Tests of Zero Correlation Using Modified RV Coefficient for High-Dimensional Vectors

M. Rauf Ahmad¹

Published online: 8 May 2019
© The Author(s) 2019

Abstract

Tests of zero correlation between two or more vectors with large dimension, possibly larger than the sample size, are considered when the data may not necessarily follow a normal distribution. A single-sample case for several vectors is first proposed, which is then extended to the common covariance matrix under the assumption of homogeneity across several independent populations. The test statistics are constructed using a recently proposed modification of the RV coefficient (a correlation coefficient for vector-valued random variables) for high-dimensional vectors. The accuracy of the tests is shown through simulations.

Keywords Block-diagonal structure · Cross-correlations · High-dimensional inference

1 Introduction

Let $\mathbf{X}_k = (X_{k1}, \dots, X_{kp})^T$, $k = 1, \dots, n$, be iid random vectors drawn from a population with $E(\mathbf{X}_k) = \boldsymbol{\mu} \in \mathbb{R}^p$ and $\text{Cov}(\mathbf{X}_k) = \boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$, where $\boldsymbol{\Sigma} > 0$ can be expressed as a partitioned matrix $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})_{i,j=1}^b$ with blocks $\boldsymbol{\Sigma}_{ij} \in \mathbb{R}^{p_i \times p_j}$, $\boldsymbol{\Sigma}_{ji} = \boldsymbol{\Sigma}_{ij}^T$, $\boldsymbol{\Sigma}_{ii}^T = \boldsymbol{\Sigma}_{ii}$. We are interested to test

$$H_{0b} : \boldsymbol{\Sigma}_{ij} = \mathbf{0} \forall i < j \text{ vs. } H_{1b} : \boldsymbol{\Sigma}_{ij} \neq \mathbf{0} \text{ for at least one pair } i < j \quad (1)$$

when the block dimensions, p_i , may exceed the sample size, n , and the data may not necessarily follow the multivariate normal distribution. Under H_{0b} , $\boldsymbol{\Sigma}$ reduces to a block-diagonal structure, $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_{bb})$, $\boldsymbol{\Sigma}_{ii} \in \mathbb{R}^{p_i \times p_i}$. Obviously, under normality, the test of H_{0b} is equivalent to testing independence of the corresponding vectors.

Now, consider $g \geq 2$ independent populations with $\mathbf{X}_{lk} = (X_{lk1}, \dots, X_{lkp})^T$, $k = 1, \dots, n_l$, as iid random vectors drawn from l th population with $E(\mathbf{X}_{lk}) = \boldsymbol{\mu}_l$,

✉ M. Rauf Ahmad
rauf.ahmad@statistik.uu.se

¹ Department of Statistics, Uppsala University, Uppsala, Sweden

$\text{Cov}(\mathbf{X}_{lk}) = \mathbf{\Sigma}_l > 0$, $l = 1, \dots, g$, defined similarly as for the single population above. Assuming homogeneity of covariance matrices across g populations, i.e., $\mathbf{\Sigma}_l = \mathbf{\Sigma} \forall l$, we are further interested to test hypotheses in (1) for common $\mathbf{\Sigma}$. To distinguish the two cases, we denote multi-sample hypotheses as H_{0g} and H_{1g} , respectively.

Tests of H_{0b} or H_{0g} are frequently needed in multivariate applications, where the special case of $p_i = 1 \forall i = 1, \dots, b$ leads to the well-known single- and multi-sample tests of identity or sphericity hypotheses; see e.g., [1, 2] and the references therein. Tests of H_{0b} or H_{0g} for a classical case, $n > p$, have often been addressed in the multivariate literature, including for multiple blocks; see [8, 16, 20, 21]. [14] give extensions to functional data, and [6] discuss permutation and bootstrap tests. Other recent extensions include nonparametric tests ([10, 13]) and a kernel-based approach ([22]); see also [9, 19] for general theory, in the context of canonical correlation analysis.

A potential advantage of H_{0b} in (1) is a huge dimension reduction, which adds further motivation to testing such hypotheses for high-dimensional data, $p \gg n$. As the classical testing theory, mostly based on likelihood ratio approach, collapses in this case, several modifications have recently been introduced to cope with the problem; see e.g., [26, 33, 34], where [32] provide a test based on rank correlations. Most of these tests assume normality and are constructed using a Frobenius norm measuring the distance between the null and alternative hypotheses. Use of a Frobenius norm as a distance measure is a common approach for estimation and testing of high-dimensional covariance matrices; for its use in other contexts, see e.g., [3] and the references cited therein.

Our construction of the test statistics for (1) is, however, based on a different approach. We use the RV coefficient (see Sect. 2) as the basic measure of orthogonality of vectors and extend it to test H_{0b} and H_{0g} . [5] recently proposed this modification to the RV coefficient for two high-dimensional vectors and used it to introduce a test of orthogonality. For details, including historical overview, comparison with its competitors and extensive bibliography, see the reference.

Apart from theoretical analysis of the modified coefficient and its competitors, extensive simulations are used to show the accuracy of the modified estimator and the significance test. The modified coefficient is composed of computationally very efficient estimators as simple functions of the empirical covariance matrix. In this paper, we first extend the theory to propose a test of no correlation between $b \geq 2$ large vectors, possibly of different dimension.

We further extend this one-sample multi-block test to a multi-sample case with g independent populations. Assuming homogeneity of covariance matrices across g populations, we use pooled information to construct a test for the common covariance matrix. A particularly attractive aspect of the multi-sample test is its simplicity of pooling information across g independent samples, in such a way that the single-sample computations conveniently carry over to the multi-sample case, resulting in a highly efficient test statistic.

We begin in the next section with the basic notational set up, where we also briefly recap the modified RV coefficient in [5] for further use and reference. We keep it as short as possible, referring for details to the original paper. Using these rudiments, we propose a test for H_{0b} for multiple blocks in Sect. 3, with an extension

to the test of H_{0g} in Sect. 4. Accuracy of the tests through simulations is shown in Sect. 5, where technical proofs are deferred to Appendix.

2 Notations and Preliminaries

Let the random vectors $\mathbf{X}_k \in \mathbb{R}^p$, $k = 1, \dots, n$, with $E(\mathbf{X}_k) = \boldsymbol{\mu} \in \mathbb{R}^p$, $\text{Cov}(\mathbf{X}_k) = \boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$, be partitioned as $\mathbf{X}_k = (\mathbf{X}_{1k}^T, \dots, \mathbf{X}_{bk}^T)^T$, $\mathbf{X}_{ik} \in \mathbb{R}^{p_i}$ so that accordingly $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^T, \dots, \boldsymbol{\mu}_b^T)^T$, $\boldsymbol{\mu}_i \in \mathbb{R}^{p_i}$, $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})_{i,j=1}^b$, $\boldsymbol{\Sigma}_{ij} = E(\mathbf{X}_{ik} - \boldsymbol{\mu}_i) \otimes (\mathbf{X}_{jk} - \boldsymbol{\mu}_j)^T \in \mathbb{R}^{p_i \times p_j}$, $i \neq j$, $\boldsymbol{\Sigma}_{ji} = \boldsymbol{\Sigma}_{ij}^T$, $\boldsymbol{\Sigma}_{ii} = \boldsymbol{\Sigma}_{ii}$, where $\mathbb{R}^{a \times b}$ is the space of real matrices and $\otimes$ is the Kronecker product. We assume $\boldsymbol{\Sigma} > 0$, $\boldsymbol{\Sigma}_{ii} > 0$ $\forall i$.

Let $\bar{\mathbf{X}} = \sum_{k=1}^n \mathbf{X}_{ik}/n$ and $\hat{\boldsymbol{\Sigma}} = \sum_{k=1}^n (\tilde{\mathbf{X}}_k \otimes \tilde{\mathbf{X}}_k)/(n-1)$ be unbiased estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, likewise partitioned as $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_b)^T$, $\hat{\boldsymbol{\Sigma}} = (\hat{\boldsymbol{\Sigma}}_{ij})_{i,j=1}^b$, so that $\bar{\mathbf{X}}_i = \sum_{k=1}^n \mathbf{X}_{ik}/n$ and $\hat{\boldsymbol{\Sigma}}_{ij} = \sum_{k=1}^n (\tilde{\mathbf{X}}_{ik} \otimes \tilde{\mathbf{X}}_{jk})/(n-1)$ are unbiased estimators of $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_{ij}$, where $\tilde{\mathbf{X}}_k = \mathbf{X}_k - \bar{\mathbf{X}}$, $\tilde{\mathbf{X}}_{ik} = \mathbf{X}_{ik} - \bar{\mathbf{X}}_i$. Denoting $\mathbf{X} = (\mathbf{X}_1^T \dots \mathbf{X}_n^T)^T \in \mathbb{R}^{n \times p}$ as the entire data matrix and $\mathbf{X}_i = (\mathbf{X}_{i1}^T, \dots, \mathbf{X}_{in}^T)^T \in \mathbb{R}^{n \times p_i}$ as the data matrix for i th block, we can express the estimators more succinctly as

$$\bar{\mathbf{X}} = \frac{1}{n} \mathbf{X}^T \mathbf{1}, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n-1} \mathbf{X}^T \mathbf{C} \mathbf{X}, \quad \bar{\mathbf{X}}_i = \frac{1}{n} \mathbf{X}_i^T \mathbf{1}, \quad \hat{\boldsymbol{\Sigma}}_{ij} = \frac{1}{n-1} \mathbf{X}_i^T \mathbf{C} \mathbf{X}_j, \quad (2)$$

where $\mathbf{C} = \mathbf{I} - \mathbf{J}/n$ is the $n \times n$ centering matrix, $\mathbf{I}$ is the identity matrix, $\mathbf{J} = \mathbf{1}\mathbf{1}^T$ with $\mathbf{1}$ a vector of 1s. The null hypotheses in (1) thus imply the nullity of all off-diagonal blocks $\boldsymbol{\Sigma}_{ij}$ in $\boldsymbol{\Sigma}$. Consider first the simplest case of $b = 2$, with only one off-diagonal block $\boldsymbol{\Sigma}_{12} = \text{Cov}(\mathbf{X}_{1k}, \mathbf{X}_{2k})$. Obviously, a test of zero correlation (or, under normality, of independence) between $\mathbf{X}_{1k}$ and $\mathbf{X}_{2k}$ can be based on an empirical estimator of $\|\boldsymbol{\Sigma}_{12}\|^2 = \text{tr}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21})$ or an appropriately normed version of it. One such normed measure is proposed in [5] as

$$\hat{\eta} = \frac{\widehat{\|\boldsymbol{\Sigma}_{12}\|^2}}{\sqrt{\widehat{\|\boldsymbol{\Sigma}_{11}\|^2} \widehat{\|\boldsymbol{\Sigma}_{22}\|^2}}}, \quad (3)$$

where $\widehat{\|\boldsymbol{\Sigma}_{ij}\|^2}$ is an unbiased (and consistent) estimator of $\|\boldsymbol{\Sigma}_{ij}\|^2$, $i, j = 1, 2$; see Sect. 3 for formal definition. The $\hat{\eta}$ is a modified form of the original RV coefficient, $\tilde{\eta} = \widehat{\|\boldsymbol{\Sigma}_{12}\|^2} / \|\hat{\boldsymbol{\Sigma}}_{11}\| \|\hat{\boldsymbol{\Sigma}}_{22}\|$, as an extension of the scalar correlation coefficient, to measure correlation between vectors of possibly different dimension. Note that $\hat{\eta}$ is constructed using unbiased estimators in the true RV coefficient $\eta = \|\boldsymbol{\Sigma}_{12}\|^2 / \|\boldsymbol{\Sigma}_{11}\| \|\boldsymbol{\Sigma}_{22}\|$. The RV coefficient extends the usual scalar correlation coefficient to data vectors of possibly different dimension, is often used to study relationship between different data configurations, and has many attractive properties; see also [25].

