = \dfrac{\theta u + 2\gamma\sqrt{\frac{p\theta}{n}}v + p\gamma^2}{(\sqrt{\frac{\theta}{p}}v + \sqrt{n}\gamma)^2} - 1$, then John's test statistic can be written as

$$U = G\left(u = \sum_{i=1}^{n} \widetilde{\lambda}_i^2, v = \sum_{i=1}^{n} \widetilde{\lambda}_i\right).$$

According to Lemma 3.1, when $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$,

$$\begin{pmatrix} \sum_{i=1}^n \widetilde{\lambda}_i^2 - n - \left(\frac{\omega}{\theta}(\nu_4 - 3) + 1\right) \\ \sum_{i=1}^n \widetilde{\lambda}_i \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & \frac{\omega}{\theta}(\nu_4 - 3) + 2 \end{pmatrix}\right)$$

Then by the Delta Method,

$$n\left(U - G(u,v)|_{u=n+\frac{\omega}{\theta}(\nu_4-3)+1, v=0}\right) \xrightarrow{d} N\left(0,\ n^2 \nabla G \begin{pmatrix} 4 & 0 \\ 0 & \frac{\omega}{\theta}(\nu_4 - 3) + 2 \end{pmatrix} \nabla G'\right),$$

where $\nabla G = \left(\dfrac{\partial U}{\partial u}, \dfrac{\partial U}{\partial \nu}\right)\Bigg|_{u=n+\frac{\omega}{\theta}(\nu_4-3)+1, v=0}$ is the corresponding gradient vector.

We have, for $(u,v) = \left(n + \frac{\omega}{\theta}(\nu_4 - 3) + 1, 0\right)$,

$$G = \frac{p}{n} + \frac{\theta}{\gamma^2} - 1 + \frac{(\omega(\nu_4 - 3) + \theta)}{n\gamma^2},$$

and

$$\begin{aligned} &\nabla G \begin{pmatrix} 4 & 0 \\ 0 & \frac{\omega}{\theta}(\nu_4 - 3) + 2 \end{pmatrix} \nabla G' \\ &= \frac{4\theta^2}{n^2\gamma^4} + \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\left(\frac{4\theta\left(\theta + \omega(\nu_4 - 3) + n\theta\right)^2}{\gamma^6 n^3 p}\right). \end{aligned}$$

The result thus follows. □

3.3. Quasi-likelihood ratio test

Consider the Quasi-LRT statistic $\mathcal{L}_n$ in (1.4) based on the eigenvalues of n-dimensional matrix $\frac{1}{p}X'X$. Similarly with John's test statistic, it can be seen that, under the alternative hypothesis H_1, the $\mathcal{L}_n$ statistic can be represented as

$$\mathcal{L}_n = \frac{p}{n} \log \frac{\left(\frac{1}{n}\sum_{i=1}^n \eta_i\right)^n}{\prod_{i=1}^n \eta_i}$$

where $\{\eta_i,\ 1 \le i \le n\}$ are eigenvalues of $\frac{1}{p}Z'\Sigma_p Z$. The main result of the power of the Quasi-LRT test is as follows.

Theorem 3.3. *With the same assumptions as in Theorem 3.2, when* $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$,

$$\mathcal{L}_n - \left(\frac{\theta}{2\gamma^2} n + \left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3} \right) \frac{n^2}{p} \right) \xrightarrow{d} N \left(\frac{\theta}{2\gamma^2} + \frac{\omega}{2\gamma^2}(\nu_4 - 3),\ \frac{\theta^2}{\gamma^4} \right). \quad (3.2)$$

Note that the theorem above reveals the limit distribution of the Quasi-LRT statistic under alternative hypothesis H_1. Nevertheless, if let $\Sigma_p = \sigma^2 I_p$, then $\gamma = \sigma^2$, $\theta = \omega = \sigma^4$, Theorem 3.3 reduces to Theorem 2.3, which states the null distribution of the Quasi-LRT test statistic under H_0. Similarly, with the two limit distributions of QLRT statistic under H_0 and H_1, power of the test is derived as below.

Proposition 3.2. *With the same assumptions as in Theorem 3.2, when $p/n \to \infty,\ n \to \infty,\ n^3/p = O(1)$, the power of QLRT $\beta_{QLRT}(H_1)$ is*

$$1 - \Phi\left(\frac{\gamma^2}{\theta} Z_\alpha + \left(\frac{\gamma^2 - \theta}{2\theta} \right) n + \left(\frac{\gamma^2}{6\theta} - \frac{\theta}{2\gamma^2} + \frac{\sqrt{\theta}}{3\gamma} \right) \frac{n^2}{p} + \left(\frac{\gamma^2(\nu_4 - 2) - \theta - \omega(\nu_4 - 3)}{2\theta} \right) \right) \to 1,$$

where α is the nominal test level, Z_α, $\Phi(\cdot)$ are the alpha upper quantile and cdf of standard normal distribution respectively.

For QLRT statistic $\mathcal{L}$, under H_0,

$$\mathcal{L}_n - \frac{n}{2} - \frac{n^2}{6p} \xrightarrow{d} N \left(\frac{\nu_4 - 2}{2}, 1 \right),$$

under H_1,

$$\mathcal{L}_n - \frac{\theta}{2\gamma^2} n - \left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3} \right) \frac{n^2}{p} \xrightarrow{d} N \left(\frac{\theta}{2\gamma^2} + \frac{\omega}{2\gamma^2}(\nu_4 - 3),\ \frac{\theta^2}{\gamma^4} \right).$$

$$\begin{aligned}
&\beta_{\mathrm{QLRT}}(H_1) \\
&\quad = P\left(\mathcal{L}_n - \frac{n}{2} - \frac{n^2}{6p} - \frac{\nu_4 - 2}{2} > Z_\alpha \,\middle|\, H_1 \right) \\
&\quad = P\left(\frac{\mathcal{L}_n - \frac{\theta}{2\gamma^2} n - \left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3} \right) \frac{n^2}{p} - \left(\frac{\theta}{2\gamma^2} + \frac{\omega}{2\gamma^2}(\nu_4 - 3) \right)}{\frac{\theta}{\gamma^2}} \right. \\
&\quad \left. > \frac{Z_\alpha + \frac{n}{2} + \frac{n^2}{6p} + \frac{\nu_4 - 2}{2} - \frac{\theta}{2\gamma^2} n - \left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3} \right) \frac{n^2}{p} - \left(\frac{\theta}{2\gamma^2} + \frac{\omega}{2\gamma^2}(\nu_4 - 3) \right)}{\frac{\theta}{\gamma^2}} \right) \\
&\quad = 1 - \Phi\left(\frac{\gamma^2}{\theta} Z_\alpha + \left(\frac{\gamma^2 - \theta}{2\theta} \right) n + \left(\frac{\gamma^2}{6\theta} - \frac{\theta}{2\gamma^2} + \frac{\sqrt{\theta}}{3\gamma} \right) \frac{n^2}{p} \right. \\
&\qquad \left. + \left(\frac{\gamma^2(\nu_4 - 2) - \theta - \omega(\nu_4 - 3)}{2\theta} \right) \right),
\end{aligned}$$

since $\gamma^2 \leq \theta$, Proposition 3.2 follows.

The proof of Theorem 3.3 is based on lemma 3.2.

