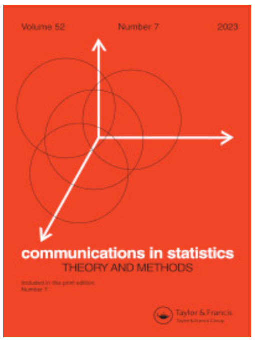

Communications in Statistics - Theory and Methods

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Some correlation tests for vectors of large dimension

M. Rauf Ahmad & S. Ejaz Ahmed

To cite this article: M. Rauf Ahmad & S. Ejaz Ahmed (2023) Some correlation tests for vectors of large dimension, Communications in Statistics - Theory and Methods, 52:7, 2144-2160, DOI: 10.1080/03610926.2021.1945631

To link to this article: https://doi.org/10.1080/03610926.2021.1945631

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC

Published online: 08 Jul 2021.

Submit your article to this journal

Article views: 834

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=lsta20

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2023, VOL. 52, NO. 7, 2144–2160
https://doi.org/10.1080/03610926.2021.1945631

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Some correlation tests for vectors of large dimension

M. Rauf Ahmad[a] and S. Ejaz Ahmed[b]

[a]Department of Statistics, Uppsala University, Uppsala, Sweden; [b]Department of Mathematics and Statistics, Brock University, Saint Catharines, Ontario, Canada

ABSTRACT

For a random sample of n iid p-dimensional vectors, each partitioned into b sub-vectors of dimensions p_i, $i = 1, \ldots, b$, tests for zero correlation of sub-vectors are presented when $p_i \gg n$ and the distribution need not be normal. The test statistics are composed of U-statistics based estimators of the Frobenius norm measuring the distance between the null and alternative hypotheses. Asymptotic distributions of the tests are provided for $n, p_i \to \infty$, with their finite-sample performance demonstrated through simulations. Some related tests are discussed. A real data application is also given.

ARTICLE HISTORY
Received 5 June 2020
Accepted 15 June 2021

KEYWORDS
Canonical correlation; covariance tests; high-dimensional inference; cross-covariance

1. Introduction

A considerable part of multivariate statistics concerns studying correlations and their structures. Often an observed vector can be partitioned into several sub-vectors, possibly of different dimensions, and the interest focuses on testing independence, or reveal the cross-correlations, among sub-vectors; canonical correlation analysis being an important application. We discuss such tests of independence or zero correlation of two or more sub-vectors when the dimensions of the sub-vectors may exceed their number and the data may follow a non-normal distribution. Let

$$\mathbf{X}_k = (\mathbf{X}'_{k1}, \ldots, \mathbf{X}'_{kb})' \in \mathbb{R}^p, \ \ k = 1, \ldots, n,$$

be iid vectors partitioned into $b \geq 2$ sub-vectors $\mathbf{X}_{ki} \in \mathbb{R}^{p_i}$ with $\mathrm{E}(\mathbf{X}_k) = \boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \ldots, \boldsymbol{\mu}'_b)' \in \mathbb{R}^p$, $\mathrm{Cov}(\mathbf{X}_k) = \mathrm{E}(\mathbf{X}_k - \boldsymbol{\mu})(\mathbf{X}_k - \boldsymbol{\mu})' = \boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})_{i,j=1}^{b} \in \mathbb{R}^{p \times p}$, where $\mathrm{E}(\mathbf{X}_{ki}) = \boldsymbol{\mu}_i \in \mathbb{R}^{p_i}$ and $\mathrm{Cov}(\mathbf{X}_{ki}, \mathbf{X}_{kj}) = \mathrm{E}(\mathbf{X}_{ki} - \boldsymbol{\mu}_i)(\mathbf{X}_{kj} - \boldsymbol{\mu}_j)' = \boldsymbol{\Sigma}_{ij} \in \mathbb{R}^{p_i \times p_j}$. A hypothesis of frequent interest in multivariate theory is of independence of $\mathbf{X}_{ki}$ (see e.g. Anderson 2003; Muirhead 2005)

$$H_0 : \mathbf{X}_{ki} \perp\!\!\!\perp \mathbf{X}_{kj} \ \forall \ i \neq j \ \text{vs.} \quad H_1 : \mathbf{X}_{1i} \not\perp\!\!\!\perp \mathbf{X}_{1j} \text{ for at least one pair } i \neq j. \tag{1}$$

As it leads to a drastic dimension reduction under H_0, the test is even more desirable for large parameter spaces which motivates our main objective, that is, to present a test of H_0 when $p_i \gg n$. Multivariate theory offers likelihood ratio tests of (1) under

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

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

normality, that is, $\mathbf{X}_k \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, whence $\mathbf{X}_{ki} \perp\!\!\!\perp \mathbf{X}_{kj} \iff \boldsymbol{\Sigma}_{ij} = \mathbf{0}$ and H_0 in (1) reduces to testing significance of cross-correlations, that is,

$$H_0 : \boldsymbol{\Sigma}_{ij} = \mathbf{0} \ \forall \ i \neq j \quad \text{vs.} \quad H_1 : \boldsymbol{\Sigma}_{\mathrm{ij}} \neq \mathbf{0} \quad \text{for at least one pair } i \neq j. \tag{2}$$

Thus, under normality, (1) $\iff$ (2), where in general (1) $\Rightarrow$ (2). There is extensive literature on the tests of H_0 in the classical set up, that is, $n > p_i$. Anderson (1999) and Eaton and Tyler (1994) established basic asymptotic theory with an extension for non-normal case in Muirhead and Waternaux (1980), where Nkiet (2017) discussed the case of multiple blocks. A nonparametric test is considered in Gretton and Györfi (2010), Pfister et al. (2018) provide a kernel based test, Horváth, Hušková, and Rice (2013) treat the case for functional data and Albert et al. (2015) give permutation tests.