Based on the aforementioned modification and subsequent test of zero correlation, we here extend the same concept and construct tests of hypotheses in (1) and their multi-sample variants. The proposed tests will be valid for the following general multivariate model which helps us relax the normality assumption.

Define $\mathbf{Y}_{ik} = \mathbf{X}_{ik} - \boldsymbol{\mu}_i$, $i = 1, \dots, b$, so that $\mathbf{Y}_k = (\mathbf{Y}_{1k}^T, \dots, \mathbf{Y}_{bk}^T)^T$ can replace $\mathbf{X}_k$. We define the multivariate model as

$$\mathbf{Y}_k = \boldsymbol{\Sigma}^{1/2} \mathbf{Z}_k, \quad (4)$$

with $\mathbf{Z}_k \sim \mathcal{F}$, $E(\mathbf{Z}_k) = \mathbf{0}_p$, $\text{Cov}(\mathbf{Z}_k) = \mathbf{I}_p$, $k = 1, \dots, n$, where $\mathcal{F}$ denotes a p -dimensional distribution function having fourth moment finite; see the assumptions below. Model (4) covers a wide class of distributions such as the elliptical class including multivariate normal. Further, it will help us enhance the validity of the proposed test to a variety of applications under fairly general conditions.

3 One-Sample Test with Multiple Blocks

Consider any two data matrices $\mathbf{X}_i = (\mathbf{X}_{i1}^T, \dots, \mathbf{X}_{in}^T)^T \in \mathbb{R}^{n \times p_i}$ and $\mathbf{X}_j = (\mathbf{X}_{j1}^T, \dots, \mathbf{X}_{jn}^T)^T \in \mathbb{R}^{n \times p_j}$ in $\mathbf{X} \in \mathbb{R}^{n \times p}$, where $\mathbf{X}_{ik} \in \mathbb{R}^{p_i}$, $\mathbf{X}_{jk} \in \mathbb{R}^{p_j}$, $k = 1, \dots, n$, $i, j = 1, \dots, b$, $i < j$. Following $\hat{\eta}$ in Eq. (3), the RV coefficient of correlation between $\mathbf{X}_{ik}$ and $\mathbf{X}_{jk}$ can be computed as

$$\hat{\eta}_{ij} = \frac{\widehat{\|\boldsymbol{\Sigma}_{ij}\|^2}}{\sqrt{\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2} \widehat{\|\boldsymbol{\Sigma}_{jj}\|^2}}}, \quad (5)$$

where $\hat{\eta}_{ij}$ estimates $\eta_{ij} = \|\boldsymbol{\Sigma}_{ij}\|^2 / \|\boldsymbol{\Sigma}_{ii}\| \|\boldsymbol{\Sigma}_{jj}\|$. Since $\boldsymbol{\Sigma}_{ii} > 0$, $\eta_{ij} = 0 \Leftrightarrow \|\boldsymbol{\Sigma}_{ij}\|^2 = 0 \forall i > j$, which implies that a test for $\eta_{ij} = 0$ can be equivalently based on $\boldsymbol{\Sigma}_{ij}$. In fact, as will be shown shortly, the denominator of $\hat{\eta}_{ij}$ adjusts itself through $\text{Var}(\widehat{\|\boldsymbol{\Sigma}_{ij}\|^2})$ so that it suffices to use $\widehat{\|\boldsymbol{\Sigma}_{ij}\|^2}$ to construct a test statistic. We begin by defining the estimators used to compose $\hat{\eta}_{ij}$ and their moments, which will also be useful in subsequent computations. For notational convenience, denote $\boldsymbol{\Sigma}_{ii}^{1/2} = \boldsymbol{\Delta}_{ii}$ so that $\|\boldsymbol{\Delta}_{ii}\|^2 = \text{tr}(\boldsymbol{\Sigma}_{ii})$ and correspondingly $\|\hat{\boldsymbol{\Delta}}_{ii}\|^2 = \text{tr}(\hat{\boldsymbol{\Sigma}}_{ii})$. For the proof of the following theorem, see ‘‘Appendix 1.2.1’’.

Theorem 1 *The unbiased estimators of $\|\boldsymbol{\Sigma}_{ii}\|^2$, $\|\boldsymbol{\Sigma}_{ij}\|^2$, and $\|\boldsymbol{\Delta}_{ii}\| \|\boldsymbol{\Delta}_{jj}\|$ are defined as*

$$\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2} = \nu \left[(n-1)(n-2) \|\hat{\boldsymbol{\Sigma}}_{ii}\|^2 + \{\|\hat{\boldsymbol{\Delta}}_{ii}\|^2\}^2 - nQ_{ii} \right] \quad (6)$$

$$\widehat{\|\boldsymbol{\Sigma}_{ij}\|^2} = \nu \left[(n-1)(n-2) \|\hat{\boldsymbol{\Sigma}}_{ij}\|^2 + \|\hat{\boldsymbol{\Delta}}_{ii}\|^2 \|\hat{\boldsymbol{\Delta}}_{jj}\|^2 - nQ_{ij} \right] \quad (7)$$

$$[\widehat{\|\boldsymbol{\Delta}_{ii}\|^2}]^2 = \nu \left[2 \|\hat{\boldsymbol{\Sigma}}_{ii}\|^2 + (n^2 - 3n + 1) \|\hat{\boldsymbol{\Delta}}_{ii}\|^2 - nQ_{ii} \right], \quad (8)$$

$$\|\hat{\boldsymbol{\Delta}}_{ii}\|^2 \|\hat{\boldsymbol{\Delta}}_{jj}\|^2 = \nu \left[2 \|\hat{\boldsymbol{\Sigma}}_{ij}\|^2 + (n^2 - 3n + 1) \|\hat{\boldsymbol{\Delta}}_{ii}\|^2 \|\hat{\boldsymbol{\Delta}}_{jj}\|^2 - nQ_{ij} \right], \quad (9)$$

where $v = (n-1)/n(n-2)(n-3)$, $Q_{ij} = \sum_{k=1}^n \tilde{z}_{ik} \tilde{z}_{jk} / (n-1)$ and $Q_{ii} = \sum_{k=1}^n \tilde{z}_{ik}^2 / (n-1)$ with $\tilde{z}_{ik} = \tilde{\mathbf{X}}_{ik}^T \tilde{\mathbf{X}}_{ik}$ and $\tilde{\mathbf{X}}_{ik} = \mathbf{X}_{ik} - \mathbf{X}_i$, $i, j = 1, \dots, b$, $i < j$.

Note that the terms Q_{ij} are needed to make the estimators, and hence the subsequent test statistics, valid under Model (4) beyond the normality assumption. Essentially, these terms involve fourth-order elements of $\mathbf{X}_{ik}$ due to the variances of bilinear forms to be computed. This in turn calls for bounding such moments of $\mathbf{X}_{ik}$ to be finite (see assumptions below). For this, define

$$\kappa_{ij} = E(z_{ik} z_{jk}) - 2\|\Sigma_{ij}\|^2 - \|\Delta_{ii}\|^2 \|\Delta_{jj}\|^2 \Rightarrow \kappa_{ii} = E(z_{ik}^2) - 2\|\Sigma_{ii}\|^2 - [\|\Delta_{ii}\|^2]^2 \quad (10)$$

with $z_{ik} = \mathbf{Y}_{ik}^T \mathbf{Y}_{ik}$. As $\kappa_{ij} = 0$ under normality, it serves as a measure of non-normality and, given finite fourth moment, helps extend the results to a wide class of distributions under Model (4). The results of Theorem 1 also help obtain unbiased estimators of κ_{ij} and κ_{ii} as (see ‘‘Appendix 1.2.1’’)

$$\hat{\kappa}_{ij} = -\frac{nv}{n-1} \left[2(n-1)^2 \|\hat{\Sigma}_{ij}\|^2 + (n-1)^2 \|\hat{\Delta}_{ii}\|^2 \|\hat{\Delta}_{jj}\|^2 - n(n+1)Q_{ij} \right] \quad (11)$$

$$\hat{\kappa}_{ii} = -\frac{nv}{n-1} \left[2(n-1)^2 \|\hat{\Sigma}_{ii}\|^2 + (n-1)^2 [\|\hat{\Delta}_{ii}\|^2]^2 - n(n+1)Q_{ii} \right]. \quad (12)$$

The estimators in Theorem 1 are composed of $\hat{\Sigma}_{ij}$ (see Eq. 2) and Q_{ij} , both of which are simple functions of mean-deviated vectors $\tilde{\mathbf{X}}_{ik}$. It makes the estimators very simple and computationally highly efficient for practical use. For mathematical amenability, however, an alternative form of the same estimators in terms of U -statistics helps us study their properties due to the attractive projection and asymptotic theory of U -statistics.

