



Testing high-dimensional covariance matrices under the elliptical distribution and beyond

Xinxin Yang^a, Xinghua Zheng^{b,*}, Jiaqi Chen^c

^a School of Statistics and Mathematics, Central University of Finance and Economics, China

^b Department of ISOM, Hong Kong University of Science and Technology, Hong Kong

^c School of Mathematics, Harbin Institute of Technology, China

ARTICLE INFO

Article history:

Received 8 December 2018

Received in revised form 14 December 2019

Accepted 17 May 2020

Available online 10 August 2020

JEL classification:

C12

C55

C58

Keywords:

Covariance matrix

High-dimension

Elliptical model

Linear spectral statistics

Central limit theorem

ABSTRACT

We develop tests for high-dimensional covariance matrices under a generalized elliptical model. Our tests are based on a central limit theorem for linear spectral statistics of the sample covariance matrix based on self-normalized observations. For testing sphericity, our tests neither assume specific parametric distributions nor involve the kurtosis of data. More generally, we can test against any non-negative definite matrix that can even be not invertible. As an interesting application, we illustrate in empirical studies that our tests can be used to test uncorrelatedness among idiosyncratic returns.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Tests for high-dimensional covariance matrices

Testing covariance matrices is of fundamental importance in multivariate analysis. There has been a long history of study on testing (i) the covariance matrix Σ is equal to a given matrix, or (ii) the covariance matrix Σ is proportional to a given matrix. Specifically, for a given non-negative definite matrix Σ_0 , one aims to test

$$H_0 : \Sigma = \Sigma_0 \quad \text{vs.} \quad H_a : \Sigma \neq \Sigma_0, \quad (1)$$

or

$$H_0 : \Sigma \propto \Sigma_0 \quad \text{vs.} \quad H_a : \Sigma \not\propto \Sigma_0, \quad (2)$$

where “ $\propto$ ” stands for “proportional to”. When $\Sigma_0 = \mathbf{I}$, test (1) is referred to as the identity test and (2) as the sphericity test. If Σ_0 is invertible, then testing (1) or (2) can be reduced to the identity or sphericity test by multiplying the observations with $\Sigma_0^{-1/2}$.

* Correspondence to: Department of Information Systems, Business Statistics and Operations Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.

E-mail address: xhzheng@ust.hk (X. Zheng).

In the classical setting where the dimension p is fixed and the sample size n goes to infinity, the sample covariance matrix is a consistent estimator, and further inference can be made based on the associated *central limit theory* (CLT). Examples include the likelihood ratio tests (see, e.g., [Muirhead \(1982\)](#), Sections 8.3 and 8.4), and the locally most powerful invariant tests ([John, 1971](#); [Nagao, 1973](#)).

In the high-dimensional setting, because the sample covariance matrix is inconsistent, conventional tests may not apply. For the identity and sphericity tests, new methods have been developed, first under the multivariate normal distribution, then for more generally distributed data:

- Multivariate normally distributed data. When $p/n \rightarrow y \in (0, \infty)$, [Ledoit and Wolf \(2002\)](#) show that John's test of sphericity is still consistent and propose a modified Nagao's identity test. [Srivastava \(2005\)](#) introduces a new test of sphericity under a more general condition that $n = O(p^\delta)$ for some $\delta \in (0, 1]$. [Birke and Dette \(2005\)](#) show that the asymptotic null distributions of John's and the modified Nagao's test statistics in [Ledoit and Wolf \(2002\)](#) are still valid when $p/n \rightarrow \infty$. Relaxing the normality assumption but still assuming the kurtosis equals 3, [Bai et al. \(2009\)](#) develop a corrected likelihood ratio test of identity when $p/n \rightarrow y \in (0, 1)$. For testing sphericity, [Jiang and Yang \(2013\)](#) derive the asymptotic distribution of the likelihood ratio test statistic under the multivariate normal distribution with $p/n \rightarrow y \in (0, 1]$.
- More generally distributed data. [Chen et al. \(2010\)](#) generalize the results in [Ledoit and Wolf \(2002\)](#) without assuming normality nor an explicit relationship between p and n . By relaxing the kurtosis assumption, [Wang et al. \(2013\)](#) extend the corrected likelihood ratio test in [Bai et al. \(2009\)](#) and the modified Nagao's test in [Ledoit and Wolf \(2002\)](#). Along this line, [Wang and Yao \(2013\)](#) propose two tests by correcting the likelihood ratio test and John's test.

1.2. The elliptical distribution and its applications

The elliptically distributed data can be expressed as

$$\mathbf{Y} = \omega \Sigma^{1/2} \mathbf{Z},$$

where ω is a positive random scalar, $\mathbf{Z}$ is a p -dimensional random vector from $N(\mathbf{0}, \mathbf{I})$, and further ω and $\mathbf{Z}$ are independent of each other. It is a natural generalization of the multivariate normal distribution, and contains many widely used distributions as special cases including the multivariate t -distribution, the symmetric multivariate Laplace distribution and the symmetric multivariate stable distribution. See [Fang et al. \(1990\)](#) for further details.

One of our motivations of this study arises from the wide applicability of the elliptical distribution. The “mixture coefficient” ω can feature heteroskedasticity that are widely present in real data. Furthermore, just like the ARCH ([Engle, 1982](#)) and GARCH ([Bollerslev, 1986](#)) models, the marginal distribution of $\mathbf{Y}$ is a mixture of normal hence heavy-tailed. Therefore, the elliptical distribution can feature both heteroskedasticity and heavy-tailedness. In finance, stock returns have been extensively documented to exhibit such two features, dating back at least to [Fama \(1965\)](#) and [Mandelbrot \(1967\)](#). Accommodating heteroskedasticity and heavy-tailedness makes the elliptical distribution a more admissible candidate for stock-return models than the Gaussian distribution; see, e.g., [Owen and Rabinovitch \(1983\)](#) and [Bingham and Kiesel \(2002\)](#). [McNeil et al. \(2005\)](#) state that “elliptical distributions ... provided far superior models to the multivariate normal for daily and weekly US stock-return data” and that “multivariate return data for groups of returns of similar type often look roughly elliptical”.

1.3. Performance of existing tests under the elliptical model

Given the wide applicability of the elliptical distribution, it is important to check whether existing tests for covariance matrices are applicable to the elliptical distribution. Both numerical and theoretical analyses give a negative answer.

We start with a simple numerical study to investigate the empirical sizes. Consider the observations $\mathbf{Y}_i = \omega_i \mathbf{Z}_i$, $i = 1, \dots, n$, where

- ω_i 's are absolute values of i.i.d. standard normal random variables,
- $\mathbf{Z}_i$'s are i.i.d. p -dimensional standard multivariate normal random vectors, and
- ω_i 's and $\mathbf{Z}_i$'s are independent of each other.

Under such a setting, $\mathbf{Y}_i$'s are i.i.d. random vectors with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}$. We will test both $H_0 : \Sigma = \mathbf{I}$ and $H_0 : \Sigma \propto \mathbf{I}$.

To test $H_0 : \Sigma = \mathbf{I}$, we use the tests in [Ledoit and Wolf \(2002\)](#) (LW_1 test), [Bai et al. \(2009\)](#) (BJYZ test), [Chen et al. \(2010\)](#) (CZZ₁ test) and [Wang et al. \(2013\)](#) (WYMC-LR and WYMC-LW tests). For testing $H_0 : \Sigma \propto \mathbf{I}$, we apply the tests proposed in [Ledoit and Wolf \(2002\)](#) (LW_2 test), [Srivastava \(2005\)](#) (S test), [Chen et al. \(2010\)](#) (CZZ₂ test) and [Wang and Yao \(2013\)](#) (WY-LR and WY-JHN tests). [Table 1](#) reports the empirical sizes for these tests at 5% significance level.

We observe from [Table 1](#) that the empirical sizes of all these tests are far higher than the nominal level of 5%, suggesting that they are inconsistent under the elliptical distribution.

It is worth discussing why these tests fail. Firstly, the reason is not that these tests do not apply to the high-dimensional setting. In fact, when $\omega_i \equiv 1$, for the same pairs of dimensions and sample sizes, all these tests yield sizes close to the

Table 1

Empirical sizes (%) of the existing tests for testing $H_0: \Sigma = \mathbf{I}$ or $H_0: \Sigma \propto \mathbf{I}$ at 5% significance level. Data are generated as $\mathbf{Y}_i = \omega_i \mathbf{Z}_i$ where ω_i 's are absolute values of i.i.d. $N(0, 1)$, $\mathbf{Z}_i$'s are i.i.d. $N(\mathbf{0}, \mathbf{I})$, and further ω_i 's and $\mathbf{Z}_i$'s are independent of each other. The results are based on 10,000 replications for each pair of p and n .