Proof. Denote the eigenvalues of $n \times n$ matrix $\frac{1}{p}X'X = \frac{1}{p}Z'\Sigma_p Z$ in descending order by $\widetilde{l}_i (1 \le i \le n)$, and eigenvalues of $n \times n$ matrix $\widetilde{A} = \sqrt{\frac{1}{n}}\left(\frac{1}{\sqrt{tr(\Sigma_p^2)}}Z'\Sigma_p Z - \frac{tr(\Sigma_p)}{\sqrt{tr(\Sigma_p^2)}}I_n\right)$ by $\widetilde{\lambda}_i (1 \le i \le n)$. These eigenvalues are related as

$$\sqrt{\frac{n\, tr(\Sigma_p^2)}{p^2}}\widetilde{\lambda}_i + \frac{1}{p}\, tr(\Sigma_p) = \widetilde{l}_i, \quad 1 \le i \le n.$$

We have, for the Quasi-LRT test statistic

$$\begin{aligned}
\mathcal{L}_n &= \frac{p}{n}\log\left[\left(\frac{1}{n}\sum_{i=1}^{n}\eta_i\right)^n \Big/ \prod_{i=1}^{n}\eta_i\right] \\
&= p\log\left(\gamma + \sqrt{\frac{\theta}{np}}\sum_{i=1}^{n}\widetilde{\lambda}_i\right) - \frac{p}{n}\sum_{i=1}^{n}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}\widetilde{\lambda}_i\right),
\end{aligned}$$

Define the function

$$G(u, v) = p\log\left(\gamma + \sqrt{\frac{\theta}{np}}u\right) - \sqrt{\frac{p}{n}}v,$$

then the Quasi-LRT test statistic can be written as

$$\mathcal{L}_n = G\left(u = \sum_{i=1}^{n}\widetilde{\lambda}_i, v = \sqrt{\frac{p}{n}}\sum_{i=1}^{n}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}\widetilde{\lambda}_i\right)\right).$$

According to Lemma 3.2, when $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$,

$$\begin{pmatrix} \sum_{i=1}^{n}\widetilde{\lambda}_i \\ \sqrt{\frac{p}{n}}\sum_{i=1}^{n}\log\left(\gamma + \widetilde{\lambda}_i\sqrt{\frac{n\theta}{p}}\right) - \sqrt{pn}\log(\gamma) + \frac{\theta}{2\gamma^2}\sqrt{\frac{n^3}{p}} + \left(\left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{n^2}{p} + \frac{\theta+\omega(\nu_4-3)}{2\gamma^2}\right)\sqrt{\frac{n}{p}} \end{pmatrix}$$
$$= \xi_n + o_p(1),$$

where

$$\xi_n \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\omega}{\theta}(\nu_4 - 3) + 2 & \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\left(\frac{\sqrt{\theta}}{\gamma} + \frac{\theta\sqrt{\theta}}{\gamma^3}\frac{n}{p}\right) \\ \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\left(\frac{\sqrt{\theta}}{\gamma} + \frac{\theta\sqrt{\theta}}{\gamma^3}\frac{n}{p}\right) & \frac{\left(\frac{\omega}{\theta}(\nu_4-3)+2\right)\theta}{\gamma^2} + \frac{\left(\frac{2\omega}{\theta}(\nu_4-3)+5\right)\theta^2 n}{\gamma^4 p} \end{pmatrix}\right)$$

By the Delta Method,

$$\mathcal{L}_n - G(u,v)\Big|_{u=0, v=\sqrt{pn}\log(\gamma) - \frac{\theta}{2\gamma^2}\sqrt{\frac{n^3}{p}} - \left(\left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{n^2}{p} + \frac{\theta}{2\gamma^2} + \frac{\omega}{2\gamma^2}(\nu_4-3)\right)\sqrt{\frac{n}{p}}} \xrightarrow{d}$$
$$N\left(0,\ \nabla G\begin{pmatrix} \frac{\omega}{\theta}(\nu_4 - 3) + 2 & \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\left(\frac{\sqrt{\theta}}{\gamma} + \frac{\theta\sqrt{\theta}}{\gamma^3}\frac{n}{p}\right) \\ \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\left(\frac{\sqrt{\theta}}{\gamma} + \frac{\theta\sqrt{\theta}}{\gamma^3}\frac{n}{p}\right) & \frac{\left(\frac{\omega}{\theta}(\nu_4-3)+2\right)\theta}{\gamma^2} + \frac{\left(\frac{2\omega}{\theta}(\nu_4-3)+5\right)\theta^2 n}{\gamma^4 p} \end{pmatrix}\nabla G'\right),$$

where $\nabla G = \left(\frac{\partial U}{\partial u}, \frac{\partial U}{\partial \nu}\right)\Big|_{u=0, v=\sqrt{pn}\log(\gamma)-\frac{\theta}{2\gamma^2}\sqrt{\frac{n^3}{p}}-\left(\left(\frac{\theta^2}{2\gamma^4}-\frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{n^2}{p}+\frac{\theta+\omega(\nu_4-3)}{2\gamma^2}\right)\sqrt{\frac{n}{p}}}$ is the corresponding gradient vector.

Then we have, for

$$(u,v) = \left(0, \sqrt{pn}\log(\gamma) - \frac{\theta}{2\gamma^2}\sqrt{\frac{n^3}{p}} - \left(\left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{n^2}{p} + \frac{\theta + \omega(\nu_4 - 3)}{2\gamma^2}\right)\sqrt{\frac{n}{p}}\right),$$

$$G(u,v) = \frac{\theta}{2\gamma^2}n + \left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{n^2}{p} + \frac{\theta + \omega(\nu_4 - 3)}{2\gamma^2},$$

and

$$\nabla G \; Cov\,(u,v)\,\nabla G' = \frac{\theta^2}{\gamma^4}.$$

The result thus follows. □

3.4. Simulation experiments

Empirical power of the two tests are shown in this section to testify the theoretical results presented in Proposition 3.1 and 3.2. Specifically, we consider two different scenarios to generate sample data:

(1) $\{Z_j,\ 1 \le j \le n\}$ i.i.d p-dimensional random vector generated from multivariate normal population $N_p(\mathbf{0}, I_p)$, $\mathbb{E}z_{ij}^4 = \nu_4 = 3$, $X_j = \Sigma_p^{1/2} Z_j$, $1 \le j \le n$;

(2) $\{z_{ij},\ 1 \le i \le p,\ 1 \le j \le n\}$ i.i.d follow $Gamma(4,2) - 2$ distribution, then $\mathbb{E}z_{ij} = 0$, $\mathbb{E}z_{ij}^2 = 1$, $\mathbb{E}z_{ij}^4 = \nu_4 = 4.5$. $X_{p\times n} = \Sigma_p^{1/2} Z_{p\times n}$.

To cover multiple alternative hypothesis, Σ_p is configured as a diagonal matrix with elements 0.5 and 1. The proportion of "1" is δ. The nominal test level is set as $\alpha = 0.05$. $(p,n) = (2400, 64)$ and empirical power are generated from 5000 replications. Theoretical values are displayed for comparison.

It can be seen from Table 6 that the empirical and theoretical power coincide with each other and both tests have very large power even when δ is small.

3.5. On the dimension-proof property of John's test under the alternative

In this paper, we have proved in Theorem 3.2 that, assume $X = \Sigma_p^{1/2} Z$ where $Z = \{z_{ij}\}$ is a $p\times n$ matrix with i.i.d. entries z_{ij} satisfying $\mathbb{E}(z_{ij}) = 0$, $\mathbb{E}(z_{ij}^2) = 1$, $\mathbb{E}|z_{ij}|^4 = \nu_4 < \infty$, Σ_p is a sequence of $p \times p$ non-negative definite matrices with bounded spectral norm and the following limit exist,

TABLE 6
Empirical(Empi) and Theoretical(Thry) Power of two tests

δ	Gaussian				Non-Gaussian			
	John's test		QLRT		John's test		QLRT	
	Empi	Thry	Empi	Thry	Empi	Thry	Empi	Thry
0	0.046	0.050	0.049	0.050	0.051	0.050	0.052	0.050
0.1	0.738	0.745	0.727	0.759	0.736	0.746	0.727	0.761
0.2	0.958	0.953	0.954	0.959	0.950	0.954	0.951	0.960
0.3	0.984	0.979	0.982	0.982	0.981	0.979	0.981	0.982
0.4	0.978	0.976	0.978	0.980	0.978	0.976	0.978	0.980
0.5	0.958	0.953	0.958	0.959	0.951	0.954	0.950	0.960

(a) $\gamma = \lim_{p\to\infty} \frac{1}{p} tr(\Sigma_p)$,
(b) $\theta = \lim_{p\to\infty} \frac{1}{p} tr(\Sigma_p^2)$,
(c) $\omega = \lim_{p\to\infty} \frac{1}{p} \sum_{i=1}^{p} (\Sigma_{ii})^2$,

where Σ_{ii} is the i-th diagonal entry of Σ_p, then when $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$,

$$nU - p - \left(\frac{\theta}{\gamma^2} - 1\right) n \xrightarrow{d} N\left(\frac{\theta + \omega(\nu_4 - 3)}{\gamma^2},\ \frac{4\theta^2}{\gamma^4}\right).$$

To check the *dimension-proof* property of the power of John's test, we need to look into the asymptotic distributions of U under the alternative hypothesis with different (n,p)-asymptotics.