Given $\{\mathbf{X}_{ik}, \mathbf{X}_{ir}, \mathbf{X}_{il}, \mathbf{X}_{is}\}$, $k \neq r \neq l \neq s$, from $\mathbf{X}_i \in \mathbb{R}^{n \times p_i}$, define $\mathbf{D}_{ikr} = \mathbf{X}_{ik} - \mathbf{X}_{ir}$ with $E(\mathbf{D}_{ikr}) = \mathbf{0}$, $\text{Cov}(\mathbf{D}_{ikr}) = 2\Sigma_{ii}$, $E(\mathbf{D}_{ikr} \mathbf{D}_{jkr}^T) = 2\Sigma_{ij}$. For $A_{ikr} = \mathbf{D}_{ikr}^T \mathbf{D}_{ikr}$, $E(A_{ikr}) = 2\|\Delta_{ii}\|^2$ and $E(A_{iks} A_{jlr}) = 4\|\Delta_{ii}\|^2 \|\Delta_{jj}\|^2$. With $\mathbf{D}_{ils}$ defined similarly, let $A_{ijlskr} = A_{ilskr} A_{jkr}$ where $A_{ilskr} = \mathbf{D}_{ils}^T \mathbf{D}_{ikr}$, so that $E(A_{ilskr}^2) = 4\|\Sigma_{ij}\|^2$, $E(A_{ijlskr}) = 4\|\Sigma_{ij}\|^2$. Denoting further $B_{iikrsl} = A_{ilskr}^2 + A_{ilksr}^2 + A_{ilrsk}^2$, $C_{ijkrls} = A_{ijlskr} + A_{ijlksr} + A_{ijlrks}$, $D_{ijkrls} = A_{ikr} A_{jls} + A_{ikl} A_{jrs} + A_{iks} A_{jlr}$ and $P(n) = n(n-1)(n-2)(n-3)$, the U -statistics versions of the estimators of $\|\Sigma_{ii}\|^2$, $\|\Sigma_{ij}\|^2$ and $\|\Delta_{ii}\|^2 \|\Delta_{jj}\|^2$ in Theorem 2, denoted U_1, U_2, U_3 , are defined, respectively, as

$$U_1 = \frac{1}{12P(n)} \sum_{\substack{k,r,l,s \\ \pi(\cdot)}}^n B_{iikrsl}, \quad U_2 = \frac{1}{12P(n)} \sum_{\substack{k,r,l,s \\ \pi(\cdot)}}^n C_{ijkrls}, \quad U_3 = \frac{1}{12P(n)} \sum_{\substack{k,r,l,s \\ \pi(\cdot)}}^n D_{ijkrls}, \quad (13)$$

where the sum is quadruple over $\{k, l, r, s\}$ and $\pi(\cdot) = \pi(k, r, l, s)$ implies $k \neq r \neq l \neq s$. The moments in the following theorem follow conveniently using the alternative form of estimators in (13) and will be very useful for further computations in the sequel.

Theorem 2 For the unbiased estimators $\widehat{\|\Sigma_{ij}\|^2}$ and $\widehat{\|\Sigma_{ii}\|^2}$ in Theorem 2, we have

$$\begin{aligned} \text{Var}(\widehat{\|\Sigma_{ij}\|^2}) = & \frac{2}{P(n)} \left[4a(n)\|\Sigma_{ij}\Sigma_{ji}\|^2 + 2b(n)\text{tr}(\Sigma_{ii}\Sigma_{ij}\Sigma_{jj}\Sigma_{ji}) \right. \\ & \left. + c(n)\left\{ \left[\|\Sigma_{ij}\|^2\right]^2 + \|\Sigma_{ii}\|^2\|\Sigma_{jj}\|^2 + KO(1) \right\} \right] \end{aligned} \quad (14)$$

$$\text{Var}(\widehat{\|\Sigma_{ii}\|^2}) = \frac{4}{P(n)} \left[d(n)\|\Sigma_{ii}^2\|^2 + e(n)\left[\|\Sigma_{ii}^2\|^2 + KO(n^2) \right] \right] \quad (15)$$

where $a(n) = 3n^3 - 24n^2 + 44n + 20$, $b(n) = 6n^3 - 40n^2 + 22n + 181$, $c(n) = 2n^2 - 12n + 21$, $d(n) = 2n^3 - 9n^2 + 9n - 16$, $e(n) = n^2 - 3n + 8$ and K contains terms involving κ_{ij} .

We skip the proof of Theorem 2 which follows from that of Theorem 2 in [5]. The terms $KO(\cdot)$ sum up constants that eventually vanish under the assumptions. In particular, K involves κ_{ij} in (10) and terms involving Hadamard products such as $\Sigma_{ii} \odot \Sigma_{jj}$ which converge to zero. Now, we have the required tools to proceed with the test of H_{0b} in (1). As mentioned above,

$$H_{0b} : \Sigma_{ij} = \mathbf{O} \Leftrightarrow H_{0b} : \eta_{ij} = 0 \quad \forall i < j, \quad i, j = 1, \dots, b. \quad (16)$$

Moreover, $\eta_{ij} = 0 \Leftrightarrow \sum_{i < j}^b \eta_{ij} = 0$, so that a test of H_{0b} can be based on a sum of Frobenius norms over all off-diagonal blocks, i.e., $\sum_{i < j} \|\Sigma_{ij}\|^2$. We thus define the test statistic for H_{0b} as

$$T_b = \sum_{i=1}^b \sum_{\substack{j=1 \\ i < j}}^b T_{ij}, \quad \text{where} \quad T_{ij} = \frac{1}{p_i p_j} \widehat{\|\Sigma_{ij}\|^2}. \quad (17)$$

Here, T_{ij} is a statistic to test $H_{0ij} : \eta_{ij} = 0$ for any single off-diagonal block Σ_{ij} . Moreover, the scaling factor $p_i p_j$ will help us obtain the limit of T_b for $p_i \rightarrow \infty$ along with $n \rightarrow \infty$, under the following assumptions. Recall $\mathbf{Y}_k \in \mathbb{R}^p$ in Model (4).

Assumption 3 $E(Y_{ks}^4) = \gamma_s \leq \gamma < \infty, \forall s = 1, \dots, p$.

Assumption 4 $\lim_{p_i \rightarrow \infty} \frac{\|\Delta_{ii}\|^2}{p_i} = O(1), \forall i = 1, \dots, b$.

Assumption 5 $\lim_{n, p_i \rightarrow \infty} \frac{p_i}{n} \rightarrow \zeta_0 \in (0, \infty), \forall i = 1, \dots, b$.

Assumption 3 bounds fourth moment of $\mathbf{Y}_k$ so that, by Cauchy–Schwarz inequality, $E(y_{k1s}^2 y_{k2s}^2) < \infty$ which implies that $\kappa_{ij} / \|\Sigma_{ii}\|^2 \|\Sigma_{jj}\|^2 \rightarrow 0$. This conforms to the definition of κ_{ij} and helps us present the test under Model (4), and it also makes the terms involving K vanish in Theorem 2. Assumption 4 bounds the average of the eigenvalues of diagonal blocks. It is a mild assumption, often used in high-dimensional testing, and as its consequence, $\|\Sigma_{ij}\|^2 / p_i p_j = O(1)$.

Whereas Assumption 4 is needed for limits under H_{0b} and H_{1b} , its consequence is only needed under H_{1b} since it neither holds nor is required under H_{0b} . Assumption 5 controls joint growth of n and p_i so that the limit holds under a high-dimensional framework, and it is also needed only under H_{1b} . Now, for T_b , we have $E(T_b) = \sum_{i < j}^b \|\Sigma_{ij}\|^2 / p_i p_j = 0$ under H_{0b} and

$$\begin{aligned} \text{Var}(T_b) = & \sum_{i=1}^b \sum_{j=1}^b \text{Var}(T_{ij}) + 2 \sum_{i=1}^b \sum_{j=1}^b \sum_{j'=1}^b \text{Cov}(T_{ij}, T_{ij'}) \\ & + 2 \sum_{i=1}^b \sum_{i'=1}^b \sum_{j=1}^b \text{Cov}(T_{ij}, T_{i'j}) + \sum_{i=1}^b \sum_{i'=1}^b \sum_{j=1}^b \sum_{j'=1}^b \text{Cov}(T_{ij}, T_{i'j'}) . \end{aligned} \quad (18)$$

Equation (18) gives $\text{Var}(T_{ij})$ and the covariance, say C_1, C_2, C_3 , respectively, given as below.

$$\begin{aligned} C_1 = & \frac{4}{P(n)p_i^2 p_j p_{j'}} \left[2a_1(n) \text{tr}(\Sigma_{ij} \Sigma_{ji} \Sigma_{ij'} \Sigma_{j'i}) + b_1(n) \text{tr}(\Sigma_{ii} \Sigma_{ij} \Sigma_{jj'} \Sigma_{j'i}) \right. \\ & \left. + (n-4) \{ 2\|\Sigma_{ij}\|^2 \|\Sigma_{ij'}\|^2 + 5\|\Sigma_{jj'}\|^2 \|\Sigma_{ii}\|^2 \} \right] \end{aligned} \quad (19)$$

$$\begin{aligned} C_2 = & \frac{4}{P(n)p_i p_j p_{j'}} \left[2a_1(n) \text{tr}(\Sigma_{ii'} \Sigma_{i'i} \Sigma_{ij} \Sigma_{ji}) + b_1(n) \text{tr}(\Sigma_{ii} \Sigma_{ii'} \Sigma_{i'j} \Sigma_{ji}) \right. \\ & \left. + (n-4) \{ 2\|\Sigma_{ij}\|^2 \|\Sigma_{i'j}\|^2 + 5\|\Sigma_{ii'}\|^2 \|\Sigma_{jj}\|^2 \} \right] \end{aligned} \quad (20)$$

$$\begin{aligned} C_3 = & \frac{4}{P(n)p_i p_j p_{j'}} \left[2a_2(n) \text{tr}(\Sigma_{ij} \Sigma_{ji'} \Sigma_{i'j'} \Sigma_{j'i}) + b_2(n) \text{tr}(\Sigma_{ij} \Sigma_{jj'} \Sigma_{j'i'} \Sigma_{i'i}) \right. \\ & + (4n-11) \text{tr}(\Sigma_{ii'} \Sigma_{i'j} \Sigma_{jj'} \Sigma_{j'i}) + 3(n-3) \|\Sigma_{i'j}\|^2 \|\Sigma_{ij'}\|^2 \\ & \left. + (3n-10) \|\Sigma_{ii'}\|^2 \|\Sigma_{jj'}\|^2 \right] \end{aligned} \quad (21)$$

$a_1(n) = 3n^3 - 38n^2 + 170n - 262$, $b_1(n) = 6n^3 - 57n^2 + 178n - 182$,
 $a_2(n) = 3n^3 - 39n^2 + 176n - 269$, $b_2(n) = 6n^3 - 70n^2 + 286n - 199$. Under H_{0b} , the covariances vanish and the variance reduces to

$$\text{Var}(T_b) = \frac{2(2n^2 - 12n + 21)}{P(n)} \sum_{i=1}^b \sum_{j=1}^b \frac{1}{p_i^2 p_j^2} \|\Sigma_{ii}\|^2 \|\Sigma_{jj}\|^2 = O\left(\frac{1}{n^2}\right) \sum_{i=1}^b \sum_{j=1}^b v_{ij}, \quad (22)$$

where $V_{ij} = \|\Sigma_{ii}\|^2 \|\Sigma_{jj}\|^2 / p_i^2 p_j^2$. It implies that $\text{Var}(nT_b)$ is bounded and a non-degenerate limit of nT_b may exist. That this indeed holds under the assumptions follow from the limit of T_{ij} given in Ahmad ([5], Theorem 3), by noting that the covariance terms above essentially vanish under the same assumptions. The following theorem summarizes the result, the proof of which, along with that of Theorem 12 for the multi-sample case, is given in “Appendix 1.2.3”.