$H_0: \Sigma = \mathbf{I}$									
p	$p/n = 0.5$					$p/n = 2$			
	LW ₁	BJYZ	CZZ ₁	WYMC-LR	WYMC-LW	LW ₁	CZZ ₁	WYMC-LW	
100	100	100	54.0	100	100	100	50.2	100	
200	100	100	51.6	100	100	100	53.0	100	
500	100	100	52.3	100	100	100	53.3	100	
$H_0: \Sigma \propto \mathbf{I}$									
p	$p/n = 0.5$					$p/n = 2$			
	LW ₂	S	CZZ ₂	WY-LR	WY-JHN	LW ₂	S	CZZ ₂	WY-JHN
100	100	100	51.8	100	100	100	100	50.2	100
200	100	100	53.0	100	100	100	100	52.3	100
500	100	100	52.3	100	100	100	100	53.5	100

nominal level of 5%; see Table 2 for details. Secondly, it is not due to heavy-tailedness either. The marginal distribution of $\mathbf{Y}$, although is heavier than normal, still has exponentially decaying tails with a finite moment generating function.

The real reason that these tests fail lies in the presence of $\{\omega_i\}$. Denote $\mathbf{S}_n = n^{-1} \sum_{i=1}^n \mathbf{Y}_i \mathbf{Y}_i^\top = n^{-1} \sum_{i=1}^n \omega_i^2 \mathbf{Z}_i \mathbf{Z}_i^\top$. The celebrated Marčenko–Pastur theorem states that, when ω_i 's are constant, the empirical spectral distribution (ESD) of $\mathbf{S}_n$ converges to the Marčenko–Pastur law. This convergence turns out to be crucial for all the aforementioned tests as they all involve certain moments of the limiting ESD (LSD). When ω_i 's are not constant, Theorem 1 of Zheng and Li (2011) implies that the ESD of $\mathbf{S}_n$ will not converge to the Marčenko–Pastur law. Consequently, the asymptotic null distributions of the aforementioned test statistics change, and the tests no longer apply.

1.4. Our model and aim of this study

The previous section shows that existing tests do not apply when observations are heteroskedastic, a feature that is commonly encountered in finance, economics and many other fields.

In this paper, we study tests for high-dimensional covariance matrices when data may exhibit heteroskedasticity. Specifically, we consider the following model. Denote by $\mathbf{Y}_i$, $i = 1, \dots, n$, the observations, which can be written as

$$\mathbf{Y}_i = \omega_i \Sigma^{1/2} \mathbf{Z}_i, \quad (3)$$

where

- (i) ω_i 's are positive random scalars reflecting heteroskedasticity,
- (ii) $\Sigma \in \mathbb{R}^{p \times p}$ is a non-negative definite matrix,
- (iii) $\mathbf{Z} := (\mathbf{Z}_1, \dots, \mathbf{Z}_n) = (\mathbf{Z}_{ij})_{p \times n}$ consists of i.i.d. standardized random variables,
- (iv) ω_i 's can depend on each other and on $\{\mathbf{Z}_i : i = 1, \dots, n\}$ in an arbitrary way, and
- (v) ω_i 's do not need to be stationary.

Model (3) incorporates the elliptical distribution as a special case. This general model further possesses several important advantages:

- It can be considered as a multivariate extension of the ARCH/GARCH model and accommodates conditional heteroskedasticity. In the ARCH/GARCH model, the volatility process is serially dependent and depends on past information. Such dependence is excluded from the elliptical distribution; however, it is perfectly compatible with Model (3).
- The dependence between $\{\omega_i\}$ and $\{\mathbf{Z}_i\}$ can feature the leverage effect in financial econometrics, which accounts for the negative correlation between asset return and change in volatility. Various research has been conducted to study the leverage effect; see, e.g., Schwert (1989), Campbell and Hentschel (1992), Aït-Sahalia et al. (2013), Wang and Mykland (2014) and Kalnina and Xiu (2017).
- Furthermore, it can capture (conditional) asymmetry by allowing the entries of $\mathbf{Z}_i$'s to be asymmetrically distributed. The asymmetry is another stylized fact of financial data. For instance, the empirical study in Singleton and Wingender (1986) shows high skewness in individual stock returns. Skewness is also reported in exchange rate returns in Peiro (1999). Christoffersen (2012) documents that asymmetry exists in standardized returns; see Chapter 6 therein.

Because of the heteroskedasticity induced by $\{\omega_i\}$, in this paper we focus on testing

$$[H_0: \Sigma \propto \Sigma_0 \text{ vs. } H_a: \Sigma \not\propto \Sigma_0,]$$

in the high-dimensional setting where both p and n grow to infinity with the ratio $p/n \rightarrow y \in (0, \infty)$.

Note that in many applications, knowing the covariance matrix up to a constant is good enough. For example, the minimum variance portfolio is given by $\Sigma^{-1}\mathbf{1}/(\mathbf{1}^T\Sigma^{-1}\mathbf{1})$, where $\mathbf{1} = (1, \dots, 1)^T$. The portfolio is therefore invariant to scaling in the covariance matrix.

1.5. Summary of main results

To deal with heteroskedasticity, we propose to self-normalize the observations. To be specific, we focus on the self-normalized observations $\mathbf{Y}_i/|\mathbf{Y}_i|$, where $|\cdot|$ stands for the Euclidean norm. Observe that

$$\frac{\mathbf{Y}_i}{|\mathbf{Y}_i|} = \frac{\Sigma^{1/2}\mathbf{Z}_i}{|\Sigma^{1/2}\mathbf{Z}_i|}, \quad i = 1, \dots, n.$$

Hence ω_i 's no longer play a role, and this is exactly the reason why we make no assumption on ω_i 's. There is, however, no such thing as a free lunch. Self-normalization introduces a new challenge in that the entries of $\Sigma^{1/2}\mathbf{Z}_i/|\Sigma^{1/2}\mathbf{Z}_i|$ are dependent in an unusual fashion. To see this, consider the simplest case where $\Sigma = \mathbf{I}$ and $\mathbf{Z}_i$'s are i.i.d. standard multivariate normal random vectors. In this case, the entries of $\mathbf{Z}_i$'s are i.i.d. standard normal random variables. However, the self-normalized random vector $\mathbf{Z}_i/|\mathbf{Z}_i|$ is uniformly distributed over the p -dimensional unit sphere, and its p entries are dependent on each other in an unconventional way.

To conduct tests, we need some kind of CLTs. Our strategy is to establish a CLT for the *linear spectral statistic* (LSS) of the sample covariance matrix based on the self-normalized observations, namely,

$$\tilde{\mathbf{S}}_n = \frac{\text{tr}(\Sigma)}{n} \sum_{i=1}^n \frac{\mathbf{Y}_i \mathbf{Y}_i^T}{|\mathbf{Y}_i|^2} = \frac{\text{tr}(\Sigma)}{n} \sum_{i=1}^n \frac{\Sigma^{1/2} \mathbf{Z}_i \mathbf{Z}_i^T \Sigma^{1/2}}{|\Sigma^{1/2} \mathbf{Z}_i|^2}. \quad (4)$$

When $|\mathbf{Y}_i|$ or $|\Sigma^{1/2}\mathbf{Z}_i| = 0$, we adopt the convention that $0/0 = 0$.

As we shall see below, our CLT is different from the ones for the usual sample covariance matrix. One important advantage of our result is that, when testing $H_0 : \Sigma \propto \mathbf{I}$, applying our CLT requires neither $\mathbb{E}(Z_{11}^4) = 3$ as in [Bai and Silverstein \(2004\)](#), nor the estimation of $\mathbb{E}(Z_{11}^4)$, which is inevitable in [Najim and Yao \(2016\)](#). Based on the new CLT, we propose two sphericity tests by modifying the likelihood ratio test and John's test. More tests based on general moments of the ESD of $\tilde{\mathbf{S}}_n$ are also constructed. Numerical studies show that, for the sphericity hypothesis, our proposed tests work well even when $\mathbb{E}(Z_{11}^4)$ does not exist. Because heavy-tailedness and heteroskedasticity are commonly encountered in practice, such relaxations are appealing in many real applications. More generally, we can also test $H_0 : \Sigma \propto \Sigma_0$, where Σ_0 is a general non-negative definite matrix and can even be not invertible. We illustrate such a test in Section 3.4 for a case when Σ_0 contains a substantial proportion (1/4 to be precise) of zero eigenvalues.

Remark 1. Independently, [Li and Yao \(2018\)](#) study high-dimensional covariance matrix test under a mixture model. Their test relies on comparing two John's test statistics: one is based on the original data and the other is based on randomly permuted data. There are a couple of major differences between our paper and theirs. Firstly, they only consider sphericity test and so can only test $H_0 : \Sigma \propto \Sigma_0$ when Σ_0 is invertible. Secondly, in [Li and Yao \(2018\)](#), the mixture coefficients (ω_i 's in (3)) are assumed to be i.i.d. and drawn from a distribution with a bounded support. Thirdly, [Li and Yao \(2018\)](#) require independence between the mixture coefficients and the innovation process ($\mathbf{Z}_i$). In our paper, we do not put any assumptions on the mixture coefficients. As we discussed in Section 1.4, such relaxations allow us to accommodate several important stylized features of real data, consequently, make our tests more suitable in many real applications. It can be shown that the test in [Li and Yao \(2018\)](#) is not necessarily consistent under our setup even if Σ in our model (3) is identity. Furthermore, as we can see from the simulation studies, the test in [Li and Yao \(2018\)](#) is less powerful than the existing tests in the i.i.d. Gaussian setting and, in general, substantially less powerful than our tests.