For this exploration, we consider two different types of alternative hypothesis, which are, the single and multi-spiked population model studied in [11] and [16] with $n, p \to, p/n \to c \in (0,\infty)$, and the general covariance matrix $\Sigma = \mathrm{diag}(\lambda_1, \cdots, \lambda_p)$ model in [13] with $n = O(p^\delta)$, $0 < \delta \le 1$.

a. Alternative with a spiked covariance matrix Σ_p whose eigenvalues of are all one except for a few fixed number of spikes, i.e.

$$H_1 : \mathrm{Spec}\,(\Sigma_p) = \mathrm{diag}\left(\underbrace{b_1, \cdots, b_1}_{n_1}, \cdots, \underbrace{b_k, \cdots, b_k}_{n_k}, \underbrace{1, \cdots, 1}_{p-M}\right),$$

where the multiplicity numbers $n_i's$ are fixed and satisfying $\sum_{i=1}^k n_i = M$. [16] proved that, under H_1, when $p, n \to \infty$, $p/n \to c \in (0,\infty)$,

$$nU - p - \frac{n}{p} \sum_{i=1}^{k} n_i (b_i - 1)^2 \xrightarrow{d} N(\nu_4 - 2, 4).$$

An important observation here is that, the power of John's test drops significantly when p/n becomes larger since the term $\frac{n}{p} \sum_{i=1}^{k} n_i (b_i - 1)^2$ vanishes when $p/n \to \infty$. Actually, this phenomenon has already been noticed in Figure 3 in [16]. For instance, when $p/n = 2.5$, the power already goes lower than 0.2.

In fact, this phenomenon can be also explained from our ultra-high dimensional point of view. Since all the spikes are with fixed multiplicity $n_i's$, which means, under H_1, all the essential parameters $\theta,\ \gamma,\ \omega$ remain the same with H_0. Therefore, the asymptotic distribution of U are the same under H_0 and H_1, thus explains the loss of power above in [16].
When the population mean is unknown, the situation is similar and the above consistency still holds. Conclusively, the *dimension-proof* property of the power of John's test holds here when the alternative hypothesis is multi-spiked population model.

b. For the general population model, since the testing problem remains invariant under orthogonal transformations, we may assume without loss of generality,
$$H_1 : \Sigma_p = \operatorname{diag}(\lambda_1, \cdots, \lambda_p),$$
define $a_i = (tr\,\Sigma_p^i)/p$, [13] proved in their Theorem 3.1 that if $X_1, \cdots, X_n$ are independently drawn from $N_p(\mu, \Sigma_p)$, $a_i \to a_i^0 \in (0, \infty)$ as $p \to \infty$, $n = O(p^\delta),\ 0 < \delta \le 1$, then
$$\frac{(n-1)^3}{(n-2)(n+1)} \cdot \frac{\frac{1}{p}\,tr(\widehat{S}_n^2)}{\left(\frac{1}{p}\,tr\,\widehat{S}_n\right)^2} - \frac{p(n-1)^2}{(n-2)(n+1)} - \frac{a_2}{a_1^2}(n-1) \to N(0, 4\tau_1^2), \tag{3.3}$$
where $\tau_1^2 = \frac{2n\left(a_4a_1^2 - 2a_1a_2a_3 + a_2^3\right)}{pa_1^6} + \frac{a_2^2}{a_1^4}$, $\widehat{S}_n = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})(X_i - \overline{X})'$.
It can be seen that as $p/n \to \infty$, $\tau_1^2 \to \frac{a_2^2}{a_1^4} = \frac{\theta^2}{\gamma^4}$. Since Σ_p is diagonal, $\omega = \theta = a_2,\ \gamma = a_1,\ \nu_4 = 3$ for Gaussian distribution. By simple calculations, it can be seen that, when $p/n \to \infty$, $n \to \infty$, $n^3/p = O(1)$, equation 3.3 is asymptotically equivalent to 4.2, the power of John's test with unknown population mean which is presented in the following section.
Conclusively, the *dimension-proof* property of the power of John's test holds here again when the alternative hypothesis is a general population model.

4. Generalization to the case when the population mean is unknown

Previously we have assumed that the data matrix consists of n i.i.d p dimensional centered random vectors $\{X_i\}_{1\le i\le n}$. Both test statistics are based on eigenvalues $\{l_i\}_{1\le i\le p}$ of the sample covariance matrix $S_n = \frac{1}{n}\sum_{i=1}^n X_iX_i'$. However in practice, the population mean $\mu = \mathbb{E}X_i$ is usually unknown and the sample covariance matrix should be defined as
$$\widehat{S}_n = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})(X_i - \overline{X})',$$
where $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is the sample mean. Then the corresponding John's test statistic becomes

$$\widehat{U} = \frac{1}{p} tr \left[\left(\frac{\widehat{S}_n}{(1/p)tr(\widehat{S}_n)} - I_p \right)^2 \right] = p \frac{\sum_{i=1}^p \hat{l}_i^2}{\left(\sum_{i=1}^p \hat{l}_i\right)^2} - 1,$$

where $\{\hat{l}_i\}_{1 \le i \le p}$ are the eigenvalues of $\widehat{S}_n$.

If we assume first that the data are normal-distributed, then $(n-1)\widehat{S}_n \sim$ Wishart$(n-1, \Sigma)$, that is, $\widehat{S}_n \overset{\mathcal{D}}{=\!=} \frac{1}{n-1} \sum_{j=1}^{n-1} Y_j Y_j'$, where $Y_j \sim N_p(\mathbf{0}, \Sigma), i.i.d.$. As a consequence, functionals of $\{\hat{l}_i\}_{1 \le i \le p}$, including $\sum_{i=1}^p \hat{l}_i^2$, $\sum_{i=1}^p \hat{l}_i$ etc, are in distribution equivalent to those of some centered sample covariance matrix with $n-1$ degree of freedom. Therefore we can get the asymptotic distribution of these functionals for $\widehat{S}_n$ by substituting $n-1$ for n in those for S_n.

Actually, this substitution principle has been well established in [18] for an arbitrary population distribution. [16] also applied this principle to get the CLT of John's test statistic with unknown population mean. Accordingly, the same procedure is adopted here to deduce the asymptotic distribution of John's test statistic under both the null and the alternative hypothesis in the ultra-high dimensional setting.

More precisely, denote $n-1$ by N, equations (2.3) and (3.1) in Theorem 2.2 and 3.2 become

$$N\widehat{U} - p \xrightarrow{d} N(\nu_4 - 2, 4). \tag{4.1}$$

and

$$N\widehat{U} - p - \left(\frac{\theta}{\gamma^2} - 1 \right) N \xrightarrow{d} N \left(\frac{\theta + \omega(\nu_4 - 3)}{\gamma^2}, \ \frac{4\theta^2}{\gamma^4} \right). \tag{4.2}$$

Next multiplying by $\frac{n}{N}$ to both sides of equations (4.1) and (4.2), we get

$$n\widehat{U} - \frac{n}{N} p - \frac{n}{N}(\nu_4 - 2) \xrightarrow{d} N(0, 4),$$

$$n\widehat{U} - \frac{n}{N} p - \frac{n}{N} \left(\frac{\theta + \omega(\nu_4 - 3)}{\gamma^2} \right) - \left(\frac{\theta}{\gamma^2} - 1 \right) n \xrightarrow{d} N \left(0, \ \frac{4\theta^2}{\gamma^4} \right).$$

These results coincide with the generalization results in [16] as well and the dimensional-proof property of John's test statistic holds still with unknown population mean.