Theorem 6 *Given D , T_b in (17) and Assumptions 3–5. Then, $(T_b - E(T_b))/\sigma_{T_b} \rightarrow N(0, 1)$, as $n, p_i \rightarrow \infty$, with $\sigma_{T_b}^2 = \text{Var}(T_b)$ in (18). Under H_0 and Assumptions 3–4, $nT_b/\sigma_{T_b} \rightarrow N(0, 1)$ with $\sigma_{T_b}^2$ as in Eq. (22). Further, the same limits hold when $\text{Var}(T_b)$ is replaced with $\widehat{\text{Var}}(T_b)$.*

The last part of Theorem 6 makes the statistic T_b applicable, when a consistent estimator, i.e.,

$$\widehat{\text{Var}}(nT_b) = \sum_{i=1}^b \sum_{\substack{j=1 \\ i < j}}^b \widehat{V}_{ij}, \quad (23)$$

is plugged in for the true variance, where $\widehat{V}_{ij} = \|\widehat{\Sigma}_{ii}\|^2 \|\widehat{\Sigma}_{jj}\|^2 / p_i^2 p_j^2$. From the proof of the theorem, we also note that the moments and the limit of T_b are functions of $\eta_{ij} = \|\Sigma_{ij}\|^2 / \|\Sigma_{ii}\| \|\Sigma_{jj}\|$. For the power of T_b , let $\beta(\theta)$ be the power function with $\theta \in \Theta$, where

$$\theta_0 = \{\Sigma_{ii}, i \in \{1, \dots, b\}\}, \quad \Theta_1 = \{\Sigma_{ii}, \Sigma_{ij}, i, j \in \{1, \dots, b\}, i < j\}$$

are the parameter subspaces under H_{0b} and H_{1b} , respectively. Let z_α denote $100\alpha\%$ quantile of $N(0, 1)$ distribution, so that $P(T_b/\sigma_{T_b} \geq z_\alpha) = \alpha$, by Theorem 6, where $\sigma_{T_b}^2$ denotes $\text{Var}(T_b)$ under H_{0b} . Also, let $\gamma = \sigma_{T_{b0}}/\sigma_{T_b}$, $\delta = E(T_b)/\sigma_{T_b}$ with $E(T_b) = \sum_{i < j} \|\Sigma_{ij}\|^2 / p_i p_j$, $\sigma_{T_b}^2 = \text{Var}(T_b)$. Then, $\beta(\theta|H_1) = P(T_b/\sigma_{T_b} \leq \gamma z_\alpha - \delta) = 1 - \Phi(\gamma z_\alpha - \delta)$ defines the power of T_b , where $\Phi(\cdot)$ is the distribution function of $N(0, 1)$. From Theorem 2, $\gamma = O(1)$ and $\delta = O(n)$ under the assumptions, so that $1 - \Phi(\gamma z_\alpha - \delta) \rightarrow 1$ as $n, p_i \rightarrow \infty$. Finally, letting $\Sigma_{ij} = \mathbf{A}_{ij}/\sqrt{n}$ for a fixed matrix $\mathbf{A}_{ij}$, $i, j = 1, \dots, b$, $i < j$, similar arguments imply that the local power also converges to 1 as $n, p_i \rightarrow \infty$.

4 Multi-Sample Extension with Multiple Blocks

Consider the multi-sample set up given after (1) with $\mathbf{X}_{lk} = (\mathbf{X}_{l1k}^T, \dots, \mathbf{X}_{lbk}^T)^T \in \mathbb{R}^p$, $\mathbf{X}_{lik} \in \mathbb{R}^{p_i}$, $k = 1, \dots, n_l$, as iid vectors from population l , where $E(\mathbf{X}_{lk}) = \boldsymbol{\mu}_l = (\boldsymbol{\mu}_{l1}^T, \dots, \boldsymbol{\mu}_{lb}^T)^T \in \mathbb{R}^p$, $\boldsymbol{\mu}_{li} \in \mathbb{R}^{p_i}$, $\text{Cov}(\mathbf{X}_{lk}) = \Sigma_l = (\Sigma_{lij})_{i,j=1}^b \in \mathbb{R}^{p \times p}$, $l = 1, \dots, g$, $\Sigma_{lij} = E(\mathbf{X}_{lik} - \boldsymbol{\mu}_{li})(\mathbf{X}_{ljk} - \boldsymbol{\mu}_{lj})^T \in \mathbb{R}^{p_i \times p_j}$, $i < j$, $\Sigma_{lji} = \Sigma_{lij}^T$, $\Sigma_{lji} = \Sigma_{lji}^T$. Then $\bar{\mathbf{X}}_l = \sum_{k=1}^{n_l} \mathbf{X}_{lk}/n_l$, $\bar{\Sigma}_l = \sum_{k=1}^{n_l} (\tilde{\mathbf{X}}_{lk} \otimes \tilde{\mathbf{X}}_{lk})/(n_l - 1)$ are unbiased estimators of $\boldsymbol{\mu}_l$, Σ_l , correspondingly partitioned as $\bar{\mathbf{X}}_l = (\bar{\mathbf{X}}_{l1}, \dots, \bar{\mathbf{X}}_{lb})^T$, $\bar{\Sigma}_l = (\hat{\Sigma}_{lij})_{i,j=1}^b$ with

$\bar{\mathbf{X}}_{li} = \sum_{k=1}^{n_l} \tilde{\mathbf{X}}_{lik}/n_l$, $\hat{\Sigma}_{lij} = \sum_{k=1}^{n_l} (\tilde{\mathbf{X}}_{lik} \otimes \tilde{\mathbf{X}}_{ljk})/(n_l - 1)$ as unbiased estimators of μ_{li} and Σ_{lij} , where $\tilde{\mathbf{X}}_k = \mathbf{X}_k - \bar{\mathbf{X}}$ and $\tilde{\mathbf{X}}_{lik} = \mathbf{X}_{lik} - \bar{\mathbf{X}}_{li}$. Denoting $\mathbf{X}_l = (\mathbf{X}_{l1}^T \dots \mathbf{X}_{ln_l}^T)^T \in \mathbb{R}^{n_l \times p}$ as the entire data matrix for l th sample, with $\mathbf{X}_{li} = (\mathbf{X}_{li1}^T, \dots, \mathbf{X}_{lin_l}^T)^T \in \mathbb{R}^{n_l \times p_i}$ as i th data matrix in $\mathbf{X}_l$, we can rewrite the estimators, using $\mathbf{C}_l = \mathbf{I}_{n_l} - \mathbf{J}_{n_l}/n_l$, $l = 1, \dots, g$, as

$$\bar{\mathbf{X}}_l = \frac{1}{n_l} \mathbf{X}_l^T \mathbf{1}_{n_l}, \quad \hat{\Sigma}_l = \frac{1}{n_l - 1} \mathbf{X}_l^T \mathbf{C}_l \mathbf{X}_l, \quad \bar{\mathbf{X}}_{li} = \frac{1}{n_l} \mathbf{X}_{li}^T \mathbf{1}_{n_l}, \quad \hat{\Sigma}_{lij} = \frac{1}{n_l - 1} \mathbf{X}_{li}^T \mathbf{C}_l \mathbf{X}_{lj}. \quad (24)$$

The RV coefficient in Eq. (3) can now be computed from l th sample, $l = 1, \dots, g$, as

$$\hat{\eta}_{lij} = \frac{\|\hat{\Sigma}_{lij}\|^2}{\sqrt{\|\hat{\Sigma}_{lii}\|^2 \|\hat{\Sigma}_{ljj}\|^2}} \quad (25)$$

which estimates $\eta_{lij} = \|\Sigma_{lij}\|^2 / \|\Sigma_{lii}\| \|\Sigma_{ljj}\|$. We are interested to test hypotheses in (1) for common Σ , assuming $\Sigma_l = \Sigma \forall l$, i.e., to test if Σ is block diagonal, $\Sigma = \text{diag}(\Sigma_{11}, \dots, \Sigma_{bb})$. Formally,

$$H_{0g} : \Sigma_{ij} = \mathbf{0} \forall i \neq j \text{ vs. } H_{1g} : \Sigma_{ij} \neq \mathbf{0} \text{ for at least one pair } i \neq j, \quad (26)$$

where the subscript g refers to g populations. A test of H_{0g} can thus be constructed by pooling information from g populations. For this, we state Assumptions 3–5 for the multi-sample case.

Assumption 7 $E(Y_{lks}^4) = \gamma_s \leq \gamma < \infty$, $\forall s = 1, \dots, p$, $l = 1, \dots, g$.

Assumption 8 $\lim_{p_i \rightarrow \infty} \frac{\|\Delta_{lii}\|^2}{p_i} = O(1)$, $\forall i = 1, \dots, b$, $l = 1, \dots, g$.

Assumption 9 $\lim_{n_l, p_i \rightarrow \infty} \frac{p_i}{n_l} \rightarrow \zeta_l \leq \zeta_0 \in (0, \infty)$, $\forall i = 1, \dots, b$, $l = 1, \dots, g$.