Empirically, we apply the proposed tests to study the correlations among idiosyncratic returns. There have been numerous factor models developed. A very interesting question is whether the idiosyncratic returns under a factor model are uncorrelated. If the answer is yes, then one can conclude that there is no missing factor. Answering such a question is challenging because typically there are tens or hundreds of stocks involved, resulting in thousands of correlations to be tested. As an innovative application, we demonstrate that our tests can be utilized to test uncorrelatedness among idiosyncratic returns. We illustrate the testing procedure by using the CAPM ([Sharpe, 1964](#)) and the Fama–French Three-Factor model ([Fama and French, 1992](#)). The analysis can be translated directly to more comprehensive factor models.

The rest of the paper is organized as follows. In Section 2, we state the CLT for the LSS of $\tilde{\mathbf{S}}_n$, based on which, for testing $H_0 : \Sigma \propto \mathbf{I}$, we derive the asymptotic null distributions of the modified likelihood ratio test statistic and John's test statistic, as well as other test statistics based on general moments of the ESD of $\tilde{\mathbf{S}}_n$. Section 3 examines the finite-sample performance of the proposed sphericity tests, and illustrates how to test $H_0 : \Sigma \propto \Sigma_0$ for a non-invertible non-negative definite matrix Σ_0 . Section 4 is dedicated to a real data analysis, in which we show how our tests can be used to test uncorrelatedness among idiosyncratic returns. Section 5 concludes. All proofs are collected in Appendix.

Finally, we collect some notation that will be used throughout the paper. For any symmetric matrix $\mathbf{A} \in \mathbb{R}^{p \times p}$, $\|\mathbf{A}\|$ stands for the spectral norm and $F^{\mathbf{A}}$ denotes the ESD, that is,

$$\|\mathbf{A}\| = \max_i |\lambda_i^{\mathbf{A}}|, \text{ and } F^{\mathbf{A}}(x) = \frac{1}{p} \sum_{i=1}^p \mathbb{1}_{\{\lambda_i^{\mathbf{A}} \leq x\}}, \text{ for all } x \in \mathbb{R},$$

where $\lambda_i^{\mathbf{A}}$, $i = 1, \dots, p$, are the eigenvalues of $\mathbf{A}$ and $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function. We also denote by $\lambda_{\min}^{\mathbf{A}}$ the smallest eigenvalue of $\mathbf{A}$. For any function f , the associated LSS of $\mathbf{A}$ is given by

$$\int_{-\infty}^{+\infty} f(x) dF^{\mathbf{A}}(x) = \frac{1}{p} \sum_{i=1}^p f(\lambda_i^{\mathbf{A}}).$$

The Stieltjes transform of a distribution G is defined as

$$m_G(z) = \int_{-\infty}^{\infty} \frac{1}{\lambda - z} dG(\lambda), \text{ for all } z \notin \text{supp}(G),$$

where $\text{supp}(G)$ denotes the support of G . For any $y \in (0, \infty)$ and distribution G , $F_{y,G}$ denotes the distribution whose Stieltjes transform $m_{F_{y,G}}(z)$ is the unique solution to

$$m(z) = \int_0^{\infty} \frac{dG(t)}{t(1 - y - yzm(z)) - z}, \quad z \in \mathbb{C}^+ := \{z \in \mathbb{C}, \Im(z) > 0\},$$

in the set $\{m(z) \in \mathbb{C} : -(1 - y)/z + ym(z) \in \mathbb{C}^+\}$, where $\Im(z)$ denotes the imaginary part of z . Finally, for any $x \in \mathbb{R}$, δ_x is the Dirac measure at x , and we denote F_{y,δ_1} as F_y .

2. Main results

2.1. CLT for the LSS of $\tilde{\mathbf{S}}_n$

As discussed above, we focus on the sample covariance matrix based on the self-normalized observations, namely, $\tilde{\mathbf{S}}_n$ defined in (4).

We make the following assumptions:

Assumption A. The random variables Z_{ij} 's are i.i.d. with $\mathbb{E}(Z_{11}) = 0$, $\mathbb{E}(Z_{11}^2) = 1$ and $\mathbb{E}(Z_{11}^4) < \infty$.

Assumption B. The matrix $\Sigma = \Sigma^{(p)}$ satisfies that $H_p \xrightarrow{D} H$ as $p \rightarrow \infty$, where $H_p = F^{\Sigma^{(p)}}$ and H is a distribution. Furthermore, $\text{tr}(\Sigma^{(p)}) \asymp p$ and $\sup_p \|\Sigma^{(p)}\| \leq C$ for some constant $C > 0$.

Assumption C. $y_n := p/n \rightarrow y \in (0, \infty)$ as $n \rightarrow \infty$.

Theorem 2 in Zheng and Li (2011) states that under some regularity conditions, $\tilde{\mathbf{S}}_n$ shares the same LSD as the sample covariance matrix $\mathbf{S}_n := n^{-1} \sum_{i=1}^n \Sigma^{1/2} \mathbf{Z}_i \mathbf{Z}_i^T \Sigma^{1/2}$. To conduct tests, we need the associated CLT. The CLTs for the LSS of $\mathbf{S}_n$ have been established in Bai and Silverman (2004) and Najim and Yao (2016), under the Gaussian and non-Gaussian kurtosis conditions, respectively. Given that $\tilde{\mathbf{S}}_n$ and $\mathbf{S}_n$ have the same LSD, one naturally asks whether their LSSs also have the same CLT. The following theorem gives a negative answer. Hence, an important message is:

Self-normalization does not change the LSD, but it does affect the CLT.

To be more specific, for any function f , define the following centered and scaled LSS:

$$G_{\tilde{\mathbf{S}}_n}(f) := p \int_0^{+\infty} f(x) d(F^{\tilde{\mathbf{S}}_n}(x) - F_{y_n, H_p}(x)). \quad (5)$$

Theorem 1. Let $a_l(\Sigma, y) = \lim_{n \rightarrow \infty} \lambda_{\min}^{\Sigma} \mathbb{1}_{\{0 < y < 1\}} (1 - \sqrt{y})^2$, $a_r(\Sigma, y) = \overline{\lim}_{n \rightarrow \infty} \|\Sigma\| (1 + \sqrt{y})^2$, $\mathcal{H}$ be the set of functions that are analytic on a domain containing $[a_l(\Sigma, y), a_r(\Sigma, y)]$, and $f_1, \dots, f_k \in \mathcal{H}$.

(i) Under Assumptions A–C, the sequence of random vectors $\{(G_{\tilde{\mathbf{S}}_n}(f_1), \dots, G_{\tilde{\mathbf{S}}_n}(f_k))\}$ is tight. Furthermore, if $\mathbb{E}(Z_{11}^4) = 3$, then the random vector $(G_{\tilde{\mathbf{S}}_n}(f_1), \dots, G_{\tilde{\mathbf{S}}_n}(f_k))$ converges weakly to a Gaussian vector $(G(f_1), \dots, G(f_k))$ with mean

$$\begin{aligned} \mathbb{E}(G(f_i)) = & -\frac{1}{2\pi i} \oint_{\mathcal{C}} f_i(z) \int_0^{\infty} \frac{ym^3(z)t^2 dH(t)}{(1 + t\overline{m}(z))^3} \left(1 - \int_0^{\infty} \frac{ym^2(z)t^2 dH(t)}{(1 + t\overline{m}(z))^2}\right)^{-2} dz \\ & + \frac{1}{2\pi i} \oint_{\mathcal{C}} f_i(z) R_0(z) \overline{m}(z) \left(1 - \int_0^{\infty} \frac{ym^2(z)t^2 dH(t)}{(1 + t\overline{m}(z))^2}\right)^{-1} dz, \quad i = 1, \dots, k, \end{aligned} \quad (6)$$

where

$$\begin{aligned} R_0(z) = & \frac{2 \int_0^\infty t^2 dH(t)}{(\int_0^\infty t dH(t))^2} \int_0^\infty \frac{t \underline{m}(z) dH(t)}{(1 + t \underline{m}(z))^2} - \frac{2}{\int_0^\infty t dH(t)} \int_0^\infty \frac{t^2 \underline{m}(z) dH(t)}{(1 + t \underline{m}(z))^2} \\ & + \frac{2y}{\int_0^\infty t dH(t)} \int_0^\infty \frac{t \underline{m}(z) dH(t)}{1 + t \underline{m}(z)} \int_0^\infty \frac{t^2 \underline{m}(z) dH(t)}{(1 + t \underline{m}(z))^2} \\ & + \frac{2y}{\int_0^\infty t dH(t)} \int_0^\infty \frac{t^2 \underline{m}(z) dH(t)}{1 + t \underline{m}(z)} \int_0^\infty \frac{t \underline{m}(z) dH(t)}{(1 + t \underline{m}(z))^2} \\ & - \frac{2y \int_0^\infty t^2 dH(t)}{(\int_0^\infty t dH(t))^2} \int_0^\infty \frac{t \underline{m}(z) dH(t)}{1 + t \underline{m}(z)} \int_0^\infty \frac{t \underline{m}(z) dH(t)}{(1 + t \underline{m}(z))^2}, \end{aligned}$$