As for the quasi-LRT statistic, since

$$\mathcal{L}_n = \frac{p}{n} \log \frac{\left(\frac{1}{n} \sum_{i=1}^n \eta_i\right)^n}{\prod_{i=1}^n \eta_i},$$

where $\{\eta_i\}_{1 \le i \le n}$ are the eigenvalues of n-dimensional matrix $\frac{1}{p} X'X$, a natural way of generalizing the QLRT statistic is

$$\widehat{\mathcal{L}}_N = \frac{p}{N} \log \frac{\left(\frac{1}{N} \sum_{i=1}^N \hat{\eta}_i\right)^N}{\prod_{i=1}^N \hat{\eta}_i},$$

where $\{\hat{\eta}_i\}_{1\le i\le N}$ are the N non-zero eigenvalues of n-dimensional matrix $\frac{1}{p}(X-\overline{X})'(X-\overline{X})$. Similarly, the same substitution principle applies under both the null and the alternatives. Equations in (2.4) and (3.2) in Theorem 2.3 and 3.3 become

$$\widehat{\mathcal{L}}_N - \frac{N}{2} - \frac{N^2}{6p} - \frac{\nu_4-2}{2} \xrightarrow{d} N\left(0,1\right) \tag{4.3}$$

and

$$\widehat{\mathcal{L}}_N - \left(\frac{\theta}{2\gamma^2}N + \left(\frac{\theta^2}{2\gamma^4} - \frac{\theta\sqrt{\theta}}{3\gamma^3}\right)\frac{N^2}{p}\right) \xrightarrow{d} N\left(\frac{\theta}{2\gamma^2} + \frac{\omega}{2\gamma^2}(\nu_4-3),\ \frac{\theta^2}{\gamma^4}\right), \tag{4.4}$$

respectively.

Notice that the unknown fourth moment ν_4 in equations 4.1 and 4.3 now corresponds to the centered random variables. The estimator given in Remark 2.1 still apply but to the centered variables, i.e.

$$\tilde{\nu}_4 = \frac{1}{np}\sum_{i=1}^{n}\sum_{j=1}^{p}(x_{ij}-\overline{x})^4.$$

This estimator is consistent and its substitution for the unknown ν_4 does not modify the corresponding limit distributions.

5. Discussions and auxiliary results

In summary, we found in the considered ultra-dimension ($p \gg n$) situations, QLRT is the most recommended procedure regarding its maximal power for sphericity test. However, from the application perspective where the dimension p and n are explicitly known, it becomes very difficult to decide which asymptotic scheme to use, namely, " p fixed, $n\to\infty$", "$p/n \to c \in (0,\infty),\ p,n\to\infty$", or "$p/n\to\infty,\ p,n\to\infty$" etc. Combining our study with the existing literature, we would like to recommend a *dimension-proof* procedure like John's test or Chen's test, with a slight preference for John's test as it has a slightly higher power and an easier implementation.

We conclude the paper by mentioning some surprising consequence of the main results of the paper as follows.

Corollary 5.1. *Assume $X=\{x_{ij}\}_{p\times n}$ are i.i.d. satisfying $\mathbb{E}(x_{ij})=0$, $\mathbb{E}(x_{ij}^2)=1$, $\mathbb{E}|x_{ij}|^4=\nu_4<\infty$, then when $n/p\to\infty$, $n,p\to\infty$,*

$$-\frac{2}{p}\log L_n - \frac{p}{2} - \frac{p^2}{6n} - \frac{\nu_4-2}{2} \xrightarrow{d} N\left(0,1\right).$$

where $-\frac{2}{p}\log L_n = \frac{n}{p}\log\left(\frac{\overline{l}^p}{\prod_{i=1}^p l_i}\right)$, $\{l_i\}_{1\le i\le p}$ are the eigenvalues of p dimensional sample covariance matrix $\frac{1}{n}XX'$.

Note that if we fix p while let $n \to \infty$, under normality assumption, the Corollary 5.1 reduces to

$$-\frac{2}{p}\log L_n - \frac{p+1}{2} \xrightarrow{d} N(0,1),$$

which is consistent with the classic LRT asymptotic, i.e. $-2\log L_n \xrightarrow{d} \chi^2_{\frac{1}{2}p(p+1)-1}$.

Corollary 5.2. *Assume $X = \{x_{ij}\}_{p\times n}$ are i.i.d. satisfying $\mathbb{E}(x_{ij}) = 0$, $\mathbb{E}(x_{ij}^2) = 1$, $\mathbb{E}|x_{ij}|^4 = \nu_4 < \infty$, then when $n/p \to \infty$, $n, p \to \infty$,*

$$nU - p \xrightarrow{d} N(\nu_4 - 2, 4).$$

Proof. Interchanging the role of n and p in Theorem 2.2, keeping all other assumptions unchanged, it can be seen that, when $n/p \to \infty$, $n, p \to \infty$,

$$p\widetilde{U} - n \xrightarrow{d} N(\nu_4 - 2, 4),$$

where

$$\widetilde{U} = \frac{n^{-1}\sum_{i=1}^n \eta_i^2}{\left(\frac{1}{n}\sum_{i=1}^n \eta_i\right)^2} - 1,$$

$\eta_i (1 \le i \le n)$ are eigenvalues of $n\times n$ matrix $\frac{1}{p}X'X$, l_i are eigenvalues of $\frac{1}{n}XX'$, then

$$\begin{aligned} p\widetilde{U} - n &= \frac{\frac{p}{n}\sum_{i=1}^n \eta_i^2}{\left(\frac{1}{n}\sum_{i=1}^n \eta_i\right)^2} - p - n \\ &= \frac{\frac{n}{p}\sum_{i=1}^p l_i^2}{\bar{l}^2} - n - p = nU - p. \end{aligned}$$ □

Henceforth, the *dimension-proof* property of John's test statistic, i.e. regardless of normality, under any (n, p)-asymptotic, $n/p \to [0, \infty]$, has been completely testified.

Appendix A: Technique lemmas and additional proofs

Lemma A.1. *In the central limit theorem of linear functions of eigenvalues of the re-normalized sample covariance matrix A when the dimension p is much larger than the sample size n derived by [4], Let $\mathscr{S}$ denote any open region on the complex plane including $[-2, 2]$, the support of the semicircle law $F(x)$, we denote the Stieltjes transform of the semicircle law F by $m(z)$. Let $\mathscr{M}$ be the set of functions which are analytic on $\mathscr{S}$, for any analytic function $f \in \mathcal{M}$, the mean correction term is defined as*

$$\frac{n}{2\pi i}\oint_{|m|=\rho} f\left(-m - m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m.$$

Define functions $f_1(x) = x^2$, $f_2(x) = x$, $f_3(x) = \frac{p}{n}\log(1 + \sqrt{\frac{n}{p}}x)$, then the mean correction term in equation (2.1) for these functions are as follows:

$$\frac{n}{2\pi i}\oint_{|m|=\rho} f_1\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m=\nu_4-2,$$

$$\frac{n}{2\pi i}\oint_{|m|=\rho} f_2\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m=0,$$

$$\frac{n}{2\pi i}\oint_{|m|=\rho} f_3\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m=-\frac{\nu_4-2}{2}+\frac{n^2}{3p}.$$

Proof. Since

$$\begin{aligned}\chi_n^{\mathrm{Calib}}(m) &\triangleq \frac{-\mathcal{B}+\sqrt{\mathcal{B}^2-4\mathcal{A}\mathcal{C}^{\mathrm{Calib}}}}{2\mathcal{A}},\ \mathcal{A}=m-\sqrt{\frac{n}{p}}(1+m^2),\\ \mathcal{B}&=m^2-1-\frac{n}{p}m(1+2m^2),\\ \mathcal{C}^{\mathrm{Calib}}&=\frac{m^3}{n}\left[\nu_4-2+\frac{m^2}{1-m^2}-2(\nu_4-1)m\sqrt{\frac{n}{p}}\right]-\sqrt{\frac{n}{p}}m^4,\end{aligned}$$

the integral's contour is taken as $|m|=\rho$ with $\rho<1$.