Now, we begin by writing the estimators in Eqs. (6)–(9) for the l th sample as

$$\|\hat{\Sigma}_{lii}\|^2 = v_l \left[(n_l - 1)(n_l - 2) \|\hat{\Sigma}_{lii}\|^2 + \{\|\hat{\Delta}_{lii}\|^2\}^2 - n_l Q_{lii} \right] \quad (27)$$

$$\|\hat{\Sigma}_{lij}\|^2 = v_l \left[(n_l - 1)(n_l - 2) \|\hat{\Sigma}_{lij}\|^2 + \|\hat{\Delta}_{lii}\|^2 \|\hat{\Delta}_{ljj}\|^2 - n_l Q_{lij} \right] \quad (28)$$

$$[\|\hat{\Delta}_{lii}\|^2]^2 = v \left[2 \|\hat{\Sigma}_{ii}\|^2 + (n_l^2 - 3n_l + 1) [\|\hat{\Delta}_{ii}\|^2]^2 - n_l Q_{ii} \right], \quad (29)$$

$$\|\hat{\Delta}_{lii}\|^2 \|\hat{\Delta}_{ljj}\|^2 = v_l \left[2 \|\hat{\Sigma}_{lij}\|^2 + (n_l^2 - 3n_l + 1) \|\hat{\Delta}_{lii}\|^2 \|\hat{\Delta}_{ljj}\|^2 - n_l Q_{lij} \right], \quad (30)$$

where $v_l = (n_l - 1)/n_l(n_l - 2)(n_l - 3)$, $Q_{lij} = \sum_{k=1}^n \tilde{z}_{lik} \tilde{z}_{ljk}/(n_l - 1)$ and $Q_{lii} = \sum_{k=1}^n \tilde{z}_{lik}^2/(n_l - 1)$ with $\tilde{z}_{lik} = \tilde{\mathbf{X}}_{lik}^T \tilde{\mathbf{X}}_{lik}$ and $\tilde{\mathbf{X}}_{lik} = \mathbf{X}_{lik} - \bar{\mathbf{X}}_{li}$, $i, j = 1, \dots, b$, $i < j$.

Note that these are unbiased estimators of $\|\Sigma_{ii}\|^2$, $\|\Sigma_{ij}\|^2$, $[\|\Delta_{ii}\|^2]^2$ and $\|\Delta_{ii}\|\|\Delta_{jj}\|$, respectively, obtained from sample l , so that they can be used to construct pooled estimators of the same parameters, as the following.

Theorem 10 Denote $v_0 = \sum_{l=1}^g v_l^{-1}$ with $v_l = (n_l - 1)/n_l(n_l - 2)(n_l - 3)$. The pooled unbiased estimators of $\|\Sigma_{ii}\|^2$, $\|\Sigma_{ij}\|^2$, $[\|\Delta_{ii}\|^2]^2$ and $\|\Delta_{ii}\|\|\Delta_{jj}\|$ are defined, respectively, as

$$\widehat{\|\Sigma_{p ii}\|^2} = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} \widehat{\|\Sigma_{l ii}\|^2} \quad (31)$$

$$\widehat{\|\Sigma_{p ij}\|^2} = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} \widehat{\|\Sigma_{l ij}\|^2} \quad (32)$$

$$[\widehat{\|\Delta_{p ii}\|^2}]^2 = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} [\widehat{\|\Delta_{l ii}\|^2}]^2, \quad (33)$$

$$\widehat{\|\Delta_{p ii}\|^2} \widehat{\|\Delta_{p jj}\|^2} = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} \widehat{\|\Delta_{l ii}\|^2} \widehat{\|\Delta_{l jj}\|^2}. \quad (34)$$

In pooling information across g samples, Eqs. (31)–(34) use weights $1/v_l$, where the pooled estimators correspond to $\Sigma = (\Sigma_{ij})_{i,j=1}^b$ for which H_{0g} is defined. Thus, an appropriate test statistic for H_{0g} can be defined as

$$T_g = \sum_{i=1}^b \sum_{\substack{j=1 \\ i < j}}^b T_{p ij} \quad \text{where} \quad T_{p ij} = \frac{1}{p_i p_j} \widehat{\|\Sigma_{p ii}\|^2} \quad (35)$$

which extends Eq. (17) for the multi-sample case under homogeneity. Equivalently, we can write

$$T_g = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} T_{lb} \quad \text{with} \quad T_{lb} = \sum_{i=1}^b \sum_{\substack{j=1 \\ i < j}}^b T_{lij} \quad \text{where} \quad T_{lij} = \frac{1}{p_i p_j} \widehat{\|\Sigma_{lij}\|^2}. \quad (36)$$

In this form, T_{lij} , hence T_{lb} , are the same as T_{ij} and T_b in Sect. 3 but now defined for l th population. In either case, the main focus on the formulation of T_g is simplicity, so that, by independence of the g samples, the computations for the one-sample case will mainly suffice for the multi-sample case, as for example the results of the following theorem; see “Appendix 1.2.2” for proof.

Theorem 11 *The pooled estimators $\widehat{\|\Sigma_{pii}\|^2}$ and $\widehat{\|\Sigma_{pij}\|^2}$ in Eqs. (31)–(32) are unbiased for $\|\Sigma_{ii}\|^2$ and $\|\Sigma_{ij}\|^2$, respectively. Further, under Assumptions 7–9, as $n, p_i \rightarrow \infty$,*

$$\text{Var} \left(\sqrt{\nu_0} \frac{\widehat{\|\Sigma_{pii}\|^2}}{\|\Sigma_{ii}\|} \right) = 4[1 + o(1)] \quad (37)$$

$$\text{Var} \left(\sqrt{\nu_0} \frac{\widehat{\|\Sigma_{pij}\|^2}}{\|\Sigma_{ij}\|} \right) = 4 \left(1 + \frac{1}{\eta_{ij}^2} \right) [1 + o(1)], \quad (38)$$

where $\nu_0 = \sum_{l=1}^g \nu_l^{-1}$ and $\eta_{ij} = \|\Sigma_{ij}\|^2 / \|\Sigma_{ii}\| \|\Sigma_{jj}\|$, $i, j = 1, \dots, b$, $i < j$.

Theorem 11 provides bounds on the variance ratios which are more important for our purpose, where the exact variances, which follow from those of single-sample case in Theorem 2, are given in “Appendix 1.2.2”. Moreover, Eq. (38) also implies that $\widehat{\|\Sigma_{pii}\|^2} / p_i^2$ is a consistent estimator of $\|\Sigma_{pii}\|^2 / p_i^2$, $i = 1, \dots, b$, as $n, p_i \rightarrow \infty$. Now, from Theorem 11,

$$\text{E}(\text{T}_g) = \sum_{i=1}^b \sum_{j=1}^b \frac{1}{p_i p_j} \|\Sigma_{ij}\|^2 \quad (39)$$

$$\text{Var}(\text{T}_g) = \frac{1}{\nu_0^2} \sum_{l=1}^g \frac{1}{\nu_l^2} \text{Var}(\text{T}_{lb}), \quad (40)$$

since $\text{Cov}(\text{T}_{lb}, \text{T}_{l'b}) = 0$, $l \neq l'$, by independence, where $\text{Var}(\text{T}_{lb})$ follows directly from Eq. (18) as

$$\begin{aligned} \text{Var}(\text{T}_{lb}) &= \sum_{i=1}^b \sum_{j=1}^b \text{Var}(\text{T}_{lij}) + 2 \sum_{i=1}^b \sum_{j=1}^b \sum_{j'=1}^b \text{Cov}(\text{T}_{lij}, \text{T}_{lij'}) \\ &\quad + 2 \sum_{i=1}^b \sum_{i'=1}^b \sum_{j=1}^b \text{Cov}(\text{T}_{lij}, \text{T}_{li'j}) + \sum_{i=1}^b \sum_{i'=1}^b \sum_{j=1}^b \sum_{j'=1}^b \text{Cov}(\text{T}_{lij}, \text{T}_{li'j'}) \end{aligned} \quad (41)$$

$l = 1, \dots, g$. Equations (19)–(21) provide variances and covariances in $\text{Var}(\text{T}_{lb})$, using the results of Theorem 2 with T_{ij} 's replaced with T_{lij} and n with n_l . We therefore skip unnecessary repetitions. Now, under H_{0g} , $\text{E}(\text{T}_g) = 0$ and

$$\begin{aligned} \text{Var}(\mathbf{T}_g) &= \frac{1}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \frac{2(2n_l^2 - 12n_l + 21)}{P(n_l)} \sum_{i=1}^b \sum_{j=1, i < j}^b \frac{1}{p_i^2 p_j^2} \|\boldsymbol{\Sigma}_{lii}\|^2 \|\boldsymbol{\Sigma}_{lij}\|^2 \\ &= O\left(\frac{1}{v_0}\right) \sum_{i=1}^b \sum_{j=1, i < j}^b v_{lij}, \end{aligned} \quad (42)$$

where $v_0 = \sum_{l=1}^g v_l^{-1}$, $P(n_l) = n_l(n_l - 1)(n_l - 2)(n_l - 3)$ and $v_{lij} = \|\boldsymbol{\Sigma}_{lii}\|^2 \|\boldsymbol{\Sigma}_{lij}\|^2 / p_i^2 p_j^2$. It is obvious from the construction and moments of $\mathbf{T}_g$ that its limit, by the independence of samples, follows along the same lines as that of $\mathbf{T}_b$ without much new computation, and likewise holds for the consistency of $\|\boldsymbol{\Sigma}_{lii}\|^2 / p_i^2$ using that of $\|\boldsymbol{\Sigma}_{ii}\|^2 / p_i^2$, so that, plugged in, they yield $\widehat{\text{Var}}(\mathbf{T}_g)$ as a consistent estimator of $\text{Var}(\mathbf{T}_g)$ as the following, where $\widehat{v}_{pij} = \|\widehat{\boldsymbol{\Sigma}_{pii}}\|^2 \|\widehat{\boldsymbol{\Sigma}_{pij}}\|^2 / p_i^2 p_j^2$:

$$\widehat{\text{Var}}(\sqrt{v_0} \mathbf{T}_g) = \sum_{i=1}^b \sum_{j=1, i < j}^b \widehat{v}_{pij}. \quad (43)$$

The following theorem extends Theorem 6 to the multi-sample case; see ‘‘Appendix 1.2.3’’ for proof.

Theorem 12 For $\mathbf{T}_g$ in (35) and Assumptions 7–9, $(\mathbf{T}_g - \mathbf{E}(\mathbf{T}_g)) / \sigma_{\mathbf{T}_g} \xrightarrow{D} N(0, 1)$, as $n_l, p_i \rightarrow \infty$, with $\sigma_{\mathbf{T}_g}^2 = \text{Var}(\mathbf{T}_g)$ in Eq. (40). Under H_{0g} and Assumptions 7–8, $\sqrt{v_0} \mathbf{T}_g / \sigma_{\mathbf{T}_g} \xrightarrow{D} N(0, 1)$ with $\sigma_{\mathbf{T}_g}^2$ in Eq. (42). Further, the limits hold if $\text{Var}(\mathbf{T}_g)$ is replaced with a consistent estimator $\widehat{\text{Var}}(\mathbf{T}_g)$.