and covariance

$$\begin{aligned} & \text{Cov}(G(f_i), G(f_j)) \\ = & -\frac{1}{2\pi^2} \oint_{c_2} \oint_{c_1} \frac{f_i(z_1) f_j(z_2) \underline{m}'(z_1) \underline{m}'(z_2)}{(\underline{m}(z_2) - \underline{m}(z_1))^2} dz_1 dz_2 \\ & + \frac{y}{2\pi^2 \int_0^\infty t dH(t)} \oint_{c_2} \oint_{c_1} \int_0^\infty \frac{t f_i(z_1) \underline{m}'(z_1) dH(t)}{(1 + t \underline{m}(z_1))^2} \int_0^\infty \frac{t^2 f_j(z_2) \underline{m}'(z_2) dH(t)}{(1 + t \underline{m}(z_2))^2} dz_1 dz_2 \\ & + \frac{y}{2\pi^2 \int_0^\infty t dH(t)} \oint_{c_2} \oint_{c_1} \int_0^\infty \frac{t^2 f_i(z_1) \underline{m}'(z_1) dH(t)}{(1 + t \underline{m}(z_1))^2} \int_0^\infty \frac{t f_j(z_2) \underline{m}'(z_2) dH(t)}{(1 + t \underline{m}(z_2))^2} dz_1 dz_2 \\ & - \frac{y \int_0^\infty t^2 dH(t)}{2\pi^2 (\int_0^\infty t dH(t))^2} \oint_{c_2} \oint_{c_1} \int_0^\infty \frac{t f_i(z_1) \underline{m}'(z_1) dH(t)}{(1 + t \underline{m}(z_1))^2} \int_0^\infty \frac{t f_j(z_2) \underline{m}'(z_2) dH(t)}{(1 + t \underline{m}(z_2))^2} dz_1 dz_2, \end{aligned} \quad (7)$$

$i, j = 1, \dots, k$, where $\underline{m}(z)$ is the Stieltjes transform of $F_{y,H} := (1-y)\mathbb{1}_{[0,\infty)} + yF_{y,H}$, and c_1 and c_2 are two non-overlapping contours contained in the domain and enclosing the interval $[a_l(\Sigma, y), a_r(\Sigma, y)]$;

(ii) If $\Sigma = \mathbf{I}$, then under [Assumptions A](#) and [C](#) and without assuming $\mathbb{E}(Z_{11}^4) = 3$, the weak convergence in (i) still holds, and the mean and covariance admit the following simpler expressions:

$$\begin{aligned} \mathbb{E}(G(f_i)) = & -\frac{1}{2\pi i} \oint_c f_i(z) \left(\frac{y \underline{m}^3(z)}{(1 + \underline{m}(z))^3} \right) \left(1 - \frac{y \underline{m}^2(z)}{(1 + \underline{m}(z))^2} \right)^{-2} dz \\ & + \frac{1}{\pi i} \oint_c f_i(z) \left(\frac{y \underline{m}^3(z)}{(1 + \underline{m}(z))^3} \right) \left(1 - \frac{y \underline{m}^2(z)}{(1 + \underline{m}(z))^2} \right)^{-1} dz; \end{aligned} \quad (8)$$

$$\begin{aligned} \text{Cov}(G(f_i), G(f_j)) = & -\frac{1}{2\pi^2} \oint_{c_2} \oint_{c_1} \frac{f_i(z_1) f_j(z_2) \underline{m}'(z_1) \underline{m}'(z_2)}{(\underline{m}(z_2) - \underline{m}(z_1))^2} dz_1 dz_2 \\ & + \frac{y}{2\pi^2} \oint_{c_2} \oint_{c_1} \frac{f_i(z_1) f_j(z_2) \underline{m}'(z_1) \underline{m}'(z_2)}{(1 + \underline{m}(z_1))^2 (1 + \underline{m}(z_2))^2} dz_1 dz_2, \end{aligned} \quad (9)$$

where $i, j = 1, \dots, k$.

Remark 2. The first terms in (6) and (7) appear in equations (1.6) and (1.7) of [Bai and Silverstein \(2004\)](#). The first terms in (8) and (9) equal the terms in equations (1.6) and (1.7) of [Bai and Silverstein \(2004\)](#) with $\Sigma = \mathbf{I}$. The other terms in (6)–(9) are new and are due to the self-normalization in $\tilde{\mathbf{S}}_n$. It is worth emphasizing that, when $\Sigma = \mathbf{I}$, our CLT neither requires $\mathbb{E}(Z_{11}^4) = 3$ as in [Bai and Silverstein \(2004\)](#) nor involves $\mathbb{E}(Z_{11}^4)$ as in [Najim and Yao \(2016\)](#).

Remark 3. Our CLT allows Σ to be not invertible and so can be used to test $H_0 : \Sigma \propto \Sigma_0$ even when Σ_0 is not invertible. This is an important contribution in the covariance matrix testing literature because existing methods rely on transforming testing $H_0 : \Sigma \propto \Sigma_0$ to $H_0 : \Sigma \propto \mathbf{I}$ by multiplying observations with $\Sigma_0^{-1/2}$.

Remark 4. When $\Sigma = \mathbf{I}$, $\tilde{\mathbf{S}}_n$ can be regarded as a sample correlation matrix (without demeaning). The CLT for the LSS of sample correlation matrices when the population covariance matrix is identity has been derived in [Gao et al. \(2017\)](#).

2.2. Tests of sphericity in the presence of heteroskedasticity

For sphericity test

$$H_0 : \Sigma \propto \mathbf{I} \quad \text{vs.} \quad H_a : \Sigma \not\propto \mathbf{I}, \quad (10)$$

based on [Theorem 1](#), we propose two tests by modifying the likelihood ratio test and John's test. More tests based on general moments of the ESD of $\tilde{\mathbf{S}}_n$ are also established.

2.2.1. Likelihood ratio test based on self-normalized observations (LR-SN)

The classical likelihood ratio test statistic is

$$L_n = \log |\mathbf{S}_n| - p \log(\text{tr}(\mathbf{S}_n)) + p \log p;$$

see, e.g., Section 8.3.1 in [Muirhead \(1982\)](#). For the heteroskedastic case, we modify the likelihood ratio test statistic by replacing $\mathbf{S}_n$ with $\tilde{\mathbf{S}}_n$. Note that $\text{tr}(\tilde{\mathbf{S}}_n) = p$ on the event $\{|\mathbf{Z}_i| > 0 \text{ for } i = 1, \dots, n\}$, which, by Lemma 2 in [Bai and Yin \(1993\)](#), occurs with a probability approaching 1 as $n \rightarrow \infty$. Therefore, we are led to the following modified likelihood ratio test statistic:

$$\tilde{L}_n = \log |\tilde{\mathbf{S}}_n| = \sum_{i=1}^p \log(\lambda_i^{\tilde{\mathbf{S}}_n}).$$

It is the LSS of $\tilde{\mathbf{S}}_n$ when $f(x) = \log(x)$. In this case, when $y_n \in (0, 1)$, we have

$$\begin{aligned} G_{\tilde{\mathbf{S}}_n}(\log) &= p \int_{-\infty}^{+\infty} \log(x) d(F_{\tilde{\mathbf{S}}_n}(x) - F_{y_n}(x)) \\ &= \sum_{i=1}^p \log(\lambda_i^{\tilde{\mathbf{S}}_n}) - p \left(\frac{y_n - 1}{y_n} \log(1 - y_n) - 1 \right) \\ &= \tilde{L}_n - p \left(\frac{y_n - 1}{y_n} \log(1 - y_n) - 1 \right). \end{aligned}$$

Applying [Theorem 1](#), we obtain the following proposition.

Proposition 1. When $y_n \rightarrow y \in (0, 1)$, under [Assumption A](#), we have

$$\frac{\tilde{L}_n - p \left(\frac{y_n - 1}{y_n} \log(1 - y_n) - 1 \right) - (\log(1 - y_n))/2 - y_n}{\sqrt{-2 \log(1 - y_n) - 2y_n}} \xrightarrow{D} N(0, 1). \quad (11)$$

The convergence in (11) gives the asymptotic null distribution of the modified likelihood ratio test statistic. Because it is derived for the sample covariance matrix based on self-normalized observations, the test based on (11) will be referred to as the likelihood ratio test based on the self-normalized observations (LR-SN).

2.2.2. John's test based on self-normalized observations (JHN-SN)

John's test statistic ([John, 1971](#)) is given by

$$T_n = \frac{n}{p} \text{tr} \left(\frac{\mathbf{S}_n}{1/p \text{tr}(\mathbf{S}_n)} - \mathbf{I} \right)^2 - p.$$

Replacing $\mathbf{S}_n$ with $\tilde{\mathbf{S}}_n$ and noting again that $\text{tr}(\tilde{\mathbf{S}}_n) = p$ almost surely for all large n lead to the following modified John's test statistic:

$$\tilde{T}_n = \frac{n}{p} \text{tr}(\tilde{\mathbf{S}}_n - \mathbf{I})^2 - p = \frac{1}{y_n} \sum_{i=1}^p (\lambda_i^{\tilde{\mathbf{S}}_n})^2 - n - p.$$

It is related to the LSS of $\tilde{\mathbf{S}}_n$ when $f(x) = x^2$. In this case, we have

$$G_{\tilde{\mathbf{S}}_n}(x^2) = p \int_{-\infty}^{+\infty} x^2 d(F_{\tilde{\mathbf{S}}_n}(x) - F_{y_n}(x)) = \sum_{i=1}^p (\lambda_i^{\tilde{\mathbf{S}}_n})^2 - p(1 + y_n) = y_n \tilde{T}_n.$$

Based on [Theorem 1](#), we can prove the following proposition.