The classical tests, however, collapse or are inefficient when $p_i > n$, mainly due to singularity of $\hat{\boldsymbol{\Sigma}}$ or $\hat{\boldsymbol{\Sigma}}_{ii}$, $i = 1, ..., b$, where $\hat{\mathbf{A}}$ denotes an estimator of $\mathbf{A}$. Several modifications have recently been put forth; see for example, Schott (2008) where a test of complete independence is given in Schott (2005), an extension of which is given in Mao (2020). Yang and Pan (2015) extend the canonical correlation through regularization for high-dimensional case. Another test, using block correlation matrices, is proposed in Bao et al. (2017), where Srivastava, Kollo, and von Rosen (2011) and Xu (2017) provide diagonality tests relaxing normality assumption, where a similarity coefficient based treatment is given in Ahmad (2019).

We present tests of H_0 in (1) when the data are high-dimensional but not necessarily normal. The tests are defined as U-statistics with kernels estimating the Frobenius norm of cross-covariance matrix $\boldsymbol{\Sigma}_{ij}$. This helps us study the properties of tests under a general multivariate model with certain mild assumptions. For practical use, however, we also provide simpler, computationally more efficient versions of the same estimators.

An important property of the tests is that they are location-invariant, so that the true mean vector can be assumed zero for their use, without any loss of generality. This property follows from the kernels of the U-statistics used to compose the test statistics. Given that, a completely affine invariant test in high-dimensional set is possible only under very restricted cases, the location-invariance property provides an added value to the tests for their practical applications.

Tests under normality are presented in Section 2, with an extension to the general case in Section 3. Some related tests are discussed in Section 4. Section 5 provides simulation based assessment and a real data application is given in Section 6. Proofs are collected in the Appendix.

1.1. A note on notations

Following basic notations will be used throughout the manuscript. Given the data set above, we assume $\mathrm{Cov}(\mathbf{X}_k) = \boldsymbol{\Sigma} \in \mathbb{R}^{p\times p}_{>0}$, where $\mathbb{R}^{a\times b}$ denotes the space of real (and symmetric, positive-definite, if $a = b$) matrices, so that $\boldsymbol{\Sigma}_{ii} > 0$ $(i = j)$. We assume, without loss of generality, that $p_i \leq p_j \ \forall \ i < j, \ i, j = 1, ..., b$. For a matrix $\mathbf{A}_{p\times q}$, $\|\mathbf{A}\|^2 = \mathrm{tr}(\mathbf{A}'\mathbf{A})$ denotes the Frobenius norm. The notation $\omega_{abcd} = \mathrm{tr}(\mathbf{A}_{ab}\mathbf{A}_{bc}\mathbf{A}_{cd}\mathbf{A}_{da})$, $\mathbf{A}_{ab} \in \mathbb{R}^{a\times b}$, will help us simplify many expressions. Since the test statistics are defined as U-statistics, at times, we assume a Hilbert space $\mathbb{L}_2(\cdot)$ equipped with inner product $\langle \cdot, \cdot \rangle$:

$\mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, used to define the kernel $h(\cdot)$ as measurable, square-integrable function, $\int h^2 dP < \infty$, composed of symmetric bilinear forms denoted as $A_{kl12} = \mathbf{X}'_{k1}\mathbf{X}_{l2}$ with $A_{k1} = \mathbf{X}'_{k1}\mathbf{X}_{k1}$ corresponding quadratic form.

2. Test statistics under normality

2.1. The case of two blocks

For the set up in 1.1, let $b = 2$ so that $\mathbf{X}_k = (\mathbf{X}_{k1}, \mathbf{X}_{k2})'$. For a quadruplet $\{\mathbf{X}_{ki}, \mathbf{X}_{ri}, \mathbf{X}_{li}, \mathbf{X}_{si}\} \in \mathbf{X}_i$, $k \neq r \neq l \neq s$, let $\mathbf{D}_{kri} = \mathbf{X}_{ki} - \mathbf{X}_{ri}$, $\mathbf{D}_{lsi} = \mathbf{X}_{li} - \mathbf{X}_{si}$ with $\mathrm{E}(\mathbf{D}_{kri}) = \mathbf{0}$, $\mathrm{Cov}(\mathbf{D}_{kri}) = 2\boldsymbol{\Sigma}_{ii}$, $i = 1, 2$, $\mathrm{E}(\mathbf{D}_{kr1}\mathbf{D}'_{kr2}) = 2\boldsymbol{\Sigma}_{12}$ so that $\mathrm{E}(A_{lskr12}) = 4\|\boldsymbol{\Sigma}_{12}\|^2$ for the bilinear form $A_{lskr12} = A_{lskr1}A_{krls2}$ with $A_{lskri} = \mathbf{D}'_{lsi}\mathbf{D}_{kri}$. Denote $B_{klrs12} = A_{lskr12} + A_{lksr12} + A_{lrks12}$ and $P(n) = n(n-1)(n-2)(n-3)$. The test statistic for H_0 is defined as following where $\pi(\cdot)$ implies all indices unequal.

$$\mathrm{T}_2 = \frac{1}{P(n)} \sum_{k=1}^{n} \sum_{r=1}^{n} \sum_{l=1}^{n} \sum_{\substack{s=1 \\ \pi(k,\, r,\, l,\, s)}}^{n} \frac{1}{12 p_1 p_2} B_{lskr12}, \tag{3}$$