5 Simulations

We evaluate the performance of $\mathbf{T}_b$ and $\mathbf{T}_g$ through simulations. For the one-sample multi-block statistic, $\mathbf{T}_b$, we take $b = 3$ and sample random vectors of sizes $n \in \{20, 50, 100\}$ from multivariate normal, t and chi-square distributions with 10 degrees of freedom each for the latter two, with block dimensions $p_1 \in \{10, 25, 50, 150, 300\}$, $p_2 = 2p_1$, $p_3 = 3p_1$. Two covariance patterns are assumed for the true covariance matrix, i.e., compound symmetry (CS) and AR(1), defined, respectively, as $(1 - \rho)\mathbf{I}_p + \rho\mathbf{J}_p$ and $\boldsymbol{\Sigma} = \mathbf{B}\mathbf{A}\mathbf{B}$ with $\mathbf{A} = \rho^{|k-l|^{1/5}}$ and $\mathbf{B}$ a diagonal matrix with entries square roots of $\rho + \mathbf{p}_1/p$, where $\mathbf{p}_1 = 1 : p$, $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is matrix of 1s. We set $\rho = 0.5$. Under H_{0b} , the same structures are imposed on three diagonal blocks $\boldsymbol{\Sigma}_{ii}$ of dimensions p_i , so that $\boldsymbol{\Sigma} = \bigoplus_{i=1}^3 \boldsymbol{\Sigma}_{ii}$, where $\bigoplus$ denotes the Kronecker sum.

For the multi-sample statistic, T_g , we take $g = 2$ with $b = 2$, and generate $n_1 = \{20, 40, 75\}$ and $n_2 = \{30, 50, 100\}$ iid vectors from respective populations with $p_1 = \{20, 50, 100, 300, 500\}$ and $p_2 = 2p_1$. Under homoscedasticity, the common Σ is assumed to follow the two covariance structures under the alternative, where under the null, the two diagonal blocks are assumed to follow the same structures with their respective dimensions. The estimated size and power of T_b and T_g , reported in Tables 1, 2, 3 and 4, respectively, are an average of 1000 simulation runs.

We observe an accurate test size for both tests, T_b and T_g , for all parameter settings under all distributions. For multi-sample case, the test performs relatively more accurately although the sample sizes reasonably differ. Note that, generally, both tests are very conservative for relatively small sample sizes, but improve with increasing sample size. In particular, the accuracy of the tests remains intact for increasing dimension. The results under the t and uniform distributions point to the evidence of robustness of the tests to the normality assumption, under Model (4).

Similar performance can be witnessed for the power of both test statistics. Like test size, the accuracy of which increases for increasing sample size, and the power also improves with increasing sample size, but also with increasing dimension, even for seriously differing dimensions for different blocks. In particular, for the multi-sample case, we observe an accurate performance of T_g for unbalanced design, and the accuracy is not disturbed for increasing dimension. The robustness character of the two tests also resembles that for the test size.

Table 1 Estimated size of T_b with 3 blocks for normal, t and uniform distributions (ND,TD,UD) under compound symmetric and autoregressive (CS,AR) structures

n	p	ND		TD		CD	
		CS	AR	CS	AR	CS	AR
20	60	0.931	0.936	0.920	0.939	0.933	0.945
	150	0.931	0.938	0.928	0.938	0.941	0.947
	300	0.932	0.940	0.935	0.941	0.941	0.949
	900	0.936	0.944	0.937	0.945	0.947	0.950
	1500	0.938	0.949	0.933	0.948	0.944	0.948
50	60	0.940	0.949	0.939	0.941	0.945	0.947
	150	0.942	0.944	0.941	0.949	0.952	0.944
	300	0.947	0.943	0.944	0.951	0.944	0.951
	900	0.947	0.954	0.946	0.953	0.958	0.954
	1500	0.958	0.951	0.949	0.951	0.948	0.952
100	60	0.948	0.940	0.944	0.947	0.946	0.951
	150	0.944	0.952	0.942	0.945	0.951	0.950
	300	0.947	0.954	0.945	0.948	0.955	0.954
	900	0.948	0.950	0.948	0.955	0.951	0.949
	1500	0.952	0.947	0.950	0.953	0.949	0.946

Table 2 Estimated power of T_b with 3 blocks for normal, t and uniform distributions (ND,TD,UD) under compound symmetric and autoregressive (CS,AR) structures

n	p	ND		TD		CD	
		CS	AR	CS	AR	CS	AR
20	60	0.116	0.258	0.144	0.153	0.115	0.283
	150	0.130	0.422	0.201	0.168	0.194	0.418
	300	0.224	0.482	0.257	0.246	0.324	0.480
	900	0.365	0.625	0.359	0.330	0.378	0.624
	1500	0.596	0.726	0.571	0.453	0.508	0.738
50	60	0.213	0.389	0.243	0.254	0.228	0.531
	150	0.305	0.547	0.289	0.315	0.295	0.798
	300	0.572	0.849	0.559	0.539	0.564	0.875
	900	0.761	0.958	0.759	0.736	0.783	0.957
	1500	0.970	1.000	0.958	0.814	0.887	0.984
100	60	0.452	0.683	0.474	0.609	0.562	0.782
	150	0.619	0.792	0.587	0.902	0.641	0.835
	300	0.933	0.993	0.874	0.978	0.929	0.993
	900	1.000	1.000	0.986	0.999	0.989	0.999
	1500	1.000	1.000	0.999	1.000	0.993	1.000

Table 3 Estimated size of T_g with 3 blocks for normal, t and uniform distributions (ND,TD,UD) under compound symmetric and autoregressive (CS,AR) structures

(n_1, n_2)	p	ND		TD		CD	
		CS	AR	CS	AR	CS	AR
(20, 30)	60	0.931	0.937	0.930	0.938	0.936	0.941
	150	0.938	0.935	0.936	0.941	0.939	0.942
	300	0.940	0.943	0.934	0.943	0.939	0.939
	900	0.939	0.941	0.940	0.939	0.942	0.940
	1500	0.942	0.945	0.941	0.940	0.947	0.944
(40, 50)	60	0.939	0.942	0.940	0.939	0.946	0.942
	150	0.941	0.944	0.942	0.943	0.945	0.942
	300	0.947	0.948	0.945	0.947	0.947	0.941
	900	0.942	0.953	0.947	0.949	0.947	0.944
	1500	0.944	0.945	0.952	0.949	0.951	0.947
(75, 100)	60	0.956	0.947	0.944	0.944	0.949	0.943
	150	0.952	0.948	0.948	0.954	0.956	0.947
	300	0.951	0.946	0.949	0.953	0.952	0.947
	900	0.951	0.949	0.948	0.950	0.949	0.952
	1500	0.955	0.948	0.953	0.954	0.951	0.954

6 Discussion and Conclusions

Test statistics for correlation between two or more vectors are presented when the dimensions of the vectors, possibly unequal, may exceed the number of vectors. The

Table 4 Estimated power of T_g with 3 blocks for normal, t and uniform distributions (ND,TD,UD) under compound symmetric and autoregressive (CS,AR) structures

n	p	ND		TD		CD	
		CS	AR	CS	AR	CS	AR
(20, 30)	60	0.705	0.774	0.825	0.659	0.318	0.685
	150	0.882	0.891	0.915	0.783	0.356	0.844
	300	0.953	0.914	0.993	0.852	0.589	0.868
	900	0.985	0.934	1.000	0.872	0.704	0.902
	1500	1.000	0.951	1.000	0.895	0.795	0.934
(40, 50)	60	0.978	0.981	0.912	0.892	0.537	0.884
	150	0.995	0.997	0.958	0.927	0.729	0.989
	300	0.999	1.000	1.000	0.989	0.912	0.999
	900	1.000	1.000	1.000	0.994	0.967	1.000
	1500	1.000	1.000	1.000	0.999	0.993	1.000
(75, 100)	60	1.000	1.000	0.984	0.970	0.962	1.000
	150	1.000	1.000	1.000	0.996	0.998	1.000
	300	1.000	1.000	1.000	1.000	1.000	1.000
	900	1.000	1.000	1.000	1.000	1.000	1.000
	1500	1.000	1.000	1.000	1.000	1.000	1.000

one-sample case is further extended to two or more independent samples coming from populations assumed to have common covariance matrix. Accuracy of the tests is shown through simulations with different parameters. Among potential advantages of the tests include their simple construction, particularly for the multi-sample case whence most computations follow conveniently from the one-sample results, and their wide practical applicability under fairly general conditions and for a larger class of multivariate models including the multivariate normal.

A particularly distinguishing feature of the tests is that they are composed of computationally very efficient estimators defined as simple functions of the empirical covariance matrix. From applications perspective, it may be mentioned that the tests are constructed using the RV coefficient so that they can only be used to assess linear independence of vectors. It distinguishes them from measures such as distance correlation or kernel methods which can also measure nonlinear dependence; see also [5] for more details.

Acknowledgements The author is thankful to the editors and the two referees for their comments which helped improve the original draft of the article.

Compliance with Ethical Standards

Conflict of interest There is no conflict of interest concerning this manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Technical Results and Proofs

Some Basic Moments

Given Model (4) with $\mathbf{Y}_{ik} = \mathbf{X}_{ik} - \boldsymbol{\mu}_i$, let $A_{ik} = \mathbf{Y}_{ik}^T \mathbf{Y}_{ik}$, $A_{ikr} = \mathbf{Y}_{ik}^T \mathbf{Y}_{ir}$, $k \neq r$, and κ_{ii} , κ_{ij} be as in Eq. (10); then, the moments in the following theorem hold under Model (4).

Theorem 13 For A_{ik} , A_{ikr} defined above, $E(A_{ik}) = \text{tr}(\boldsymbol{\Sigma}_{ii})$, $E(A_{ikr}) = 0$, $E(A_{ik}^2) = \|\boldsymbol{\Sigma}_{ii}\|^2$, $E(A_{ikr}A_{iks}) = 0$, $E(A_{ik}A_{ikr}) = 0$, $E(\mathbf{Y}_{ik}^T \boldsymbol{\Sigma}_{ij} \mathbf{Y}_{jk}) = \|\boldsymbol{\Sigma}_{ij}\|^2$, $E(\mathbf{Y}_{ik}^T \boldsymbol{\Sigma}_{ij} \mathbf{Y}_{jr}) = 0$, $\text{Var}(\mathbf{Y}_{ik}^T \boldsymbol{\Sigma}_{ij} \mathbf{Y}_{jk}) = K + \text{tr}(\boldsymbol{\Sigma}_{ii} \boldsymbol{\Sigma}_{ij} \boldsymbol{\Sigma}_{jj} \boldsymbol{\Sigma}_{ji})$, $\text{Var}(\mathbf{Y}_{ik}^T \boldsymbol{\Sigma}_{ij} \mathbf{Y}_{jr}) = \text{tr}(\boldsymbol{\Sigma}_{ii} \boldsymbol{\Sigma}_{ij} \boldsymbol{\Sigma}_{jj} \boldsymbol{\Sigma}_{ji})$, $\text{Var}(A_{ikr}A_{jrk}) = K + \|\boldsymbol{\Sigma}_{ii}\|^2 \|\boldsymbol{\Sigma}_{jj}\|^2 + 2 \text{tr}(\boldsymbol{\Sigma}_{ii} \boldsymbol{\Sigma}_{ij} \boldsymbol{\Sigma}_{jj} \boldsymbol{\Sigma}_{ji})$, where K is a constant involving only κ_{ij} and κ_{ii} .