Proposition 2. Under Assumptions A and C, we have

$$\frac{\tilde{T}_n + 1}{2} \xrightarrow{D} N(0, 1). \quad (12)$$

Below we will refer to the test based on (12) as John's test based on the self-normalized observations (JHN-SN).

2.2.3. More general tests based on self-normalized observations

More tests can be constructed by choosing f in Theorem 1 to be different functions. When $f(x) = x^k$ for $k \geq 2$, the corresponding LSS is the k th moment of the ESD of $\tilde{\mathbf{S}}_n$, for which we have

$$\begin{aligned} G_{\tilde{\mathbf{S}}_n}(x^k) &= p \int_{-\infty}^{+\infty} x^k d(F_{\tilde{\mathbf{S}}_n}(x) - F_{y_n}(x)) \\ &= \sum_{i=1}^p (\lambda_i^{\tilde{\mathbf{S}}_n})^k - p(1 + y_n)^{k-1} H_F\left(\frac{1-k}{2}, 1 - \frac{k}{2}, 2, \frac{4y_n}{(1+y_n)^2}\right), \end{aligned}$$

where $H_F(a, b, c, d)$ denotes the hypergeometric function ${}_2F_1(a, b, c, d)$. By Theorem 1 again, we have the following proposition.

Proposition 3. Under Assumptions A and C, for any $k \geq 2$, we have

$$\frac{G_{\tilde{\mathbf{S}}_n}(x^k) - \mu_{n,x^k}}{\sigma_{n,x^k}} \xrightarrow{D} N(0, 1), \quad \text{where}$$

$$\begin{aligned} \mu_{n,x^k} &= \frac{1}{4} \left((1 + \sqrt{y_n})^{2k} + (1 - \sqrt{y_n})^{2k} \right) - \frac{1}{2} \sum_{i=0}^k \binom{k}{i}^2 y_n^i \\ &\quad - \frac{2k(k-1)(1+y_n)^{k-2}}{(k+1)(k+2)} \left((y_n - 1)^2 H_F\left(\frac{3-k}{2}, 1 - \frac{k}{2}, 1, \frac{4y_n}{(1+y_n)^2}\right) \right. \\ &\quad \left. + (-1 + 4ky_n - y_n^2) H_F\left(\frac{3-k}{2}, 1 - \frac{k}{2}, 2, \frac{4y_n}{(1+y_n)^2}\right) \right), \end{aligned}$$

and

$$\begin{aligned} \sigma_{n,x^k}^2 &= 2y_n^{2k} \sum_{i=0}^{k-1} \sum_{j=0}^k \binom{k}{i} \binom{k}{j} \left(\frac{1-y_n}{y_n}\right)^{i+j} \sum_{\ell=1}^{k-i} \ell \binom{2k-1-(i+\ell)}{k-1} \binom{2k-1-j+\ell}{k-1} \\ &\quad - 2y_n \left((1-y_n)^k k \sum_{i=0}^{k+1} \binom{k+1}{i} \left(\frac{1-y_n}{y_n}\right)^{1-i} \frac{(k+i-1)!}{(i-1)!(k+1)!} \right)^2. \end{aligned}$$

Remark 5. Proposition 3 enables us to consistently detect any alternative hypothesis under which the covariance matrix admits an LSD not equal to δ_1 . The reason is that, under such a situation, the LSD of $\tilde{\mathbf{S}}_n$, say $\tilde{H}$, will not be the standard Marčenko–Pastur law F_y . Therefore, there exists a $k \geq 2$ such that $\int_{-\infty}^{\infty} x^k d\tilde{H}(x) \neq \int_{-\infty}^{\infty} x^k dF_y(x)$. Consequently, $G_{\tilde{\mathbf{S}}_n}(x^k)$ in (5) will blow up, and the testing power will approach 1.

3. Simulation studies

We now demonstrate the finite-sample performance of our proposed tests in Section 2.2, as well as the test for $H_0 : \Sigma \propto \Sigma_0$, where Σ_0 is a general non-negative definite matrix.

3.1. I.i.d. Gaussian case

To have a full picture of the performance of our tests, we start with the simplest situation where observations are i.i.d. multivariate normal random vectors. We will compare our proposed tests, LR-SN and JHN-SN, with the tests mentioned in Section 1.1, namely, LW₂, S, CZZ₂ and WY-LR, and also the test in Li and Yao (2018) (LY test). In the multivariate normal case, the WY-JHN test reduces to the LW₂ test.

We start with the size evaluation by sampling observations from $N(\mathbf{0}, \mathbf{I})$. Table 2 reports the empirical sizes of these tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level.

From Table 2, we find that the empirical sizes of all tests are around the nominal level of 5%.

Next, to compare the power, we generate i.i.d. observations from $N(\mathbf{0}, \Sigma)$ under the alternative with $\Sigma = (0.1^{|i-j|})_{p \times p}$, and test $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Table 3 reports the empirical powers.

Table 2

Empirical sizes (%) of LW_2 , S , CZZ_2 , $WY-LR$, LY , and the $LR-SN$ and $JHN-SN$ tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Observations are *i.i.d.* $N(\mathbf{0}, \mathbf{I})$. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$							$p/n = 2$				
	LW_2	S	CZZ_2	$WY-LR$	LY	$LR-SN$	$JHN-SN$	LW_2	S	CZZ_2	LY	$JHN-SN$
100	4.9	4.8	4.9	4.5	4.8	4.6	5.2	5.5	5.5	5.7	5.1	4.9
200	5.2	5.0	5.1	5.1	4.8	5.1	4.9	4.6	4.5	5.1	4.8	4.5
500	4.9	5.1	5.1	4.8	5.3	4.9	5.2	5.1	5.3	4.9	5.0	5.2

Table 3

Empirical powers (%) of LW_2 , S , CZZ_2 , $WY-LR$, LY , and the $LR-SN$ and $JHN-SN$ tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Observations are *i.i.d.* $N(\mathbf{0}, (0.1^{|i-j|}))$. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$							$p/n = 2$				
	LW_2	S	CZZ_2	$WY-LR$	LY	$LR-SN$	$JHN-SN$	LW_2	S	CZZ_2	LY	$JHN-SN$
100	50.7	51.3	50.1	36.7	28.1	35.0	48.9	8.4	8.7	9.1	6.3	8.2
200	97.3	97.3	97.2	88.0	79.4	88.7	97.0	18.3	17.9	18.1	11.9	17.2
500	100	100	100	100	100	100	100	70.7	70.6	69.8	43.3	70.5

Table 4

Empirical sizes (%) of LW_2 , S , CZZ_2 , $WY-LR$, $WY-JHN$, LY tests, and our proposed $LR-SN$, $JHN-SN$ tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Data are generated as $\mathbf{Y}_i = \omega_i \mathbf{Z}_i$ where ω_i 's are absolute values of *i.i.d.* $N(0, 1)$, $\mathbf{Z}_i$'s are *i.i.d.* $N(\mathbf{0}, \mathbf{I})$, and further ω_i 's and $\mathbf{Z}_i$'s are independent of each other. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$								$p/n = 2$					
	LW_2	S	CZZ_2	$WY-LR$	$WY-JHN$	LY	$LR-SN$	$JHN-SN$	LW_2	S	CZZ_2	$WY-JHN$	LY	$JHN-SN$
100	100	100	51.8	100	100	4.4	4.6	5.2	100	100	50.2	100	4.1	4.9
200	100	100	53.0	100	100	4.5	5.1	4.9	100	100	52.3	100	4.5	4.5
500	100	100	52.3	100	100	5.2	4.9	5.2	100	100	53.5	100	4.7	5.2

From Table 3, we find that our proposed $LR-SN$ and $JHN-SN$ tests and the tests mentioned in Section 1.1 have quite high powers especially as the dimension gets higher, and the powers are roughly comparable. Same as in the classical setting, John's test ($JHN-SN$) is more powerful than the likelihood ratio test ($LR-SN$). LY test proposed in Li and Yao (2018) is less powerful.

To sum up, while developed under a much more general setup, our tests perform just as well as the existing ones in the ideal *i.i.d.* Gaussian setting.

3.2. The elliptical case

We now investigate the performance of our proposed tests under the elliptical distribution. As in Section 1.3, we take the observations to be $\mathbf{Y}_i = \omega_i \mathbf{Z}_i$ with

- (i) ω_i 's being absolute values of *i.i.d.* standard normal random variables,
- (ii) $\mathbf{Z}_i$'s *i.i.d.* p -dimensional random vectors from $N(\mathbf{0}, \Sigma)$, and
- (iii) ω_i 's and $\mathbf{Z}_i$'s independent of each other.

Checking the size.