For $f_1(x)=x^2$, choose $\rho<\sqrt{\frac{n}{p}}<\sqrt{\frac{p}{n}}$,

$$\begin{aligned}&\frac{n}{2\pi i}\oint_{|m|=\rho} f_1\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m\\ &=\frac{n}{2\pi i}\oint_{|m|=\rho}\frac{(1-m^4)(1+m^2)}{m^4}\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\mathrm{d}m\\ &\quad\left(\text{denote } X:=\frac{1}{2}\left(-\mathcal{B}+\sqrt{\mathcal{B}^2-4\mathcal{A}\mathcal{C}^{\mathrm{Calib}}}\right)\right)\\ &=\frac{n}{2\pi i}\oint_{|m|=\rho}\frac{1+m^2}{m^4}\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\mathrm{d}m\\ &\quad(\text{Cauchy's Residue Theorem})\\ &=\nu_4-2.\end{aligned}$$

For $f_2(x)=x$, choose $\rho<\sqrt{\frac{n}{p}}<\sqrt{\frac{p}{n}}$,

$$\begin{aligned}&\frac{n}{2\pi i}\oint_{|m|=\rho} f_2\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m\\ &=\frac{n}{2\pi i}\oint_{|m|=\rho}(-m-m^{-1})\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\cdot\frac{1-m^2}{m^2}\mathrm{d}m\\ &=-\frac{n}{2\pi i}\oint_{|m|=\rho}\frac{1}{m^3}\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\mathrm{d}m\end{aligned}$$

$$= \; -\frac{1}{2!}\,\mathrm{d}^{(2)}\left(\frac{nX}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\right)\Bigg/\mathrm{d}\,m^2\Bigg|_{m=0}$$
$$= 0.$$

For $f_3(x) = \frac{p}{n}\log(1+\sqrt{\frac{n}{p}}x)$, choose $\rho < \sqrt{\frac{n}{p}} < \sqrt{\frac{p}{n}}$,

$$\begin{aligned}
&\frac{n}{2\pi i}\oint_{|m|=\rho} f_3\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m \\
&= \frac{p}{2\pi i}\oint_{|m|=\rho} \log\left(1-\sqrt{\frac{n}{p}}(m+m^{-1})\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m \\
&= \frac{p}{2\pi i}\oint_{|m|=\rho}\sum_{k=1}^{\infty}\left[-\frac{1}{k}\left(\sqrt{\frac{n}{p}}(m+m^{-1})\right)^k\right]\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\cdot\frac{1-m^2}{m^2}\,\mathrm{d}m \\
&= -\frac{1}{2\pi i}\oint_{|m|=\rho}\left[\sqrt{np}\cdot\frac{1-m^4}{m^3}+\frac{n}{2}\cdot\frac{(1-m^4)(1+m^2)}{m^4}+\frac{n}{3}\cdot\sqrt{\frac{n}{p}}\cdot\frac{(1-m^4)(1+m^2)^2}{m^5}\right. \\
&\quad\left.+\frac{n^2}{4p}\cdot\frac{(1-m^4)(1+m^2)^3}{m^6}\right]\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\,\mathrm{d}m+o\left(\frac{n^2}{p}\right) \\
&= -\frac{1}{2\pi i}\oint_{|m|=\rho}\left[\sqrt{np}\cdot\frac{1}{m^3}+\frac{n}{2}\cdot\frac{1+m^2}{m^4}+\frac{n}{3}\cdot\sqrt{\frac{n}{p}}\cdot\frac{1+2m^2}{m^5}\right. \\
&\quad\left.+\frac{n^2}{4p}\cdot\frac{1+3m^2+2m^4}{m^6}\right]\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\,\mathrm{d}m+o\left(\frac{n^2}{p}\right)
\end{aligned}$$

According to Cauchy's residue theorem, we have

$$\begin{aligned}
&-\frac{1}{2\pi i}\oint_{|m|=\rho}\sqrt{np}\cdot\frac{1}{m^3}\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\,\mathrm{d}m \\
&= \; -\frac{1}{2!}\,\mathrm{d}^{(2)}\left(\frac{\sqrt{np}X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\right)\Bigg/\mathrm{d}\,m^2\Bigg|_{m=0} = 0,
\end{aligned}$$

similarly,

$$\begin{aligned}
&-\frac{1}{2\pi i}\oint_{|m|=\rho}\left[\sqrt{np}\cdot\frac{1}{m^3}+\frac{n}{2}\cdot\frac{1+m^2}{m^4}+\frac{n}{3}\cdot\sqrt{\frac{n}{p}}\cdot\frac{1+2m^2}{m^5}\right. \\
&\quad\left.+\frac{n^2}{4p}\cdot\frac{1+3m^2+2m^4}{m^6}\right]\cdot\frac{X}{\left(1-\sqrt{\frac{n}{p}}m\right)}\cdot\frac{1}{m-\sqrt{\frac{n}{p}}}\,\mathrm{d}m+o\left(\frac{n^2}{p}\right) \\
&= \; -\frac{\nu_4-2}{2}+\frac{n^2}{3p}+o\left(\frac{n^2}{p}\right).
\end{aligned}$$

□

Proof of Lemma 2.1

Proof. According to Theorem 2.1, define function $f_1(x) = x^2$, then

$$\begin{aligned} G_n(f_1) &= n \int_{-\infty}^{+\infty} f_1(x) d\left(F^A(x) - F(x)\right) \\ &= \sum_{i=1}^{n} \lambda_i^2 - n \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{2\pi} \sqrt{4 - x^2}\, \mathrm{d}x \\ &= \sum_{i=1}^{n} \lambda_i^2 - n, \end{aligned}$$

where F^A is ESD of $A = \frac{1}{\sqrt{np}}(X'X - pI_n)$ and F represents the semicircular law. The mean correction term for $f_1(x) = x^2$ is, according to Lemma A.1,

$$\frac{n}{2\pi i} \oint_{|m|=\rho} f_1\left(-m - m^{-1}\right) \chi_n^{\text{Calib}}(m) \frac{1 - m^2}{m^2} \mathrm{d}m = \nu_4 - 2,$$

As for the mean function and covariance function of the Gaussian limit $Y(f_1)$, since

$$\Phi_1(f_1) = \frac{1}{2\pi} \int_{-\pi}^{\pi} 4 \cos^3 \theta \, \mathrm{d}\theta = 0,$$

$$\Phi_2(f_1) = \frac{1}{2\pi} \int_{-\pi}^{\pi} 4 \cos^2 \theta \cos 2\theta \, \mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos 4\theta + 1 + 2 \cos 2\theta) \, \mathrm{d}\theta = 1,$$

$$\begin{aligned} \Phi_k(f_1) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} 4 \cos^2 \theta \cos k\theta \, \mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} 2(\cos 2\theta + 1) \cos k\theta \, \mathrm{d}\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos(k-2)\theta + \cos(k+2)\theta + 2 \cos k\theta) \, \mathrm{d}\theta = 0, \quad \text{for } k \geq 3, \end{aligned}$$

therefore $Var(Y(f_1)) = 4$, in addition, $\mathbb{E}(Y(f_1)) = 0$, Conclusively, we have, when $p/n \to \infty$, $n \to \infty$,

$$\sum_{i=1}^{n} \lambda_i^2 - n - (\nu_4 - 2) \xrightarrow{d} N(0, 4).$$

Similarly, if we define function $f_2 = x$, then

$$\begin{aligned} G_n(f_2) &= n \int_{-\infty}^{+\infty} f_2(x) d\left(F^A(x) - F(x)\right) \\ &= \sum_{i=1}^{n} \lambda_i - n \int_{-\infty}^{+\infty} x \cdot \frac{1}{2\pi} \sqrt{4 - x^2} dx = \sum_{i=1}^{n} \lambda_i. \end{aligned}$$

The mean correction term for $f_2(x) = x$ is, according to Lemma A.1,

$$\frac{n}{2\pi i} \oint_{|m|=\rho} f_2\left(-m - m^{-1}\right) \chi_n^{\text{Calib}}(m) \frac{1 - m^2}{m^2} \mathrm{d}m = 0,$$

As for the mean function and covariance function of the Gaussian limit $Y(f_2)$, since

$$\Phi_0(f_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} 2\cos\theta \,\mathrm{d}\theta = 0,$$
$$\Phi_1(f_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} 2\cos^2\theta \,\mathrm{d}\theta = 1,$$
$$\Phi_2(f_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} 2\cos\theta\cos 2\theta \,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} (\cos 3\theta + \cos\theta) \,\mathrm{d}\theta = 0,$$
$$\begin{aligned}\Phi_k(f_2) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} 2\cos\theta\cos k\theta \,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} (\cos(k+1)\theta + \cos(k-1)\theta) \,\mathrm{d}\theta = 0 \text{ for } k \geq 3,\end{aligned}$$

therefore

$$\begin{aligned}Var(G_n(f_2)) &= (\nu_4 - 3)\Phi_1(f_2)\Phi_1(f_2) + 2\sum_{k=1}^{\infty} k\Phi_k(f_2)\Phi_k(f_2) \\ &= \nu_4 - 1,\end{aligned}$$

in addition, $\mathbb{E}(Y(f_2)) = 0$. In conclusion, we have, when $p/n \to \infty$, $n \to \infty$,

$$\sum_{i=1}^{n} \lambda_i^2 - n - (\nu_4 - 2) \xrightarrow{d} N(0, 4),$$
$$\sum_{i=1}^{n} \lambda_i \xrightarrow{d} N(0, \nu_4 - 1).$$