Main Proofs

Proof of Theorem 1

Since we can write $\hat{\boldsymbol{\Sigma}} = \sum_{k=1}^n \mathbf{Y}_k \mathbf{Y}_k^T / n - \sum_{k \neq r}^n \mathbf{Y}_k \mathbf{Y}_r^T / (n(n-1))$, it implies ([2, 4])

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_{ij} &= \frac{1}{n} \sum_{k=1}^n \mathbf{Y}_{ik} \mathbf{Y}_{jk}^T - \frac{1}{n(n-1)} \sum_{k=1}^n \sum_{\substack{r=1 \\ k \neq r}}^n \mathbf{Y}_{ik} \mathbf{Y}_{jr}^T, \\ \hat{\boldsymbol{\Sigma}}_{ii} &= \frac{1}{n} \sum_{k=1}^n \mathbf{Y}_{ik} \mathbf{Y}_{ik}^T - \frac{1}{n(n-1)} \sum_{k=1}^n \sum_{\substack{r=1 \\ k \neq r}}^n \mathbf{Y}_{ik} \mathbf{Y}_{ir}^T.\end{aligned}$$

Using $\hat{\boldsymbol{\Sigma}}_{ij}$ and Theorem 13, Eqs. (7) and (9) can be obtained from ‘Appendix B.1’ in [4] by replacing 1 with i and 2 with j . They express $\|\hat{\boldsymbol{\Sigma}}_{ij}\|^2$, $\|\hat{\boldsymbol{\Delta}}_{ii}\|^2 \|\hat{\boldsymbol{\Delta}}_{jj}\|^2$ and Q_{ij} in terms of functions of A_{ik} and A_{ikr} , defined above, so that, taking expectation, it follows that

$$E\left[\|\hat{\boldsymbol{\Sigma}}_{ij}\|^2\right] = \frac{1}{n} \kappa_{ij} + \frac{1}{n-1} \{n \|\boldsymbol{\Sigma}_{ij}\|^2 + \|\boldsymbol{\Delta}_{ii}\|^2 \|\boldsymbol{\Delta}_{jj}\|^2\}, \quad (44)$$

$$E\left[\|\hat{\boldsymbol{\Delta}}_{ii}\|^2 \|\hat{\boldsymbol{\Delta}}_{jj}\|^2\right] = \frac{1}{n} \kappa_{ij} + \frac{2}{n-1} \|\boldsymbol{\Sigma}_{ij}\|^2 + \|\hat{\boldsymbol{\Delta}}_{ii}\|^2 \|\hat{\boldsymbol{\Delta}}_{jj}\|^2, \quad (45)$$

$$E(Q_{ij}) = \frac{n^2 - 3n + 3}{n^2} \kappa_{ij} + \frac{n-1}{n} \{2 \|\boldsymbol{\Sigma}_{ij}\|^2 + \|\boldsymbol{\Delta}_{ii}\|^2 \|\boldsymbol{\Delta}_{jj}\|^2\}. \quad (46)$$

Solving these equations simultaneously gives Eqs. (7) and (9), which can be used to show unbiasedness of $\hat{\kappa}_{ij}$. Following the same lines, and using $\hat{\Sigma}_{ii}$ above, we can write

$$\begin{aligned}\|\hat{\Sigma}_{ii}\|^2 &= \frac{1}{n^2} \left\{ \sum_{k=1}^n A_{ik}^2 + \sum_{k=1}^n \sum_{\substack{r=1 \\ k \neq r}}^n \left[\left(1 + \frac{1}{(n-1)^2} \right) A_{ikr}^2 + A_{ik} A_{ir} \right] \right\} + A_{01} \\ \{\|\hat{\Delta}_{ii}\|\}^2 &= \frac{1}{n^2} \left\{ \sum_{k=1}^n A_{ik}^2 + \sum_{k=1}^n \sum_{\substack{r=1 \\ k \neq r}}^n \left(\frac{2}{(n-1)^2} A_{ikr}^2 + A_{ik} A_{ir} \right) \right\} + A_{02} \\ Q_{ii} &= \frac{1}{n^3(n-1)} \left\{ (n^3 - 3n + 3)(n-1) \sum_{k=1}^n A_{ik}^2 \right. \\ &\quad \left. + (2n-3) \sum_{k=1}^n \sum_{\substack{r=1 \\ k \neq r}}^n (2A_{ikr}^2 + A_{ik} A_{ir}) \right\} + A_{03},\end{aligned}$$

where, for simplicity, A_{01}, A_{02}, A_{03} contain terms the expectation of which vanishes, so that, using Theorem 13, it follows after some simplification that

$$E\left(\|\hat{\Sigma}_{ii}\|^2\right) = \frac{1}{n} \kappa_{ii} + \frac{1}{n-1} \left\{ n \|\Sigma_{ii}\|^2 + [\|\Delta_{ii}\|^2]^2 \right\}, \quad (47)$$

$$E\left([\|\hat{\Delta}_{ii}\|^2]^2\right) = \frac{1}{n} \kappa_{ii} + \frac{2}{n-1} \|\Sigma_{ii}\|^2 + [\|\Delta_{ii}\|^2]^2, \quad (48)$$

$$E(Q_{ii}) = \frac{n^2 - 3n + 3}{n^2} \kappa_{ii} + \frac{n-1}{n} \left\{ 2 \|\Sigma_{ii}\|^2 + [\|\Delta_{ii}\|^2]^2 \right\}. \quad (49)$$

Solving simultaneously gives Eqs. (6) and (8), and also the unbiasedness of $\hat{\kappa}_{ii}$ in Eq. (12).

Proof of Theorem 11

Under the assumption $\Sigma_l = \Sigma \forall l = 1, \dots, g$, the unbiasedness follows immediately since

$$E\left(\widehat{\|\Sigma_{pii}\|^2}\right) = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} \|\Sigma_{ii}\|^2 = \|\Sigma_{ii}\|^2, \quad E\left(\widehat{\|\Sigma_{pij}\|^2}\right) = \frac{1}{v_0} \sum_{l=1}^g \frac{1}{v_l} \|\Sigma_{ij}\|^2 = \|\Sigma_{ij}\|^2$$

$\forall i, j = 1, \dots, b, i < j$. For variances, we note, following Theorem 2 for any $l \in \{1, \dots, g\}$, that

$$\begin{aligned}\text{Var}\left(\widehat{\|\Sigma_{pii}\|^2}\right) &= \frac{1}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \text{Var}\left(\widehat{\|\Sigma_{lii}\|^2}\right) \\ &= \frac{1}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \frac{4}{P(n_l)} \left[d(n_l) \|\Sigma_{ii}^2\|^2 + e(n_l) [\|\Sigma_{ii}^2\|]^2 + KO(n_l^2) \right],\end{aligned}$$

where $d(n_l) = 2n_l^3 - 9n_l^2 + 9n_l - 16$, $e(n) = n_l^2 - 3n_l + 8$, so that we can write

$$\text{Var}\left(\widehat{\|\Sigma_{pii}\|^2}\right) = \frac{4}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \left[O\left(\frac{1}{n_l}\right) \|\Sigma_{ii}^2\|^2 + O\left(\frac{1}{n_l^2}\right) [\|\Sigma_{ii}^2\|]^2 + KO\left(\frac{1}{n_l^2}\right) \right].$$

The first and last terms vanish under the assumptions, so that, as $n, p_i \rightarrow \infty$,

$$\text{Var}\left(\frac{\widehat{\|\Sigma_{pii}\|^2}}{\|\Sigma_{ii}\|^2}\right) = \frac{4}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} O(n_l^{-2}) = 4O\left(\frac{1}{v_0}\right),$$

using $v_l^{-1} = O(n_l^2) \Rightarrow v_0 = O(\sum_{l=1}^g n_l^2)$, which also gives the consistency of $\widehat{\|\Sigma_{pii}\|^2}/p_i^2$. Similarly,

$$\text{Var}\left(\widehat{\|\Sigma_{pij}\|^2}\right) = \frac{1}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \text{Var}\left(\widehat{\|\Sigma_{lii}\|^2}\right) \quad (50)$$

$$\begin{aligned}&= \frac{2}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \left[\left\{ 4\|\Sigma_{ij}\Sigma_{ji}\|^2 + 2\text{tr}(\Sigma_{ii}\Sigma_{ij}\Sigma_{jj}\Sigma_{ji}) \right\} O\left(\frac{1}{n_l}\right) \right. \\ &\quad \left. + 2\left\{ [\|\Sigma_{ij}\|^2]^2 + \|\Sigma_{ii}\|^2 \|\Sigma_{jj}\|^2 + KO(1) \right\} O\left(\frac{1}{n_l^2}\right) \right]. \quad (51)\end{aligned}$$

With $\eta_{ij} = \|\Sigma_{ij}\|^2 / \|\Sigma_{ii}\| \|\Sigma_{jj}\|$, it implies under the assumptions that, as $p_i \rightarrow \infty$,

$$\text{Var}\left(\frac{\widehat{\|\Sigma_{pij}\|^2}}{\|\Sigma_{ij}\|^2}\right) = \frac{4}{v_0^2} \sum_{l=1}^g \frac{1}{v_l^2} \left\{ 1 + \frac{1}{\eta_{ij}^2} \right\} O\left(\frac{1}{n_l^2}\right) = 4\left(1 + \frac{1}{\eta_{ij}}\right) O\left(\frac{1}{v_0}\right).$$