Table 4 completes Table 1 by including the empirical sizes of our proposed $LR-SN$ and $JHN-SN$ tests, and also LY test in Li and Yao (2018).

Table 4 reveals sharp difference between the existing tests and our proposed ones: the empirical sizes of the existing tests are severely distorted, in contrast, the empirical sizes of our $LR-SN$ and $JHN-SN$ tests are around the nominal level of 5% as desired. LY test also yields the right level of size.

Checking the power.

Table 4 shows that LW_2 , S , CZZ_2 , $WY-LR$ and $WY-JHN$ tests are inconsistent under the elliptical distribution, therefore we exclude them when checking the power.

We generate observations under the elliptical distribution with $\Sigma = (0.1^{|i-j|})$. Table 5 reports the empirical powers of our proposed tests and LY test for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level.

From Table 5, we find that

- (i) Our tests, $LR-SN$ and $JHN-SN$, as well as LY test, enjoy a blessing of dimensionality: for a fixed ratio p/n , the higher the dimension p , the higher the power;

Table 5

Empirical powers (%) of LY test and our proposed LR-SN and JHN-SN tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Data are generated as $\mathbf{Y}_i = \omega_i \mathbf{Z}_i$ where ω_i 's are absolute values of i.i.d. $N(0, 1)$, $\mathbf{Z}_i$'s are i.i.d. random vectors from $N(\mathbf{0}, \Sigma)$ with $\Sigma = (0.1^{|i-j|})$, and further ω_i 's and $\mathbf{Z}_i$'s are independent of each other. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$			$p/n = 2$	
	LY	LR-SN	JHN-SN	LY	JHN-SN
100	7.6	35.0	48.9	3.5	8.2
200	14.5	88.7	97.0	5.7	17.2
500	64.9	100	100	9.0	70.5

Table 6

Empirical sizes (%) of LY test and our proposed LR-SN and JHN-SN tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Data are generated as $\mathbf{Y}_i = \omega_i \mathbf{Z}_i$ with $\omega_i^2 = 0.01 + 0.85\omega_{i-1}^2 + 0.1|\mathbf{Y}_{i-1}|^2/p$, and $\mathbf{Z}_i$'s consist of i.i.d. standardized $t(4)$ random variables. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$			$p/n = 2$	
	LY	LR-SN	JHN-SN	LY	JHN-SN
100	8.2	5.5	5.3	6.8	5.0
200	8.5	5.7	5.4	6.8	5.5
500	7.6	5.3	5.2	6.6	5.4

(ii) LY test is substantially less powerful than our tests.

3.3. Beyond elliptical, a GARCH-type case

Recall that in our general model (3), the observations $\mathbf{Y}_i$ admit the decomposition $\omega_i \Sigma^{1/2} \mathbf{Z}_i$, and ω_i 's can depend on each other and on $\{\mathbf{Z}_i : i = 1, \dots, n\}$ in an arbitrary way. To examine the performance of our tests in such a general setup, we simulate data using the following two-step procedure:

1. Sample $\mathbf{Z}_{ij}$'s from the standardized t -distribution with four degrees of freedom. The distribution is heavy-tailed and even does not have a finite fourth moment.
2. For each ω_i , inspired by the ARCH/GARCH model, we take $\omega_i^2 = 0.01 + 0.85\omega_{i-1}^2 + 0.1|\mathbf{Y}_{i-1}|^2/\text{tr}(\Sigma)$.

Checking the size.

We take $\Sigma = \mathbf{I}$ in the data generating process and test $H_0 : \Sigma \propto \mathbf{I}$. Table 6 reports the empirical sizes of our proposed tests and LY test at 5% significance level.

From Table 6, we find that, for all different values of p and p/n , the empirical sizes of our proposed tests are around the nominal level of 5%. Again, this is in sharp contrast with the results in Table 1, where the existing tests yield sizes far higher than 5%.

The second observation is that although Theorem 1 requires the finiteness of $\mathbb{E}(Z_{11}^4)$, the simulation above shows that our proposed tests work well even when $\mathbb{E}(Z_{11}^4)$ does not exist.

Another observation is that the empirical sizes of LY test are around 8%. Note that with 10,000 replications, the margin of error at 5% significance level is 1%, hence we can conclude that the sizes of LY test are statistically significantly higher than the nominal level of 5%.

Checking the power.

To evaluate the power, we again take $\Sigma = (0.1^{|i-j|})$ and generate data according to the design at the beginning of this subsection. Table 7 reports the empirical powers of our proposed tests and LY test for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level.

Table 7 shows again that our tests enjoy a blessing of dimensionality. Moreover, comparing Table 7 with Table 5, we find that for each pair of p and n , the powers of our tests are similar under the two designs. Such similarities show that our tests can not only accommodate (conditional) heteroskedasticity but also are robust to heavy-tailedness in $\mathbf{Z}_i$'s. Finally, LY test is again significantly less powerful.

3.4. A GARCH-type case with a general Σ_0

In this subsection, we first illustrate how to construct the test for general Σ_0 based on Theorem 1 and then investigate the finite-sample performance of the test. We take $\Sigma_0 = \mathbf{Q}\Lambda_0\mathbf{Q}^\top$, where $\Lambda_0 = \text{diag}(\underbrace{1, \dots, 1}_{p/2}, \underbrace{2, \dots, 2}_{p/4}, \underbrace{0, \dots, 0}_{p/4})$ and $\mathbf{Q}$ is a random orthogonal matrix. To build the test, we take $f(x) = x^2$.

Table 7

Empirical powers (%) of LY test and our proposed LR-SN and JHN-SN tests for testing $H_0 : \Sigma \propto \mathbf{I}$ at 5% significance level. Data are generated as $\mathbf{Y}_i = \omega_i \Sigma^{1/2} \mathbf{Z}_i$ with $\omega_i^2 = 0.01 + 0.85\omega_{i-1}^2 + 0.1|\mathbf{Y}_{i-1}|^2/p$ and $\Sigma = (0.1^{i-j})$, and $\mathbf{Z}_i$'s consist of i.i.d. standardized $t(4)$ random variables. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$			$p/n = 2$	
	LY	LR-SN	JHN-SN	LY	JHN-SN
100	20.7	34.4	47.9	7.8	8.7
200	54.4	87.8	96.6	10.5	17.6
500	100	100	100	26.4	69.9

Table 8

Empirical sizes (%) of our test for testing $H_0 : \Sigma \propto \Sigma_0$ at 5% significance level. Here $\Sigma_0 = \mathbf{Q}\Lambda_0\mathbf{Q}^\top$, where $\Lambda_0 = \text{diag}(\underbrace{1, \dots, 1}_{p/2}, \underbrace{2, \dots, 2}_{p/4}, \underbrace{0, \dots, 0}_{p/4})$ and $\mathbf{Q}$ is a random orthogonal matrix. Data are generated as $\mathbf{Y}_i = \omega_i \Sigma_0^{1/2} \mathbf{Z}_i$ with $\omega_i^2 = 0.01 + 0.85\omega_{i-1}^2 + 0.1|\mathbf{Y}_{i-1}|^2/p$ and $\mathbf{Z}_i$'s from $N(\mathbf{0}, \mathbf{I})$. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$	$p/n = 2$
100	4.8	4.4
200	4.9	4.7
500	5.1	4.6

We need to compute the asymptotic mean in Eq. (6). Note that by equation (1.2) in Bai and Silverstein (2004), we have

$$z(\underline{m}) = -\frac{1}{\underline{m}} + y\left(\frac{1}{2(1+\underline{m})} + \frac{1}{2(1+2\underline{m})}\right), \quad (13)$$

where $\underline{m} = \underline{m}(z)$. By the Chain rule with the transformation (13), we can find that the asymptotic mean in (6) equals the sum of some contour integrals with respect to $\underline{m}$, and the singularities are -1 , $-1/2$ and the roots of

$$\left(1 - \int_0^\infty \frac{y \underline{m}^2 t^2 dH(t)}{(1 + t \underline{m})^2}\right)^{-1} = 0, \text{ where } H = \frac{1}{2}\delta_1 + \frac{1}{4}\delta_2 + \frac{1}{4}\delta_0.$$

Then we can compute the asymptotic mean in (6) using the residual theorem. Similarly, note that $d\underline{m} = \underline{m}'(z)dz$. By the Chain rule, (13) and the residue theorem, we can compute the asymptotic covariance in (7).

The observations are still generated according to the design in Section 3.3 except that Z_{ij} 's are sampled from the standard normal distribution.

Checking the size.

We take $\Sigma = \Sigma_0$ to generate the data and test $H_0 : \Sigma \propto \Sigma_0$. Table 8 reports the empirical sizes of the test at 5% significance level.

We find that from Table 8 the empirical sizes of the test are around the nominal level of 5% for all different values of p and p/n .

Checking the power.

To examine the power, we generate the observations with $\Sigma = \mathbf{Q}\Lambda\mathbf{Q}^\top$, where $\Lambda = \text{diag}(\underbrace{1, \dots, 1}_{p/2}, \underbrace{2.5, \dots, 2.5}_{p/4}, \underbrace{0, \dots, 0}_{p/4})$

and $\mathbf{Q}$ is a random orthogonal matrix. Table 9 reports the empirical powers of our test for testing $H_0 : \Sigma \propto \Sigma_0$ at 5% significance level.