T_2 is a U-statistic with $\mathrm{E}(\mathrm{T}_2) = \|\boldsymbol{\Sigma}_{12}\|^2 / p_1 p_2$ and

$$\mathrm{Var}(\mathrm{T}_2) = \frac{2}{P(n) p_1^2 p_2^2} \left[4a(n)\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + 2b(n)\omega_{1122} + c(n)\left\{[\|\boldsymbol{\Sigma}_{12}\|]^2 + \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2\right\}\right] \tag{4}$$

$$= \frac{4}{P(n) p_1^2 p_2^2} \|\boldsymbol{\Sigma}_{11}\|^2 \|\boldsymbol{\Sigma}_{22}\|^2 c(n) = \frac{4}{p_1^2 p_2^2} \|\boldsymbol{\Sigma}_{11}\|^2 \|\boldsymbol{\Sigma}_{22}\|^2 O\left(\frac{1}{n^2}\right) \quad \text{under } H_0, \tag{5}$$

where $a(n) = 3n^3 - 24n^2 + 44n + 20$, $b(n) = 6n^3 - 40n^2 + 22n + 181$, $c(n) = 2n^2 - 12n + 21$. Tests for high-dimensional covariance matrices are often defined in terms of Frobenius norm between the null and alternative hypothesis. Following this, we can define a test as an estimator of $\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_D\|^2$ with $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_D = \oplus_{i=1}^2 \boldsymbol{\Sigma}_{ii}$ under H_0, where $\oplus$ denotes the Kronecker sum. Since $\|\boldsymbol{\Sigma}_{12}\|^2 = \|\boldsymbol{\Sigma}_{12} - \mathbf{O}\|^2$ measures the same distance between H_0 and H_1, it helps us define a simpler form of T_2 in Equation (3) as an estimator of $\|\boldsymbol{\Sigma}_{12}\|^2$.

Note that, T_2 is defined as a non-parametric (U-statistic) estimator of $\|\boldsymbol{\Sigma}_{12}\|^2$ to test H_0 in (1) which, under normality, implies (2). It holds since $\mathrm{E}(\mathrm{T}_2) = \|\boldsymbol{\Sigma}_{12}\|^2 = 0$ under H_0 in both (1) and (2). This will help us keep the same statistic for the non-normal case in Section 3. Further, T_2 is location-invariant. If, however, we can assume that $\boldsymbol{\mu} = \mathbf{0}$, then $\mathrm{E}(\mathbf{X}_{k1}\mathbf{X}'_{k2}) = \boldsymbol{\Sigma}_{12}$ so that $\mathrm{E}(A_{rk1}A_{kr2}) = \|\boldsymbol{\Sigma}_{12}\|^2$ with $A_{rki} = \mathbf{X}'_{ki}\mathbf{X}_{ri}$, $k \neq r$, $i = 1, 2$, and T_2 in (3) simplifies to $\sum_{k \neq r}^{n} A_{rk1}A_{kr2}/p_1 p_2 Q(n)$, $Q(n) = n(n-1)$. As a U-statistic of order 2, it is simpler than T_2 and has simpler properties, except that it is not location-invariant.

For the limit of T_2, we need certain assumptions. We state them in a general form to also use them later when we generalize T_2 to $b \geq 2$ blocks and to non-normal case. Let λ_{si} be the eigenvalues of $\boldsymbol{\Sigma}_{ii}$, so that $\nu_{si} = \lambda_{si}/p_i$ are those of $\boldsymbol{\Sigma}_{ii}/p_i$.

Assumption 1. *For* $p_i \to \infty, \ \sum_{s=1}^{p_i} \nu_{si} = O(1), \ i = 1, ..., b \geq 2.$
Assumption 2. *For* $n, p_i \to \infty, \ p_i/n \to \delta_0 \in (0, \infty), \ i, = 1, ..., b \geq 2.$

Let $\eta_{12}^2 = \zeta/\varphi, \ \zeta = \|\boldsymbol{\Sigma}_{ij}\|^2/p_i p_j$ ($i=j$ o r$i \neq j, \ i, j = 1, 2$), so that $\varphi = \|\boldsymbol{\Sigma}_{11}\|\|\boldsymbol{\Sigma}_{22}\|/p_1 p_2$. By Assumption 1, $\|\boldsymbol{\Sigma}_{ij}\|^2/p_i p_j = O(1)$ and $\omega_{iijj} = O(1), \ i, j = 1, 2$ (see 1.1). Thus, $\zeta, \ \varphi, \ \eta_{12}$ are each bounded, and from (4)

$$\mathrm{Var}\left[n(\mathrm{T}_2 - \|\boldsymbol{\Sigma}_{12}\|^2)/\varphi\right] = (\eta_{12}^2 + 1)O(1) + o(1), \tag{6}$$

which further implies $\mathrm{Var}[n(\mathrm{T}_2 - \|\boldsymbol{\Sigma}_{12}\|^2)/\varphi] = O(1)$. In particular, under H_0, $\mathrm{Var}(n\mathrm{T}_2/\varphi) = O(1)$ so that $n\mathrm{T}_2$ has a non-degenerate limit, under the assumptions. That this is the case for many useful covariance structures under the assumptions, consider for example, $\boldsymbol{\Sigma} = (1-\rho)\mathbf{I} + \rho\mathbf{J}$ (compound symmetric, CS) with $\mathbf{I}$ as identity matrix, $\mathbf{J} = \mathbf{1}\mathbf{1}'$, $\mathbf{1}$ a vectors of 1 s, $\rho \in \mathbb{R}, \ -1/(p-1) \leq \rho \leq 1$. Then $\mathrm{tr}(\boldsymbol{\Sigma}_{ii}^m) = O(p^m), \ m = 1, 2$, satisfying Assumption 1. Note that, CS belongs to the class of spiked structures where a few eigenvalues dominate the rest. In Section 5, we show the accuracy of T_2 under CS, and under AR(1) as non-spiked structure. Assumptions 1 and 2 will let part of $\mathrm{Var}(n\mathrm{T}_2)$ vanish and the rest uniformly bounded, providing the required limit.