Now consider the covariance between $G_n(f_1)$ and $G_n(f_2)$, then

$$Cov(G_n(f_1), G_n(f_2)) = (\nu_4 - 3)\Phi_1(f_1)\Phi_1(f_2) + 2\sum_{k=1}^{\infty} k\Phi_k(f_1)\Phi_k(f_2) = 0.$$

Consequently, when $p/n \to \infty$, $n \to \infty$,

$$\begin{pmatrix} \sum_{i=1}^{n} \lambda_i^2 - n - (\nu_4 - 2) \\ \sum_{i=1}^{n} \lambda_i \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & \nu_4 - 1 \end{pmatrix} \right), \qquad \square$$

Proof of Lemma 2.2

Proof. According to Theorem 2.1, define function $f_3(x) = \frac{p}{n}\log\left(1 + \sqrt{\frac{n}{p}}x\right)$, then

$$G_n(f_3) = n\int_{-\infty}^{+\infty} f_3(x) \,\mathrm{d}\big(F^A(x) - F(x)\big)$$

$$= \frac{p}{n}\sum_{i=1}^{n}\log\left(1+\lambda_i\sqrt{\frac{n}{p}}\right) - n\int_{-2}^{2}\frac{p}{n}\log\left(1+\sqrt{\frac{n}{p}}x\right)\cdot F(x)\,\mathrm{d}x$$

where F^A is ESD of $A = \frac{1}{\sqrt{np}}(X'X - pI_n)$ and F represents the semicircular law.

$$\begin{aligned} n\int_{-2}^{2}\frac{p}{n}\log\left(1+\sqrt{\frac{n}{p}}x\right)\cdot F(x)\,\mathrm{d}x &= p\int_{-2}^{2}\log\left(1+\sqrt{\frac{n}{p}}x\right)\cdot\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x \\ &= -\frac{n}{2}\sum_{k=0}^{\infty}\frac{(2k+1)!!}{2^{k-1}(k+1)^2(k+2)}\cdot\left(\frac{4n}{p}\right)^k \\ &= -\frac{n}{2}-\frac{n^2}{2p}+o\left(\frac{n^2}{2p}\right) \end{aligned}$$

The mean correction term for $f_3(x) = \frac{p}{n}\log\left(1+\sqrt{\frac{n}{p}}x\right)$ is, according to Lemma A.1,

$$\frac{n}{2\pi i}\oint_{|m|=\rho} f_3\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\,\mathrm{d}m = -\frac{\nu_4-2}{2}+\frac{n^2}{3p},$$

As for the mean function and covariance function of the Gaussian limit $Y(f_3)$, since

$$\begin{aligned} \log(1+x) &= \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}, \\ \Phi_1(f_3) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_3(2\cos\theta)\cdot\cos\theta\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(1+\sqrt{\frac{n}{p}}\cdot 2\cos\theta\right)\cdot\cos\theta\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sqrt{\frac{p}{n}}\cdot 2\cos\theta\cdot\cos\theta\,\mathrm{d}\theta - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\cdot(2\cos\theta)^2\cdot\cos\theta\,\mathrm{d}\theta \\ &+ \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{3}\sqrt{\frac{n}{p}}\cdot(2\cos\theta)^3\cdot\cos\theta\,\mathrm{d}\theta + o\left(\sqrt{\frac{n}{p}}\right) \\ &= \sqrt{\frac{p}{n}}+\sqrt{\frac{n}{p}}+o\left(\sqrt{\frac{n}{p}}\right), \end{aligned}$$

for $k \geq 2$,

$$\begin{aligned} \Phi_k(f_3) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_3(2\cos\theta)\cos k\theta\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sqrt{\frac{p}{n}}\cdot 2\cos\theta\cdot\cos k\theta\,\mathrm{d}\theta - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\cdot(2\cos\theta)^2\cdot\cos k\theta\,\mathrm{d}\theta \end{aligned}$$

$$+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{3}\sqrt{\frac{n}{p}}\cdot(2\cos\theta)^3\cdot\cos k\theta\,\mathrm{d}\theta+o\left(\sqrt{\frac{n}{p}}\right)$$
$$=o\left(\sqrt{\frac{n}{p}}\right)+\begin{cases}-\frac{1}{2} & k=2\\ \frac{1}{3}\sqrt{\frac{n}{p}} & k=3\\ 0 & k\geq 4\end{cases}$$

therefore

$$\begin{aligned}Var(G_n(f_3))&=(\nu_4-3)\Phi_1(f_3)\Phi_1(f_3)+2\sum_{k=1}^{\infty}k\Phi_k(f_3)\Phi_k(f_3)\\&=(\nu_4-1)\left(\sqrt{\frac{p}{n}}+\sqrt{\frac{n}{p}}\right)^2+2\cdot 2\cdot\left(-\frac{1}{2}\right)^2+2\cdot 3\cdot\frac{1}{9}\cdot\frac{n}{p}\\&=(\nu_4-1)\cdot\frac{p}{n}+2\nu_4-1+\frac{n}{p}(\nu_4-\frac{1}{3}),\end{aligned}$$

in addition, $\mathbb{E}(Y(f_3))=0$, Conclusively, we have, when $p/n\to\infty$, $n\to\infty$,

$$\begin{aligned}&\frac{p}{n}\sum_{i=1}^{n}\log\left(1+\sqrt{\frac{n}{p}}\lambda_i\right)+\frac{n}{2}+\frac{n^2}{6p}+\frac{\nu_4-2}{2}\\&\quad\xrightarrow{d}N\left(0,\frac{p}{n}(\nu_4-1)+2\nu_4-1+\frac{n}{p}\left(\nu_4-\frac{1}{3}\right)\right).\end{aligned}$$

If we define function $f_2=x$, it has been proved in Lemma 2.1 that,

$$G_n(f_2)=n\int_{-\infty}^{+\infty}f_2(x)d\left(F^A(x)-F(x)\right)=\sum_{i=1}^{n}\lambda_i.$$

The mean correction term for $f_2(x)=x$ is, according to Lemma A.1,

$$\frac{n}{2\pi i}\oint_{|m|=\rho}f_2\left(-m-m^{-1}\right)\chi_n^{\mathrm{Calib}}(m)\frac{1-m^2}{m^2}\mathrm{d}m=0,$$

As for the mean function and covariance function of the Gaussian limit $Y(f_2)$,

$$\begin{aligned}\Phi_0(f_2)&=\frac{1}{2\pi}\int_{-\pi}^{\pi}2\cos\theta\,\mathrm{d}\theta=0,\\\Phi_1(f_2)&=\frac{1}{2\pi}\int_{-\pi}^{\pi}2\cos^2\theta\,\mathrm{d}\theta=1,\\\Phi_2(f_2)&=\frac{1}{2\pi}\int_{-\pi}^{\pi}2\cos\theta\cos 2\theta\,\mathrm{d}\theta=0,\\\Phi_k(f_2)&=\frac{1}{2\pi}\int_{-\pi}^{\pi}2\cos\theta\cos k\theta\,\mathrm{d}\theta=0\text{ for }k\geq 3,\end{aligned}$$

$$Var(G_n(f_2)) = (\nu_4 - 3)\Phi_1(f_2)\Phi_1(f_2) + 2\sum_{k=1}^{\infty} k\Phi_k(f_2)\Phi_k(f_2) = \nu_4 - 1,$$

in addition, $\mathbb{E}(Y(f_2)) = 0$. In conclusion, we have, when $p/n \to \infty$, $n \to \infty$,

$$\sum_{i=1}^{n} \lambda_i \xrightarrow{d} N(0, \nu_4 - 1).$$

Now consider the covariance between $G_n(f_3)$ and $G_n(f_2)$, then

$$\begin{aligned} Cov(G_n(f_3), G_n(f_2)) &= (\nu_4 - 3)\Phi_1(f_3)\Phi_1(f_2) + 2\sum_{k=1}^{\infty} k\Phi_k(f_3)\Phi_k(f_2) \\ &= (\nu_4 - 1)\left(\sqrt{\frac{p}{n}} + \sqrt{\frac{n}{p}}\right). \end{aligned}$$