Again, we see that our test enjoys the blessing of dimensionality and has high powers.

3.5. Summary of simulation studies

Combining the observations in the four cases, we conclude that

- The existing tests, LW_2 , S , CZZ_2 , WY -LR and WY -JHN, work well in the i.i.d. Gaussian setting, however, they fail badly under the elliptical distribution and our general setup;
- The newly proposed LY test in Li and Yao (2018) is applicable to the elliptical distribution, however, it is less powerful than the existing tests in the i.i.d. Gaussian setting and substantially less powerful than ours in general situations;
- Our LR-SN and JHN-SN tests perform well under all the settings, yielding the right sizes and enjoying high powers;
- Even when Σ_0 is not invertible, our test still works well.

Table 9
Empirical power (%) of our test for testing $H_0 : \Sigma \propto \Sigma_0$ at 5% significance level. Data are generated as $\mathbf{Y}_i = \omega_i \Sigma^{1/2} \mathbf{Z}_i$ with $\omega_i^2 = 0.01 + 0.85\omega_{i-1}^2 + 0.1|\mathbf{Y}_{i-1}|^2/p$ and $\mathbf{Z}_i$'s from $N(\mathbf{0}, \mathbf{I})$. Here $\Sigma = \mathbf{Q}\Lambda\mathbf{Q}^\top$, where $\Lambda = \text{diag}(\underbrace{1, \dots, 1}_{p/2}, \underbrace{2.5, \dots, 2.5}_{p/4}, \underbrace{0, \dots, 0}_{p/4})$ and $\mathbf{Q}$ is a random orthogonal matrix. The results are based on 10,000 replications for each pair of p and n .

p	$p/n = 0.5$	$p/n = 2$
100	100	46.3
200	100	95.5
500	100	100

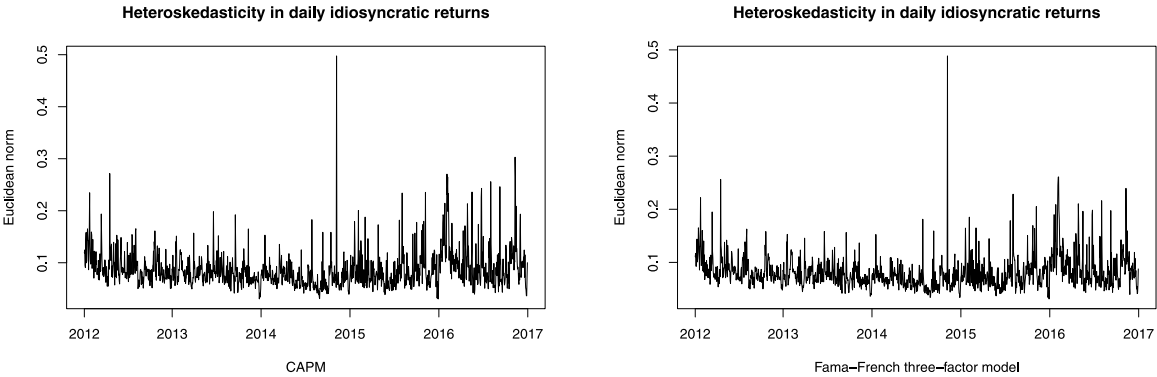


Fig. 1. Time series plots of the Euclidean norms of the daily idiosyncratic returns of 76 stocks in the S&P 500 Financials sector, by fitting the CAPM (left) and the Fama–French three-factor model (right) over the years of 2012–2016.

4. Empirical studies

Let us first explain the motivation of the empirical study, which is about idiosyncratic returns. In general, the total risk of a stock return can be decomposed into two components: systematic risk and idiosyncratic risk. Empirical studies in Campbell et al. (2001) and Goyal and Santa-Clara (2003) show that idiosyncratic risk is the major component of the total risk. It is not uncommon to assume that idiosyncratic returns are cross-sectionally uncorrelated, giving rise to the so-called strict factor model; see, e.g., Roll and Ross (1980), Brown (1989) and Fan et al. (2008). Our goal in this section is to test the cross-sectional uncorrelatedness of idiosyncratic returns.

We focus on the S&P 500 Financials sector. There are in total 80 stocks on the first trading day of 2012 (Jan 3, 2012), among which 76 stocks have complete data over the years of 2012–2016. We will focus on these 76 stocks. The stock prices that our analysis is based on are collected from the Center for Research in Security Prices (CRSP) daily database, while the Fama–French three-factor data are obtained from Kenneth French’s data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

We illustrate the testing procedure based on two most widely used factor models: the CAPM and the Fama–French Three-Factor model. We use a rolling window of six months to fit the two models. Fig. 1 reports the Euclidean norms of the fitted daily idiosyncratic returns.

We see from Fig. 1 that under both models, the Euclidean norms of the fitted daily idiosyncratic returns exhibit clear heteroskedasticity and clustering. Such features indicate that the idiosyncratic returns are unlikely to be homoskedastic, but more suitably modeled as a conditional heteroskedastic time series, which is compatible with our framework.

Now we test the cross-sectional uncorrelatedness of idiosyncratic returns. Specifically, for a diagonal matrix $\Sigma_{\mathcal{D}}$ to be chosen, we test

$$H_0 : \Sigma_{\mathcal{I}} \propto \Sigma_{\mathcal{D}} \text{ vs. } H_a : \Sigma_{\mathcal{I}} \not\propto \Sigma_{\mathcal{D}}, \tag{14}$$

where $\Sigma_{\mathcal{I}}$ denotes the covariance matrix of the idiosyncratic returns. We will test (14) by applying JHN-SN test to transformed idiosyncratic returns by multiplying $\Sigma_{\mathcal{D}}^{-1/2}$ to the (estimated) idiosyncratic returns.

4.1. Testing results

We test (14) using the same rolling window scheme as for fitting the CAPM or the Fama–French three-factor model. For each month to be tested, the diagonal matrix $\Sigma_{\mathcal{D}}$ in (14) is obtained by extracting the diagonal entries of the sample covariance matrix of the self-normalized fitted idiosyncratic returns over the previous five months. Table 10 summarizes the resulting JHN-SN test statistics.

Table 10

Summary statistics of the JHN-SN statistics for testing (14). For both the CAPM and the Fama–French three-factor model, for each month, we first estimate the idiosyncratic returns by fitting the model using the data in the current month and the previous five months. We then obtain $\Sigma_{\mathcal{D}}$ by extracting the diagonal entries of the sample covariance matrix of the self-normalized idiosyncratic returns over the previous five months, and use the fitted idiosyncratic returns in the current month to conduct the test.

CAPM						
JHN-SN	Min	Q_1	Median	Q_3	Max	Mean (SD)
	6.3	18.1	29.9	44.4	82.9	33.1 (18.5)
Fama–French three-factor model						
JHN-SN	Min	Q_1	Median	Q_3	Max	Mean (SD)
	5.0	12.9	24.6	30.3	76.4	23.9 (13.0)

Table 11

Summary statistics of the JHN-SN statistics for testing (14) based on simulated returns from Model (15). To conduct the test, with a rolling window of six months, we first estimate the idiosyncratic returns by fitting the three-factor model. We then obtain $\Sigma_{\mathcal{D}}$ by extracting the diagonal entries of the sample covariance matrix of the self-normalized fitted idiosyncratic returns over the previous five months, and use the fitted idiosyncratic returns in the current month to conduct the test.

Simulated data based on a three-factor model							
JHN-SN	Min	Q_1	Median	Q_3	Max	Mean (SD)	Percent within $[-1.96, 1.96]$
	−1.3	−0.1	0.6	1.1	2.3	0.5 (0.9)	94.5%

We observe from Table 10 that:

- The values of the JHN-SN test statistics are in general rather big, which correspond to almost zero p -values. Such a finding casts doubt on the cross-sectional uncorrelatedness of the idiosyncratic returns from fitting either the CAPM or the Fama–French three-factor model;
- Compared with the CAPM, the Fama–French three-factor model gives rise to idiosyncratic returns that are associated with less extreme test statistics. This confirms that the two additional factors, size and value, do have pervasive impacts on stock returns.

4.2. Robustness check of the testing results

The results in Table 10 are based on testing against the estimated diagonal matrix $\Sigma_{\mathcal{D}}$, which inevitably contains estimation errors. This brings up the following question: are the extreme test statistics in Table 10 due to the estimation error in $\Sigma_{\mathcal{D}}$, or, are they really due to that the idiosyncratic returns are not uncorrelated? To answer this question, we redo the test based on simulated stock returns whose idiosyncratic returns are uncorrelated and exhibit heteroskedasticity.