Theorem 1 gives the limit of T_2 which holds only under Assumptions 1 and 2 (in fact, the null limit needs only Assumption 1). In the theorem, $\sigma_{\mathrm{T}_2}^2$ denotes $\mathrm{Var}(\mathrm{T}_2)$ in Equation (4) and $\sigma_{T_{20}}^2$ denotes $\mathrm{Var}(\mathrm{T}_2)$ under H_0 in Equation in (5).

Theorem 1. *For* T_2 *in* (3), $(\mathrm{T}_2 - \|\boldsymbol{\Sigma}_{12}\|^2)/\sigma_{\mathrm{T}_2} \xrightarrow{\mathcal{D}} N(0, 1)$ *as* $n, p_i \to \infty$, *under Assumptions 1–2. In particular, under* H_0 *and Assumption 1,* $n\mathrm{T}_2/\sigma_{T_{20}} \to N(0, 1)$.

From the proof (Section A.3), we note that the kernel of T_2 is first-order degenerate under H_0 and the null limit follows through a weighted sum of χ_1^2 variates. To use T_2, we need to estimate $\|\boldsymbol{\Sigma}_{ii}\|^2$. Using the notations around Equation (3), define $C_{krik'r'}^2 = A_{lskri}^2 + A_{lksri}^2 + A_{lrski}^2$ with $A_{lskri} = \mathbf{D}_{lsi}'\mathbf{D}_{kri}$ and $\mathrm{E}(A_{lskri}^2) = 4\|\boldsymbol{\Sigma}_{ii}\|^2$. Define

$$\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2} = \frac{1}{12P(n)} \sum_{\substack{k=1 \\ \pi(k,}}^{n} \sum_{\substack{r=1 \\ r,}}^{n} \sum_{\substack{l=1 \\ l,}}^{n} \sum_{\substack{s=1 \\ s)}}^{n} C_{krlsi}^2, \tag{7}$$

where $\pi(\cdot)$ denotes that all indices are unequal. Note that, $\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2}$ is also a U-statistic and it can be shown that $\mathrm{Var}(\widehat{\|\boldsymbol{\Sigma}\|_{ii}^2}/\|\boldsymbol{\Sigma}\|_{ii}^2)$ is uniformly bounded in p_i. By Assumption 1, as $n_i, p \to \infty$

$$\widehat{\|\boldsymbol{\Sigma}\|_{ii}^2}/p_i^2 \xrightarrow{\mathcal{P}} \sum_{s=1}^{\infty} \delta_{s_i}^2, \quad i = 1, 2, \tag{8}$$

giving consistency of $\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2}/p_i^2$. We have the following corollary to Theorem 1.

Corollary 1. *Theorem 1 remains valid if* $\|\boldsymbol{\Sigma}_{ii}\|^2$ *is replaced with* $\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2}$ *in* $\mathrm{Var}(\mathrm{T}_2)$.

For power of T_2, let z_α be 100α% quantile of $N(0, 1)$ and denote $\beta(\boldsymbol{\theta})$ as the power function with $\boldsymbol{\theta} = \{\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22}, \boldsymbol{\Sigma}_{12}\}$, or $\{\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22}\}$ under H_0. By Theorem 1, $P(\mathrm{T}_2/\sigma_{T_{20}} \geq z_\alpha) =$

α. Let $\gamma = \sigma_{T_{20}}/\sigma_{T_2}$, $\delta = \|\boldsymbol{\Sigma}_{12}\|^2/\sigma_{T_2}$. Then $1-\beta = \beta(\boldsymbol{\theta}|H_1) = P(T_2/\sigma_{T_2} \le \gamma z_\alpha - \delta) =$ 1 - $\Phi(\gamma z_\alpha - \delta)$, where $\Phi(\cdot)$ is the distribution function of $N(0,\ 1)$, $\gamma^2 = 1/(1+\eta_{12}^2)$ and $\delta^2 = \eta_{12}^2/(1+\eta_{12}^2)$. With $\eta_{12} \in (0,1]$ under H_1, we have $1-\beta \to 1$ as $n, p_i \to \infty$. A similar behavior can be shown for local power, taking $\boldsymbol{\Sigma}_{12} = \mathbf{A}_{12}/\sqrt{n}$ with $\tau_2 = \|\mathbf{A}_{12}\|^2/n$ where $\mathbf{A}_{12} \ge 0$ is any fixed matrix.

2.2. Extension to b blocks

For the general case, consider $\mathbf{X}_k = (\mathbf{X}'_{k1}, \ldots, \mathbf{X}'_{kb})'$ with $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{ij})_{i,j=1}^{b}$; see 1.1. To extend T_2 for b blocks, let $B_{krlsij} = A_{lskr1ij} + A_{lksrij} + A_{lrksij}$ (Equation (3)). We define the general statistic as

$$T_b = \sum_{\substack{i=1\\ i<j}}^{b}\sum_{j=1}^{b} T_{ij} \quad \text{with} \quad T_{ij} = \frac{1}{P(n)}\sum_{\substack{k=1\\ \pi(k,}}^{n}\sum_{\substack{r=1\\ r,}}^{n}\sum_{\substack{l=1\\ l,}}^{n}\sum_{\substack{s=1\\ s)}}^{n}\frac{1}{12p_ip_j}B_{krlsij}, \tag{9}$$

where $E(T_b) = \sum_{i<j}\|\boldsymbol{\Sigma}_{ij}\|^2$, which is 0 under H_0, and

$$\begin{aligned}\operatorname{Var}(T_b) = &\sum_{\substack{i=1\\ i<j}}^{b}\sum_{j=1}^{b}\operatorname{Var}(T_{ij}) + 2\sum_{\substack{i=1\\ i<j<j'}}^{b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\operatorname{Cov}(T_{ij}, T_{ij'})\\ &+2\sum_{\substack{i=1\\ i<i'<j}}^{b}\sum_{i'=1}^{b}\sum_{j=1}^{b}\operatorname{Cov}(T_{ij}, T_{i'j}) + \sum_{\substack{i=1\\ i<i'<j<j'}}^{b}\sum_{i'=1}^{b}\sum_{j=1}^{b}\sum_{j'=1}^{b}\operatorname{Cov}(T_{ij}, T_{i'j'}).\end{aligned} \tag{10}$$