Consequently result follows. □

Proof of Lemma 3.1

Proof. According to Theorem 3.1, define function $f_1(x) = x^2$, then

$$\begin{aligned} G_n(f_1) &= n\int_{-\infty}^{+\infty} f_1(x)\, d\left(F^{\widetilde{A}}(x) - F(x)\right) - \sqrt{\frac{n^3}{p}}\Phi_3(f_1) \\ &= \sum_{i=1}^{n} \widetilde{\lambda}_i^2 - n\int_{-2}^{2} \frac{x^2}{2\pi}\sqrt{4 - x^2}\, \mathrm{d}x - \sqrt{\frac{n^3}{p}}\Phi_3(f_1) \\ &= \sum_{i=1}^{n} \widetilde{\lambda}_i^2 - n - \sqrt{\frac{n^3}{p}}\Phi_3(f_1), \end{aligned}$$

where $F^{\widetilde{A}}$ is the ESD of $\widetilde{A} = \sqrt{\frac{1}{n}}\left(\frac{1}{\sqrt{tr(\Sigma_p^2)}}Z'\Sigma_p Z - \frac{tr(\Sigma_p)}{\sqrt{tr(\Sigma_p^2)}}I_n\right)$ and F represents the semicircular law.

As for the mean function and covariance function of the Gaussian limit $Y(f_1)$, since

$$\begin{aligned} \Phi_0(f_1) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(2\cos(\theta))\, \mathrm{d}\theta = 2, \\ \Phi_k(f_1) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(2\cos(\theta))\cos(k\theta)\, \mathrm{d}\theta = \begin{cases} 0, & k = 1, \\ 1, & k = 2, \\ 0, & k \geq 3, \end{cases} \\ G_n(f_1) &= \sum_{i=1}^{n} \widetilde{\lambda}_i^2 - n, \end{aligned}$$

$$
\begin{aligned}
\mathbb{E}\left(Y\left(f_1\right)\right) &= \frac{1}{4}\left(f_1\left(2\right)+f_1\left(-2\right)\right)-\frac{1}{2}\Phi_0\left(f_1\right)+\frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_2\left(f_1\right)\\
&= 2-1+\frac{\omega}{\theta}\left(\nu_4-3\right)=\frac{\omega}{\theta}\left(\nu_4-3\right)+1,\\
Var\left(Y\left(f_1\right)\right) &= \frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_1^2\left(f_1\right)+2\sum_{k=1}^{\infty}k\Phi_k^2\left(f_1\right)=4.
\end{aligned}
$$

Similarly, if we define function $f_2\left(x\right)=x$, then

$$
\begin{aligned}
G_n\left(f_2\right) &= n\int_{-\infty}^{+\infty} f_2\left(x\right)d\left(F^{\widetilde{A}}\left(x\right)-F\left(x\right)\right)-\sqrt{\frac{n^3}{p}}\Phi_3\left(f_2\right)\\
&= \sum_{i=1}^{n}\widetilde{\lambda}_i-n\int_{-2}^{2}\frac{x}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x-\sqrt{\frac{n^3}{p}}\Phi_3\left(f_2\right)\\
&= \sum_{i=1}^{n}\widetilde{\lambda}_i-\sqrt{\frac{n^3}{p}}\Phi_3\left(f_2\right),
\end{aligned}
$$

As for the mean function and covariance function of the Gaussian limit $Y\left(f_2\right)$, since

$$
\begin{aligned}
\Phi_0\left(f_2\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_2\left(2\cos\left(\theta\right)\right)\mathrm{d}\theta=0,\\
\Phi_k\left(f_2\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_2\left(2\cos\left(\theta\right)\right)\cos\left(k\theta\right)\mathrm{d}\theta=\begin{cases}1, & k=1,\\ 0, & k\geq 2,\end{cases}\\
G_n\left(f_2\right) &= \sum_{i=1}^{n}\widetilde{\lambda}_i,\\
\mathbb{E}\left(Y\left(f_2\right)\right) &= \frac{1}{4}\left(f_2\left(2\right)+f_2\left(-2\right)\right)-\frac{1}{2}\Phi_0\left(f_2\right)+\frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_2\left(f_2\right)\\
&= -\frac{1}{2}\Phi_0\left(f_2\right)+\frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_2\left(f_2\right)=0,\\
Var\left(Y\left(f_2\right)\right) &= \frac{\omega}{\theta}\left(\nu_4-3\right)+2,\\
Cov\left(Y\left(f_1\right),Y\left(f_2\right)\right) &= \frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_1\left(f_1\right)\Phi_1\left(f_2\right)+2\sum_{k=1}^{\infty}k\Phi_k\left(f_1\right)\Phi_k\left(f_2\right)=0,
\end{aligned}
$$

therefore

$$
\begin{pmatrix}\sum_{i=1}^{n}\widetilde{\lambda}_i^2-n-\left(\frac{\omega}{\theta}\left(\nu_4-3\right)+1\right)\\ \sum_{i=1}^{n}\widetilde{\lambda}_i\end{pmatrix}\xrightarrow{d}N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}4 & 0\\ 0 & \frac{\omega}{\theta}\left(\nu_4-3\right)+2\end{pmatrix}\right),
$$

as $n\to\infty,\quad p\to\infty,\quad p/n^3=O\left(1\right).$ □

Proof of Lemma 3.2

Proof. According to Theorem 3.1, define function $f_3(x) = \frac{p}{n}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}x\right)$, then

$$
\begin{aligned}
G_n(f_3) &= n\int_{-\infty}^{+\infty} f_3(x)\,d\left(F^{\widetilde{A}}(x) - F(x)\right) - \sqrt{\frac{n^3}{p}}\Phi_3(f_3) \\
&= \sum_{i=1}^{n}\frac{p}{n}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}\widetilde{\lambda}_i\right) - n\int_{-2}^{2}\frac{p}{n}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}x\right)\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x \\
&\quad - \sqrt{\frac{n^3}{p}}\Phi_3(f_3),
\end{aligned}
$$

$$
\begin{aligned}
& n\int_{-2}^{2}\frac{p}{n}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}x\right)\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x \\
&\quad = p\int_{-2}^{2}\log\left(\gamma + \sqrt{\frac{n\theta}{p}}x\right)\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x \\
&\quad = p\int_{-2}^{2}\left(\log\gamma + \sum_{k=1}^{\infty}(-1)^{k+1}\frac{\left(\frac{\sqrt{\theta}}{\gamma}\sqrt{\frac{n}{p}}x\right)^k}{k}\right)\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x \\
&\quad = p\log\gamma - \frac{\theta}{2\gamma^2}n - \frac{\theta^2}{2\gamma^4}\frac{n^2}{p} + o\left(\frac{n^2}{p}\right),
\end{aligned}
$$

$$
\begin{aligned}
\Phi_0(f_3) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(\gamma + 2\sqrt{\frac{n\theta}{p}}\cos t\right)\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\gamma\,\mathrm{d}t + \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(1 + \frac{2\sqrt{\theta}}{\gamma}\sqrt{\frac{n}{p}}\cos t\right)\mathrm{d}t \\
&= \frac{p}{n}\log\gamma + \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}}\cos t\,\mathrm{d}t - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2\theta}{\gamma^2}(\cos t)^2\,\mathrm{d}t \\
&+ \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{8\theta\sqrt{\theta}}{3\gamma^3}(\cos t)^3\sqrt{\frac{n}{p}}\,\mathrm{d}t + o\left(\sqrt{\frac{n}{p}}\right) \\
&= \frac{p}{n}\log\gamma - \frac{\theta}{\gamma^2} + o\left(\sqrt{\frac{n}{p}}\right), \\
\Phi_1(f_3) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(\gamma + 2\sqrt{\frac{n\theta}{p}}\cos t\right)\cos t\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\gamma\cos t\,\mathrm{d}t + \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(1 + \frac{2\sqrt{\theta}}{\gamma}\sqrt{\frac{n}{p}}\cos t\right)\cos t\,\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}}\left(\cos t\right)^2\mathrm{d}t-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2\theta}{\gamma^2}\left(\cos t\right)^3\mathrm{d}t\\
&+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{8\theta\sqrt{\theta}}{3\gamma^3}\left(\cos t\right)^4\sqrt{\frac{n}{p}}\,\mathrm{d}t+o\left(\sqrt{\frac{n}{p}}\right)\\
&= \frac{\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}}+\frac{\theta\sqrt{\theta}}{\gamma^3}\sqrt{\frac{n}{p}}+o\left(\sqrt{\frac{n}{p}}\right),
\end{aligned}
$$