Specifically, we consider the following three-factor model:

$$\mathbf{r}_t = \mathbf{B}\mathbf{f}_t + \boldsymbol{\varepsilon}_t, \quad \text{with } \mathbf{f}_t \sim N(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f), \quad \boldsymbol{\varepsilon}_t = \omega_t \cdot \boldsymbol{\Sigma}_{\mathcal{X}}^{1/2} \mathbf{Z}_t \quad \text{and } \mathbf{Z}_t \sim N(\mathbf{0}, \mathbf{I}), \quad (15)$$

where $\mathbf{r}_t$ denotes return vector at time t , $\mathbf{B}$ is a factor loading matrix, $\mathbf{f}_t$ represents three factor returns, and $\boldsymbol{\varepsilon}_t$ consists of idiosyncratic returns. To mimic the real data, we calibrate the parameters as follows:

- The factor loading matrix $\mathbf{B}$ is taken to be the estimated factor loading matrix by fitting the Fama–French three-factor model to the daily returns of the 76 stocks over the years of 2012–2016;
- The mean and covariance matrix of factor returns, $\boldsymbol{\mu}_f$ and $\boldsymbol{\Sigma}_f$, are the sample mean and sample covariance matrix of the Fama–French three factor returns from 2012 to 2016;
- To generate data under the null hypothesis that the idiosyncratic returns are uncorrelated, their covariance matrix $\boldsymbol{\Sigma}_{\mathcal{X}}$ is taken to be the diagonal matrix obtained by extracting the diagonal entries of the sample covariance matrix of the self-normalized fitted idiosyncratic returns; and
- Finally, ω_t is taken to be the Euclidean norm of the fitted daily idiosyncratic returns.

With such generated data, we test (14) in parallel with the real data analysis. Table 11 summarizes the JHN-SN test statistics for testing (14) based on the simulated data.

Table 11 reveals sharp contrast with Table 10. We see that if the idiosyncratic returns are indeed uncorrelated, then even if they are heteroskedastic and even if we are testing against the estimated $\widehat{\Sigma}_{\mathcal{D}}$, the percentage of resulting test statistics that are within $[-1.96, 1.96]$ is close to 95%, the expected level under the null hypothesis. In sharp contrast, the test statistics in Table 10 are all very extreme. Such a comparison suggests that the idiosyncratic returns in the real data are indeed unlikely to be uncorrelated.

4.3. Broader usage of the proposed test

The testing procedure above can be directly translated to other factor models. Furthermore, the comparison made in Section 4.1 based on our test statistic can be viewed as a “scoring” system for different factor models. More specifically, less extreme test statistic values would suggest that the model is more effective in accounting for pervasive impact of underlying factors.

5. Conclusions

We study testing high-dimensional covariance matrices under a generalized elliptical distribution, which can feature heteroskedasticity, leverage effect, asymmetry, etc. We establish a CLT for the LSS of the sample covariance matrix based on self-normalized observations. The CLT is different from the existing ones for the usual sample covariance matrix. When the covariance matrix equals the identity matrix, our CLT neither requires $\mathbb{E}(Z_{11}^4) = 3$ as in Bai and Silverstein (2004) nor involves $\mathbb{E}(Z_{11}^4)$ as in Najim and Yao (2016). Based on the new CLT, we propose two sphericity tests by modifying the likelihood ratio test and John’s test. More general tests are also provided. Numerical studies show that our proposed tests work well no matter whether the observations are i.i.d. Gaussian or from an elliptical distribution or feature conditional heteroskedasticity or even when $\mathbf{Z}_i$ ’s do not admit the fourth moment. Moreover, we can also test against a general non-negative definite matrix which can be even not invertible. As an innovative application, we demonstrate that our tests can be utilized to test uncorrelatedness among idiosyncratic returns. The testing procedure is illustrated by using the CAPM and Fama–French Three-Factor model. The analysis can be translated directly to other factor models.

Acknowledgments

We are very grateful to the Editor, Jianqing Fan, and two referees for their valuable comments and constructive suggestions that lead to significant improvements of this paper. We thank Zhigang Bao for a suggestion that helps relax the assumptions of Lemma 2 in Appendix.

This work was partially supported by the disciplinary funding of Central University of Finance and Economics, China [No. SAM201908], Program for Innovation Research in Central University of Finance and Economics [No. 20190107], the Beijing Universities Advanced Disciplines Initiative [No. GJJ2019163] (Yang); RGC, Hong Kong [GRF16305315, GRF16304317, GRF 16304019 and T31-604/18-N] (Zheng); and NSFC 11501147 (Chen).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jeconom.2020.05.017>.

References

- Aït-Sahalia, Y., Fan, J., Li, Y., 2013. The leverage effect puzzle: Disentangling sources of bias at high frequency. *J. Financ. Econ.* 109, 224–249.
- Bai, Z., Jiang, D., Yao, J., Zheng, S., 2009. Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.* 37, 3822–3840.
- Bai, Z., Silverstein, J.W., 2004. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* 32, 553–605.
- Bai, Z.D., Yin, Y.Q., 1993. Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.* 21, 1275–1294.
- Bingham, N.H., Kiesel, R., 2002. Semi-parametric modelling in finance: theoretical foundations. *Quant. Financ.* 2, 241–250.
- Birke, M., Dette, H., 2005. A note on testing the covariance matrix for large dimension. *Statist. Probab. Lett.* 74, 281–289.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* 31, 307–327.
- Brown, S.J., 1989. The number of factors in security returns. *J. Financ.* 44, 1247–1262.
- Campbell, J.Y., Hentschel, L., 1992. No news is good news: An asymmetric model of changing volatility in stock returns. *J. Financ. Econ.* 31, 281–318.
- Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *J. Financ.* 56, 1–43.
- Chen, S.X., Zhang, L.-X., Zhong, P.-S., 2010. Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* 105, 810–819.
- Christoffersen, P., 2012. *Elements of Financial Risk Management*, second ed. Academic Press.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50, 987–1007.
- Fama, E.F., 1965. The behavior of stock-market prices. *J. Bus.* 38, 34–105.
- Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. *J. Financ.* 47, 427–465.
- Fan, J., Fan, Y., Lv, J., 2008. High dimensional covariance matrix estimation using a factor model. *J. Econometrics* 147, 186–197.
- Fang, K.T., Kotz, S., Ng, K.W., 1990. *Symmetric Multivariate and Related Distributions*. In: *Monographs on Statistics and Applied Probability*, vol. 36, Chapman and Hall, Ltd, London.
- Gao, J., Han, X., Pan, G., Yang, Y., 2017. High dimensional correlation matrices: the central limit theorem and its applications. *J. R. Statist. Soc. Ser. B. Statist. Methodol.* 79, 677–693.
- Goyal, A., Santa-Clara, P., 2003. Idiosyncratic risk matters!. *J. Finance* 58, 975–1007.
- Jiang, T., Yang, F., 2013. Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. of Statist.* 41, 2029–2074.
- John, S., 1971. Some optimal multivariate tests. *Biometrika* 58, 123–127.
- Kalnina, I., Xiu, D., 2017. Nonparametric estimation of the leverage effect: a trade-off between robustness and efficiency. *J. Amer. Statist. Assoc.* 112, 384–396.
- Ledoit, O., Wolf, M., 2002. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. of Statist.* 30, 1081–1102.

- Li, W., Yao, J., 2018. On structure testing for component covariance matrices of a high-dimensional mixture. *J. R. Statist. Soc.: Ser. B. (Statist. Methodol.)* 80, 293–318.
- Mandelbrot, B., 1967. The variation of some other speculative prices. *J. Bus.* 40, 393–413.
- McNeil, A.J., Frey, R., Embrechts, P., 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton university press.
- Muirhead, R.J., 1982. *Aspects of Multivariate Statistical Theory*. In: *wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons, Inc, New York.
- Nagao, H., 1973. On some test criteria for covariance matrix. *Ann. of Statist.* 1, 700–709.
- Najim, J., Yao, J., 2016. Gaussian fluctuations for linear spectral statistics of large random covariance matrices. *Ann. Appl. Probab.* 26, 1837–1887.
- Owen, J., Rabinovitch, R., 1983. On the class of elliptical distributions and their applications to the theory of portfolio choice. *J. Financ.* 38, 745–752.
- Peiro, A., 1999. Skewness in financial returns. *J. Bank. Financ.* 23, 847–862.
- Roll, R., Ross, S.A., 1980. An empirical investigation of the arbitrage pricing theory. *J. Financ.* 35, 1073–1103.
- Schwert, G.W., 1989. Why does stock market volatility change over time?. *J. Finance* 44, 1115–1153.
- Sharpe, W., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Financ.* 19, 425–442.
- Singleton, J.C., Wingender, J., 1986. Skewness persistence in common stock returns. *J. Financ. Quant. Anal.* 21, 335–341.
- Srivastava, M.S., 2005. Some tests concerning the covariance matrix in high dimensional data. *J. Jpn. Statist. Soc. (Nihon Tōkei Gakkai Kaihō)* 35, 251–272.
- Wang, C.D., Mykland, .P.A., 2014. The estimation of leverage effect with high-frequency data. *J. Amer. Statist. Assoc.* 109, 197–215.
- Wang, C., Yang, J., Miao, B., Cao, L., 2013. Identity tests for high dimensional data using RMT. *J. Multivariate Anal.* 118, 128–137.
- Wang, Q., Yao, J., 2013. On the sphericity test with large-dimensional observations. *Electron. J. Statist.* 7, 2164–2192.
- Zheng, X., Li, Y., 2011. On the estimation of integrated covariance matrices of high dimensional diffusion processes. *Ann. of Statist.* 39, 3121–3151.