$\operatorname{Var}(T_{ij})$ follows from Equation (4) and $\operatorname{Cov}(T_{ij}, T_{ij'})$, $\operatorname{Cov}(T_{ij}, T_{i'j})$, $\operatorname{Cov}(T_{ij}, T_{i'j'})$, say C_1, C_2, C_3, respectively, are given as following; see also Theorem 2.

$$C_1 = \frac{4}{P(n)p_i^2p_jp_{j'}}\left[2a_1(n)\omega_{ijij'} + d_1(n)\omega_{iijj'} + (n-4)\left\{2\|\boldsymbol{\Sigma}_{ij}\|^2\|\boldsymbol{\Sigma}_{ij'}\|^2 + 5\|\boldsymbol{\Sigma}_{jj'}\|^2\|\boldsymbol{\Sigma}_{ii}\|^2\right\}\right] \tag{11}$$

$$C_2 = \frac{4}{P(n)p_i^2p_jp_{i'}}\left[2a_1(n)\omega_{ii'ij} + d_1(n)\omega_{iii'j} + (n-4)\left\{2\|\boldsymbol{\Sigma}_{ij}\|^2\|\boldsymbol{\Sigma}_{i'j}\|^2 + 5\|\boldsymbol{\Sigma}_{ii'}\|^2\|\boldsymbol{\Sigma}_{jj}\|^2\right\}\right] \tag{12}$$

$$\begin{aligned}C_3 = \frac{4}{P(n)p_ip_jp_{i'}p_{j'}}&\Big[2a_2(n)\omega_{iji'j'} + b_2(n)\omega_{ijj'i'} + (4n-11)\omega_{ii'jj'} + 3(n-3)\|\boldsymbol{\Sigma}_{i'j}\|^2\|\boldsymbol{\Sigma}_{ij'}\|^2\\ &+(3n-10)\|\boldsymbol{\Sigma}_{ii'}\|^2\|\boldsymbol{\Sigma}_{jj'}\|^2\Big]\end{aligned} \tag{13}$$

where $a_1(n) = 3n^3 - 38n^2 + 170n - 262$, $b_1(n) = (n-4)(6n^2 - 47n + 104)$, $c_1(n) = 7n^2 - 57n + 117$, $a_2(n) = 3n^3 - 39n^2 + 176n - 269$, $b_2(n) = 6n^3 - 70n^2 + 286n - 199$ and $d_1(n) = \{b_1(n) + 2c_1(n)\}$ and $\omega_{abcd} = \operatorname{tr}(\mathbf{A}_{ab}\mathbf{A}_{bc}\mathbf{A}_{cd}\mathbf{A}_{da})$.

Theorem 2. *For* T_b *in (9),* $\mathrm{E}(\mathrm{T}_b) = \sum_{i<j}^{b} \|\boldsymbol{\Sigma}_{ij}\|^2$ *with* $\mathrm{Var}(\mathrm{T}_b)$ *as in (10). Under* H_0

$$\mathrm{Var}(\mathrm{T}_b) = \frac{2(2n^2 - 12n + 21)}{P(n)} \sum_{\substack{i=1\\ i<j}}^{b} \sum_{j=1}^{b} \frac{1}{p_i^2 p_j^2} \|\boldsymbol{\Sigma}_{ii}\|^2 \|\boldsymbol{\Sigma}_{jj}\|^2 \tag{14}$$

Under Assumptions 1 and 2, $\mathrm{Var}(n\mathrm{T}_b)$ and nC_1, nC_2, nC_3 are uniformly bounded, so that the limit of T_b follows similarly as of T_2. To see this precisely, write $\mathrm{T}_b = \mathbf{1}'\mathbf{T}_B$ with $\mathbf{T}_B = (\mathbf{T}_1, ..., \mathbf{T}_{b-1})'$, $\mathbf{T}_i = (T_{i,i+1}, ..., T_{ib})'$, $i = 1, ..., b-1$, where $B = b(b-1)/2$ and $\mathbf{1}_B$ is the vector of 1 s. Then $\mathrm{E}(\mathrm{T}_b) = \mathbf{1}'\mathbf{T}_B$ and $\mathrm{Var}(\mathrm{T}_b) = \mathbf{1}'\boldsymbol{\Lambda}\mathbf{1}$, where $\mathrm{Cov}(\mathbf{T}_B) = \boldsymbol{\Lambda} = (\boldsymbol{\Lambda}_{ij})_{i,j=1}^{B}$ is a partitioned matrix with diagonals $\boldsymbol{\Lambda}_{ii} : (b-i) \times (b-i)$ and off-diagonals $\boldsymbol{\Lambda}_{ij} = \boldsymbol{\Lambda}'_{ji} : (b-i) \times (b-j)$, $j > i$.