Proofs of Theorems 6 and 12

Consider $T_b = \sum_{i < j}^b T_{ij}$ with $T_{ij} = \widehat{\|\Sigma_{ij}\|^2} / p_i p_j$. In Ahmad ([5], Theorem 3), it is shown that

$$\frac{T_{ij} - E(T_{ij})}{\sqrt{\text{Var}(T_{ij})}} \xrightarrow{D} N(0, 1),$$

under Assumptions 3–5, as $n, p_i \rightarrow \infty$, where $\text{Var}(T_{ij})$ follows from Theorem 2. As the limit holds for all T_{ij} , $i < j$, and T_b is a sum of all such T_{ij} , we basically need to focus on the covariances in Eqs. (19)–(21) for the limit of T_b . Consider C_1 , where the first term, normed by $\|\Sigma_{ii}\|^2 \|\Sigma_{jj}\| \|\Sigma_{jj'}\|$, vanishes under Assumptions 3–5, as $p_i \rightarrow \infty$, and the same holds for the second term. Repeating the same for C_2 and C_3 , and noting that the terms like $\|\Sigma_{ii}\|^2 / p_i^2$ are uniformly bounded under the same assumptions, it follows that

$$\begin{aligned} C_1 &= (2\eta_{ij}\eta_{ij'} + 5\eta_{ij'})O(n^{-3}), \quad C_2 = (2\eta_{ij}\eta_{ij'} + 5\eta_{ij'})O(n^{-3}), \\ C_3 &= (3\eta_{ij}\eta_{ij'} + 5\eta_{ij'})O(n^{-3}), \end{aligned}$$

where $\eta_{ij} = \|\Sigma_{ij}\|^2 / \|\Sigma_{ii}\| \|\Sigma_{jj}\|$; see Eq. (5). Combined with Eq. (22), it implies that $\text{Var}(nT_b)$ is bounded, but the covariances with the same order vanish. Now, denote $\mathbf{T}_B = (\mathbf{T}_1, \dots, \mathbf{T}_{b-1})'$ with $T_b = \mathbf{1}'\mathbf{T}_B$ where $\mathbf{T}_i = (T_{i,i+1}, \dots, T_{ib})'$, $i = 1, \dots, b-1$, $B = b(b-1)/2$ and $\mathbf{1}_B$ is the vector of 1s. By the above arguments, as $n, p_i \rightarrow \infty$, $\text{Cov}(\mathbf{T}_B) = \mathbf{\Lambda}$ is a diagonal matrix with diagonal elements $\text{Var}(T_{ij})$, $i, j = 1, \dots, b$, $i < j$, i.e.,

$$\mathbf{\Lambda} = \text{diag}(\mathbf{\Lambda}_1, \dots, \mathbf{\Lambda}_{b-1}),$$

where $\mathbf{\Lambda}_i = \text{Cov}(\mathbf{T}_i) = \text{diag}(\text{Var}(T_{i1}), \dots, \text{Var}(T_{ib}))$, $i = 1, \dots, b-1$. Hence (see Eq. 22),

$$E(T_b) = \mathbf{1}'\mathbf{T}_B = E(T_b), \quad \text{Var}(T_b) = \mathbf{1}'\mathbf{\Lambda}\mathbf{1} = \text{Var}(T_b).$$

This gives the limit T_b by a simple application of Cramér–Wold device ([31], p. 16), including for the case under the null whence the covariances vanish exactly. For the last part of the theorem, we only need to prove that, we note, from Theorem 2, that

$$\text{Var}\left(\frac{\widehat{\|\Sigma_{ii}\|^2}}{\|\Sigma_{ii}\|^2}\right) = 4 \left[\frac{\text{tr}(\Sigma_{ii}^4)}{[\text{tr}(\Sigma_{ii}^2)]^2} O\left(\frac{1}{n}\right) + [1 + o(1)] O\left(\frac{1}{n^2}\right) \right].$$

The first and last terms vanish under the assumptions, so that $\text{Var}(\widehat{\|\Sigma_{ii}\|^2} / \|\Sigma_{ii}\|^2) \rightarrow 4O(1/n^2)$ as $p_i \rightarrow \infty$. Thus, $\widehat{\|\Sigma_{ii}\|^2} / p_i^2 \rightarrow \|\Sigma_{ii}\|^2 / p_i^2$, as $n, p_i \rightarrow \infty$. Plugging in $\|\Sigma_{ii}\|^2$ for $\widehat{\|\Sigma_{ii}\|^2}$ in $\text{Var}(T_b)$ gives $\text{Var}(T_b)$ as a consistent estimator of $\text{Var}(T_b)$. This completes the proof of Theorem 6.

The proof of Theorem 12 now essentially follows from above, by the independence of g samples. First, the covariance terms in $\text{Var}(T_{lb})$, i.e.,

$\text{Cov}(\mathbf{T}_{lij}, \mathbf{T}_{lij'})$, $\text{Cov}(\mathbf{T}_{lij}, \mathbf{T}_{li'j'})$ and $\text{Cov}(\mathbf{T}_{lij}, \mathbf{T}_{li'j'})$, asymptotically vanish similarly as C_1 , C_2 , C_3 did for $\text{Var}(\mathbf{T}_b)$. Then, with $v_l = O(n_l^2) \Rightarrow v_0 = O(\sum_{l=1}^g n_l^2)$, $\text{Var}(\mathbf{T}_g) = 4O(v_0^{-1}) \sum_i \sum_j V_{lij}$, with V_{lij} uniformly bounded under the assumptions, so that $\sqrt{v_0} \mathbf{T}_g$ has a non-degenerate limit, as $n, p_i \rightarrow \infty$. Define $\mathbf{T}_G = (\mathbf{T}_1, \dots, \mathbf{T}_g)'$ with $\mathbf{T}_g = \mathbf{1}' \mathbf{T}_G$, where $\mathbf{T}_l = (\mathbf{T}_{l1}, \dots, \mathbf{T}_{l,b-1})'$ and $\mathbf{T}_{li} = (\mathbf{T}_{li,i+1}^l, \dots, \mathbf{T}_{li,b}^l)'$, $i = 1, \dots, b$, $l = 1, \dots, g$. By the independence of g samples, $\text{Cov}(\mathbf{T}_G) = \Delta$ reduces to

$$\Delta = \text{diag}(\Delta_{11}, \dots, \Delta_{gg}),$$

where, by the above arguments, $\Delta_{ll} = \text{Cov}(\mathbf{T}_l)$ is again a diagonal matrix with diagonal elements $\text{Var}(\mathbf{T}_{lij})$, $i, j = 1, \dots, b$, $i < j$. Thus,

$$\mathbf{E}(\mathbf{T}_g) = \mathbf{1}' \mathbf{E}(\mathbf{T}_G), \quad \text{Var}(\mathbf{T}_g) = \mathbf{1}' \Delta \mathbf{1}$$

asymptotically coincide with Eqs. (39) and (40), respectively. Finally, the consistency of $\widehat{\text{Var}}(\mathbf{T}_g)$ follows from that of Σ_{pii}/p_i^2 as shown in “Appendix 1.2.2”.

References

- Ahmad MR (2016) On testing sphericity and identity of a covariance matrix with large dimensions. *Math Methods Stat* 25(2):121–132
- Ahmad MR (2017a) Location-invariant multi-sample U -tests for covariance matrices with large dimension. *Scand J Stat* 44:500–523
- Ahmad MR (2017b) Location-invariant tests of homogeneity of large dimensional covariance matrices. *J Stat Theory Pract* 11:731–745
- Ahmad MR (2018) A unified approach to testing mean vectors with large dimensions. *ASTA Adv Stat Anal*. <https://doi.org/10.1007/s10182-018-00343-z>
- Ahmad MR (2019) A significance test of the RV coefficient for large dimensions. *Comput Stat Data Anal* 131:116–130
- Albert M, Bouret Y, Fromont M, Reynaud-Bouret P (2015) Bootstrap and permutation tests of independence for point processes. *Ann Stat* 43:2537–2564
- Allaire J, Lepage Y (1992) A procedure for assessing vector correlations. *Ann Inst Stat Math* 44:755–768
- Anderson TW (1999) Asymptotic theory for canonical correlation analysis. *J Multivar Anal* 70:1–29
- Anderson TW (2003) An introduction to multivariate statistical analysis, 3rd edn. Wiley, Hoboken
- Bakirov NK, Rizzo ML, Székely GJ (2006) A multivariate nonparametric test of independence. *J Multivar Anal* 97:1742–1756
- de Wet T (1980) Cramér-von Mises tests for independence. *J Multivar Anal* 10:38–50
- Gretton A, Bousquet O, Smola A, Schölkopf B (2005) Measuring statistical dependence with Hilbert-Schmidt norms. In: Jain et al (eds) *Algorithmic learning theory*. Springer, Berlin
- Gretton A, Györfi L (2010) Consistent nonparametric tests of independence. *J Mach Learn Res* 11:1391–1423
- Horváth L, Hušková M, Rice G (2013) Test of independence for functional data. *J Multivar Anal* 117:100–119
- Johnson RA, Wichern DW (2007) *Applied multivariate data analysis*, 6th edn. Prentice Hall, Upper Saddle River
- Kettenring JR (1971) Canonical analysis of several sets of variables. *Biometrika* 58:433–451
- Koroljuk VS, Borovskich YV (1994) *Theory of U -statistics*. Kluwer Press, Dordrecht
- Lehmann EL (1999) *Elements of large-sample theory*. Springer, Berlin
- Muirhead RJ (2005) *Aspects of multivariate statistical theory*. Wiley, Hoboken
- Muirhead RJ, Waternaux CM (1980) Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika* 67:31–43

21. Nkiet GM (2017) Asymptotic theory of multiple-set linear canonical analysis. *Math Methods Stat* 26:196211
22. Pfister N, Bühlmann P, Schülkopf B, Peters J (2017) Kernel-based tests for joint independence. *JRSS B* 80(1):5–31
23. Ramsay JO, ten Berge J, Styan HPH (1984) Matrix correlation. *Psychom* 49:403–423
24. Robert P, Cleroux R, Ranger N (1985) Some results on vector correlation. *Comput Stat Data Anal* 3:25–32
25. Robert P, Escoufier Y (1976) A unifying tool for linear multivariate statistical methods: the RV-coefficient. *J RStat Soc C* 25:257–265
26. Schott JR (2008) A test for independence of two sets of variables when the number of variables is large relative to the sample size. *Stat Probab Lett* 78:3096–3102
27. Searle SR (1971) *Linear models*. Wiley, Hoboken
28. Seber GAF (2004) *Multivariate observations*. Wiley, Hoboken
29. Serfling RJ (1980) *Approximation theorems of mathematical statistics*. Wiley, Hoboken
30. Székely G, Rizzo M (2013) The distance correlation t -test of independence in high dimension. *J Multivar Anal* 117:193–213
31. van der Vaart AW (1998) *Asymptotic statistics*. Cambridge University Press, Cambridge
32. Wang G, Zou C, Wang Z (2013) A necessary test for complete independence in high dimensional using rank correlations. *J Multivar Anal* 121:224–232
33. Yang Y, Pan G (2015) Independence test for high dimensional data based on regularized canonical correlation coefficients. *Ann Stat* 43:467–500
34. Yata K, Aoshima M (2013) Correlation test for high-dimensional data using extended cross-data-matrix methodology. *J Multivar Anal* 117:313–331

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.