$$
\begin{aligned}
\Phi_k\left(f_3\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(\gamma+2\sqrt{\frac{n\theta}{p}}\cos t\right)\cos kt\,\mathrm{d}t\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\gamma\cos kt\,\mathrm{d}t+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{p}{n}\log\left(1+\frac{2\sqrt{\theta}}{\gamma}\sqrt{\frac{n}{p}}\cos t\right)\cos kt\,\mathrm{d}t\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}}\cos t\cos kt\,\mathrm{d}t-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2\theta}{\gamma^2}\left(\cos t\right)^2\cos kt\,\mathrm{d}t\\
&+\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{8\theta\sqrt{\theta}}{3\gamma^3}\sqrt{\frac{n}{p}}\left(\cos t\right)^3\cos kt\,\mathrm{d}t+o\left(\sqrt{\frac{n}{p}}\right)\\
&= \begin{cases}-\frac{\theta}{2\gamma^2}, & k=2,\\ \frac{\theta\sqrt{\theta}}{3\gamma^3}\sqrt{\frac{n}{p}}, & k=3,\\ 0, & k\geq 4,\end{cases}
\end{aligned}
$$

thus,

$$
\begin{aligned}
G_n\left(f_3\right) &= \sum_{i=1}^{n}\frac{p}{n}\log\left(\gamma+\sqrt{\frac{n\theta}{p}}\widetilde{\lambda}_i\right)\\
&\quad -n\int_{-2}^{2}\frac{p}{n}\log\left(\gamma+\sqrt{\frac{n\theta}{p}}x\right)\frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x-\sqrt{\frac{n^3}{p}}\Phi_3\left(f_3\right)\\
&= \sum_{i=1}^{n}\frac{p}{n}\log\left(\gamma+\sqrt{\frac{n\theta}{p}}\widetilde{\lambda}_i\right)-\left(p\log\gamma-\frac{\theta}{2\gamma^2}n-\frac{\theta^2}{2\gamma^4}\frac{n^2}{p}\right)-\frac{\theta\sqrt{\theta}}{3\gamma^3}\frac{n^2}{p},\\
\mathbb{E}\left(Y\left(f_3\right)\right) &= \frac{1}{4}\left(f_3\left(2\right)+f_3\left(-2\right)\right)-\frac{1}{2}\Phi_0\left(f_3\right)+\frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_2\left(f_3\right)\\
&= \frac{1}{4}\left(\frac{p}{n}\log\left(\gamma+2\sqrt{\frac{n\theta}{p}}\right)+\frac{p}{n}\log\left(\gamma-2\sqrt{\frac{n\theta}{p}}\right)\right)\\
&\quad -\frac{1}{2}\left(\frac{p}{n}\log\gamma-\frac{\theta}{\gamma^2}\right)-\frac{\omega}{\theta}\left(\nu_4-3\right)\frac{\theta}{2\gamma^2}\\
&= -\frac{\theta}{2\gamma^2}-\frac{\omega}{2\gamma^2}\left(\nu_4-3\right),\\
Var\left(Y\left(f_3\right)\right) &= \frac{\omega}{\theta}\left(\nu_4-3\right)\Phi_1^2\left(f_3\right)+2\sum_{k=1}^{\infty}k\Phi_k^2\left(f_3\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\omega}{\theta}(\nu_4 - 3)\left(\frac{\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}} + \frac{\theta\sqrt{\theta}}{\gamma^3}\sqrt{\frac{n}{p}}\right)^2 \\
&+ 2\left(\frac{\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}} + \frac{\theta\sqrt{\theta}}{\gamma^3}\sqrt{\frac{n}{p}}\right)^2 + 4\left(\frac{\theta^2}{4\gamma^4}\right) + 6\frac{\theta^3}{9\gamma^6}\frac{n}{p} \\
&= \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\frac{\theta}{\gamma^2}\frac{p}{n} + \left(\frac{2\omega}{\theta}(\nu_4 - 3) + 5\right)\frac{\theta^2}{\gamma^4} \\
&\quad + \left(\frac{\omega}{\theta}(\nu_4 - 3) + \frac{8}{3}\right)\frac{\theta^3}{\gamma^6}\frac{n}{p}.
\end{aligned}
$$

Consider function $f_2(x) = x$, from lemma 3.1, we have

$$
\Phi_k(f_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_2(2\cos(\theta))\cos(k\theta)\,\mathrm{d}\theta = \begin{cases} 0, & k = 0 \\ 1, & k = 1 \\ 0, & k \geq 2 \end{cases},
$$

therefore the covariance between $Y(f_2)$ and $Y(f_3)$ is

$$
\begin{aligned}
Cov(Y(f_2), Y(f_3)) &= \frac{\omega}{\theta}(\nu_4 - 3)\Phi_1(f_2)\Phi_1(f_3) + 2\sum_{k=1}^{\infty} k\Phi_k(f_2)\Phi_k(f_3) \\
&= \left(\frac{\omega}{\theta}(\nu_4 - 3) + 2\right)\left(\frac{\sqrt{\theta}}{\gamma}\sqrt{\frac{p}{n}} + \frac{\theta\sqrt{\theta}}{\gamma^3}\sqrt{\frac{n}{p}}\right),
\end{aligned}
$$

consequently the result follows. □

References

[1] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd edition. Wiley, New York. MR0771294

[2] Bai, Z. D., and Yao, J. (2005). On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli*, 11.6: 1059–1092. MR2189081

[3] Bai, Z. D., and Yong Q. Yin. (1988). Convergence to the semicircle law. *The Annals of Probability*, 863–875. MR0929083

[4] Chen B. B., Pan G. M. (2015). CLT for linear spectral statistics of normalized sample covariance matrices with the dimension much larger than the sample size. *Bernoulli*, 21(2): 1089–1133. MR3338658

[5] Chen S. X., Zhang L. X., Zhong P. S. (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association*, 105(490). MR2724863

[6] Fisher, Thomas J., Xiaoqian Sun, and Colin M. Gallagher (2010). A new test for sphericity of the covariance matrix for high dimensional data. *Journal of Multivariate Analysis*, 101.10: 2554–2570. MR2719881

[7] John S. (1971). Some optimal multivariate tests. *Biometrika*, 58(1): 123–127. MR0275568

[8] John S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. *Biometrika*, 59(1): 169–173. MR0312619

[9] Ledoit O., Wolf M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of Statistics*, 1081–1102. MR1926169

[10] Li Z. and Yao J. (2016). CLT for linear spectral statistics of general sample covariance matrices with dimension much larger than sample size. *Preprint*.

[11] Onatski, A., Moreira, M. J., and Hallin, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *The Annals of Statistics*, 41(3): 1204–1231. MR3113808

[12] Schott J. R. (2005). Testing for complete independence in high dimensions. *Biometrika*, 92(4): 951–956. MR2234197

[13] M. S. Srivastava (2005). Some tests concerning the covariance matrix in high-dimensional data, *Journal of The Japan Statistical Society*, 35, 251–272. MR2328427

[14] M. S. Srivastava (2006). Some tests criteria for the covariance matrix with fewer observations than the dimension, *Acta Comment. Univ. Tartu. Math*, 10: 77–93. MR2309747

[15] M. S. Srivastava, Muni S., Tõnu Kollo, and Dietrich von Rosen (2011). Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis*, 102.6: 1090–1103. MR2793878

[16] Wang Q. and Yao J. (2013). On the sphericity test with large-dimensional observations. *Electronic Journal of Statistics*, 7: 2164–2192. MR3104916

[17] Yao J., Bai Z., and Zheng S. (2015). *Large Sample Covariance Matrices and High-dimensional Data Analysis*. Cambridge University Press. Vol. 39. MR3468554

[18] Zheng, S., Bai, Z., and Yao, J. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *The Annals of Statistics*, 43(2): 546–591. MR3316190

[19] Zou C., Peng L., Feng L., et al. (2013). Multivariate sign-based high-dimensional tests for sphericity. *Biometrika*, ast040.