As for $\mathrm{Var}(\mathrm{T}_2)$, elements of $\boldsymbol{\Lambda}$ are uniformly bounded in terms of $\eta_{ij} = \|\boldsymbol{\Sigma}_{ij}\| / \|\boldsymbol{\Sigma}_{ii}\| \|\boldsymbol{\Sigma}_{jj}\|$ as $p_i \to \infty$, under the assumptions. For example, for $b = 3$, $\mathrm{Var}(nT_3) = \mathbf{1}'\boldsymbol{\Lambda}\ \mathbf{1}[1 + \mathrm{O}(1)]$, where $\boldsymbol{\Lambda}$ converges to 4 times a matrix with diagonal elements $1 + \eta_{12}^2$, $1 + \eta_{13}^2$, $1 + \eta_{23}^2$ and off-diagonal elements $\eta_{12}\eta_{13} + \eta_{23}$, $\eta_{12}\eta_{23} + \eta_{13}$, $\eta_{13}\eta_{23} + \eta_{13}$. Thus, under H_0, $\boldsymbol{\Lambda}$ converges to $4\mathbf{I}_3$. The limit of T_b follows now from that of $\mathbf{T}_B$ by Cramér–Wold device (van der Vaart 1998), and is given in Theorem 3 where $\sigma^2_{T_b} = \mathrm{Var}(T_b)$ and $\sigma^2_{T_{b0}} = \mathrm{Var}(T_b)$ under H_0 are given in Equations (10) and (14), respectively.

Theorem 3. *For* T_b *in (9),* $(\mathrm{T}_b - \mathrm{E}(\mathrm{T}_b))/\sigma_{\mathrm{T}_b} \xrightarrow{\mathcal{D}} N(0,1)$ *as* $n, p_i \to \infty$*, under Assumptions 1–2. In particular, under* H_0 *and Assumption 1,* $n\mathrm{T}_b/\sigma_{\mathrm{T}_{b0}} \xrightarrow{\mathcal{D}} N(0,1)$.

3. The non-normal case

Defined as U-statistic, T_b is a nonparametric measure of $\|\boldsymbol{\Sigma}_{12}\|^2$. Further, many of the computations in Section 2 are valid, exactly or asymptotically, without normality. It motivates us to show that T_b and its properties can be used by relaxing normality. Given the notations for $\mathbf{X}_k$ in 1.1, let $\mathbf{Y}_k = (\mathbf{Y}'_{k1}, ..., \mathbf{Y}'_{kb})'$, $\mathbf{Y}_{ki} = \mathbf{X}_{ki} - \boldsymbol{\mu}_i$ with $\mathbf{Z}_k = (\mathbf{Z}'_{k1}, ...\mathbf{Z}'_{kb})$, $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}^{1/2}$. Define the model

$$\mathbf{Y}_k = \boldsymbol{\Gamma}\mathbf{Z}_k, \quad k = 1, ..., n \tag{15}$$

where $\mathbf{Z}_{ki} \sim \mathcal{F}$, $\mathrm{E}(\mathbf{Z}_{ki}) = \mathbf{0}_{p_i}$, $\mathrm{Cov}(\mathbf{Z}_{ki}) = \mathbf{I}_{p_i}$ and $\mathcal{F}$ denotes a distribution function. Model (15) is very general and covers for example, elliptical class including multivariate normal, so that the results in Section 2 are a special case of those under Model (15). To see this precisely, first note that, working under Model (15), we need to control the fourth moment of $\mathcal{F}$ as the computations involve moments of bilinear forms. For this, we define κ_{ij} which is 0 under normality.

$$\kappa_{ij} = \mathrm{E}(A_{ki}A_{kj}) - 2\|\boldsymbol{\Sigma}_{ij}\|^2 - \|\boldsymbol{\Gamma}_{ii}\|^2 \|\boldsymbol{\Gamma}_{jj}\|^2, \tag{16}$$

with $\boldsymbol{\Gamma}_{ii} = \boldsymbol{\Sigma}_{ii}^{1/2}$, $A_{ki} = \mathbf{Y}'_{ik}\mathbf{Y}_{ik}$. To mirror this fact through assumptions, we also let

Assumption 3. $\mathrm{E}(Y_{kis}^4) = \gamma_s \le \gamma_0 < \infty$, $\forall\ s = 1, ..., p$, $\gamma_0 \in \mathbb{R}^+$.

Under this set up, the basic moments under Model (15) are either same as under normality or can be easily extended using κ_{ij}. These moments are given in Theorem 8 where the constant K is used to represent such terms. Note that, these terms also involve Hadamard products like $\mathrm{tr}(\boldsymbol{\Gamma} \odot \boldsymbol{\Gamma})$ for $\mathrm{E}(A_{ik}^2)$, but are suppressed in K since all such terms vanish under Assumptions 1 and 2 whence $\kappa_{12}/\|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2 \to 0$; Under normality, K is exactly 0; see Ahmad (2017b) for details. Now, Theorem 2 can be extended under Model (15), using the results of Theorem 8. For example, Equation (4) extends by an extra term $c(n)KO(1)$ as

$$\mathrm{Var}(\mathrm{T}_2) = \frac{2}{P(n)}\left[4a(n)\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + 2b(n)\omega_{1122} + c(n)\left\{\left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2 + KO(1)\right\}\right] \tag{17}$$

Finally, the proof of main theorem in Section 2 (Theorem 1) is in fact carried out without assuming normality, so that the same proof holds under Model (15); see Appendix A.3. We thus state the following theorem which generalizes all main results of Section 2. Note that, under Model (15), Theorem 4 pertains only to testing of zero correlation, where under normality, all results reduce to those in Section 2, pertaining to testing independence.

Theorem 4. *Theorem 3 and Corollary 1 remain valid for Model (15) under Assumptions 1–3.*

4. Some related tests

4.1. Test of complete independence

For $p_i = 1 \ \forall\ i$, $\mathbf{X}_k = (X_{k1}, \ldots, X_{kp})'$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$, $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1}^p$, $\boldsymbol{\Sigma}_{ij} = \sigma_{ij}$ so that tests of (1) or (2) reduce to complete independence, denoted $H_{0c} : X_{ki} \perp\!\!\!\perp X_{kj} \ \forall\ i \neq j$ vs. $H_{1c} : X_{ki} \not\perp\!\!\!\perp X_{kj}$ for at least one pair $i \neq j$ or, under normality, $H_{0c} : \sigma_{ij} = 0 \ \forall\ i \neq j$ vs. $H_{1c} : \sigma_{ij} \neq 0$ for at least one pair $i \neq j$, $i, j = 1, \ldots, p$, $k = 1, \ldots, n$, $\sigma_{ij} = \mathrm{Cov}(X_{ki}, X_{kj})$. We want to test H_{0c} when $p \gg n$.

For T_b in Section 2, $\mathrm{E}(\mathrm{T}_b) = \sum_{i<j} \sigma_{ij}^2 = 0$ under H_{0c}. We can thus reduce T_b to define a test, say T_c, for H_{0c}. For brevity, we only discuss the test under the null. For $p = 2$, it is the correlation test of $H_0 : \rho = 0$. Denote $b_{krlsij} = a_{lskrij} + a_{lksrij} + a_{lrksij}$, $a_{lskrij} = a_{lskri}a_{krlsj}$, $a_{lskri} = d_{lsi}d_{kri}$, $d_{kri} = X_{ki} - X_{ri}$, $\mathrm{E}(d_{kri}) = 0$, $\mathrm{Var}(d_{kri}) = 2\sigma_{ii}$, $\mathrm{E}(d_{kri}d_{krj}) = 2\sigma_{ij}$, $\mathrm{E}(a_{lskrij}) = 4\sigma_{ij}^2$. Then $\mathrm{T}_c = \sum_{i<j}^p \mathrm{T}_{ij}$ with $\mathrm{E}(\mathrm{T}_{ij}) = \sigma_{ij}^2$ where

$$\mathrm{T}_{ij} = \frac{1}{12P(n)} \underset{\pi(k,\ r,\ l,\ s)}{\sum_{k=1}^n \sum_{r=1}^n \sum_{l=1}^n \sum_{s=1}^n} b_{krlsij}. \tag{18}$$

Now $\mathrm{E}(\mathrm{T}_c) = \sum_{i<j} \sigma_{ij}^2$, $\mathrm{Var}(\mathrm{T}_c) = 2(2n^2 - 12n + 21)\sum_{i<j}^p \sigma_{ii}^2\sigma_{jj}^2/P(n)$. The covariances in (11)–(13) follow similarly, and vanish under H_{0c}. Like T_b, the moments and limit of T_c depend on $\eta_{ij} = \rho_{ij}^2$. Assumptions 1–3 simplify where, under H_{0c}, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp})$, so that $\lambda_i = \sigma_{ii}$ are the eigenvalues of $\boldsymbol{\Sigma}$. By Assumption 1, $\sum_{i<j}^p \sigma_{ii}^2\sigma_{jj}^2/p^2 = O(1)$, (7) gives $b_i = \sum_{\pi(k,r,l,s)} d_{krlsi}^2/12P(n)$ as consistent estimator of σ_{ii}^2, $i = 1, \ldots, p$, and $n\mathrm{T}_c/p$ has finite limit. Theorems 3 and 4 are thus valid with results

reduced for T_c. Theorem 5 gives the null limit under the following assumptions, where ν_i are eigenvalues of $\boldsymbol{\Sigma}/p$ and $\sigma^2_{T_{c0}}$ is $\mathrm{Var}(\mathrm{T}_c)$ under H_{0c}; in particular, $\sigma^2_{20} = 4\sigma^2_{11}\sigma^2_{22}O(1)$.

Assumption 4. For $p \to \infty$, $\sum_{i=1}^p \nu_i = O(1)$.

Assumption 5. *For* $n, p \to \infty$, $p/n \to \delta_1 \in (0, \infty)$.

Theorem 5. *Given* T_c *with* T_{ij} *in (18). Then* $n\mathrm{T}_c/p\sigma_{\mathrm{T}_c} \xrightarrow{\mathcal{D}} N(0,1)$ *as* $n, p \to \infty$, *under* H_{0c} *and Assumptions 1–5. The limit holds if* σ^2_{ii} *are replaced by* b_i *given above.*

4.2. Tests of homogeneity of diagonal blocks

Under H_0 in (1) or (2), $\boldsymbol{\Sigma} = \oplus_{i=1}^b \boldsymbol{\Sigma}_{ii}$ and it might be of interest to test equality of diagonal blocks

$$H_{0h} : \boldsymbol{\Sigma}_{ii} = \boldsymbol{\Sigma}_h \quad \forall\ i \quad \text{vs.} \quad H_{1h} : \boldsymbol{\Sigma}_{ii} \neq \boldsymbol{\Sigma}_h \quad \text{for atleast one } i. \tag{19}$$

For $p_i = q\ \forall\ i$ under H_{0h}, $p = bq$, $\boldsymbol{\Sigma} = \mathbf{I}_b \otimes \boldsymbol{\Sigma}_h$, $\boldsymbol{\Sigma}_h : q \times q$, $\|\boldsymbol{\Gamma}\|^2 = b\|\boldsymbol{\Gamma}_h\|^2$, $\|\boldsymbol{\Sigma}\|^2 = b\|\boldsymbol{\Gamma}_h\|^2$. Further, under $H_{1h}|H_0$, $\|\boldsymbol{\Sigma}\|^2 = \sum_{i=1}^b \|\boldsymbol{\Gamma}_{ii}\|^2$, $\|\boldsymbol{\Sigma}\|^2 = \sum_{i=1}^b \|\boldsymbol{\Sigma}_{ii}\|^2$, where $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}^{1/2}$ etc. (see Section 3). A high-dimensional test of homogeneity of $g \geq 2$ covariance matrices, using Frobenius norm of the difference between null and alternative hypotheses, is given in Ahmad (2017b) for models like (15). With $\mathbf{X}_{ki}$ independent under H_0, the test can be used for H_{0h}, as is briefly explained below.

First let $b = 2$ with $\tau_{2h} = \|\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{22}\|_F^2 = \sum_{i=1}^2 \|\boldsymbol{\Sigma}_{ii}\|^2 - 2\|\boldsymbol{\Gamma}_{11}\boldsymbol{\Gamma}_{22}\|^2$. With E_i and E_{12} as unbiased and consistent estimators of $\|\boldsymbol{\Sigma}_{ii}\|^2$ and $\|\boldsymbol{\Gamma}_{11}\boldsymbol{\Gamma}_{22}\|^2$, respectively, the test statistic is

$$T_{2h} = \sum_{i=1}^2 a_i \tilde{E}_i - 2a_{12}\tilde{E}_{12}$$

where $a_i = \|\boldsymbol{\Sigma}_{ii}\|^2/q^2$, $a_{12} = \|\boldsymbol{\Gamma}_{11}\boldsymbol{\Gamma}_{22}\|^2/q^2$ and $\tilde{E} = E/\mathrm{E}(E) - 1$. Using all pairwise norms, $\tau_{bh} = \sum_{i<j}^b \|\boldsymbol{\Sigma}_{ii} - \boldsymbol{\Sigma}_{jj}\|_F^2$, the statistic for b blocks is

$$\mathrm{T}_{bh} = \sum_{i<j}^b T_{2hij} = 4(b-1)\sum_{i=1}^b T_i - 2\sum_{i<j}^b \mathrm{T}_{ij}$$

with $T_i = a_i\tilde{E}_i$, $\mathrm{T}_{ij} = a_{ij}\tilde{E}_{ij}$, T_{h2ij} is T_{2h} for $(i,\ j)$th pair, E_i is as in (7) and $E_{ij} = \|\hat{\boldsymbol{\Gamma}}_{ii}\hat{\boldsymbol{\Gamma}}_{jj}\|^2$, where $\hat{\boldsymbol{\Sigma}}_{ii} = \sum_{k=1}^n (\mathbf{X}_{ki} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ki} - \bar{\mathbf{X}}_i)'/(n-1)$; see also Section 4.3. Now $\mathrm{E}(\mathrm{T}_{bh}) = 0$ and

$$\begin{aligned}\sigma^2_{\mathrm{T}_{bh}} = {} & (b-1)^2\sum_{i=1}^b a_i^2 \mathrm{Var}(\tilde{E}_i) + 4\sum_{\substack{i=1 \\ i<j}}^b\sum_{j=1}^b a_{ij}^2\mathrm{Var}(\tilde{E}_{ij}) + 8\sum_{\substack{i=1 \\ i<j<j'}}^b\sum_{j=1}^b\sum_{j'=1}^b a_{ij}a_{ij'}\mathrm{Cov}(\tilde{E}_{ij}, \tilde{E}_{ij'}) \\ & + 8\sum_{\substack{i=1 \\ i<i'<j}}^b\sum_{i'=1}^b\sum_{j=1}^b a_{ij}a_{i'j}\mathrm{Cov}(\tilde{E}_{ij}, \tilde{E}_{i'j}) - 8(b-1)\sum_{\substack{i=1 \\ i<j}}^b\sum_{j=1}^b a_i a_{ij}\mathrm{Cov}(\tilde{E}_i, \tilde{E}_{ij})\end{aligned} \tag{20}$$

The moments composing $\sigma^2_{\mathrm{T}_{bh}}$ follow from Ahmad (2017b), where $\mathrm{Var}(\tilde{E}_i)$, $\mathrm{Var}(\tilde{E}_{ij})$ and $\mathrm{Cov}(\tilde{E}_i, \tilde{E}_{ij})$ are each $O(1/n)$ and uniformly bounded in p_i which help determine the limit of T_{bh} under the following assumptions in addition to Assumption 1. Let $\kappa = \|\boldsymbol{\Sigma}_h^2\|^2/[\|\boldsymbol{\Sigma}_h\|^2]^2$. For $n, p_i \to \infty$

Assumption 6. $\inf_p \kappa \gg 0$.

Assumption 7. $p_i/n \to \delta_i \le \delta_2 \in (0, \infty)$.

Under H_{0h}, with $a_h = \|\boldsymbol{\Sigma}_h\|^2$, we have

$$\mathrm{T}_{bh} = a_h \left[(b-1) \sum_{i=1}^{b} \tilde{E}_i - 2 \sum_{i \neq j}^{b} \tilde{E}_{ij} \right]$$

where $\sigma^2_{T_{bh0}} \approx 4a_h^2 b^2(b-1)/n^2$ with consistent estimator $\hat{\sigma}^2_{T_{bh0}}$ using $\hat{a}_h = \sum_{i \neq j}^{b} E_{ij}/b(b-1)$.

Theorem 6. *Given T_{bh}, $\sigma^2_{\mathrm{T}_{bh}}$ and $\sigma^2_{\mathrm{T}_{bh0}}$. Then $\sigma^{-1}_{\mathrm{T}_{bh}}(\mathrm{T}_{bh} - \tau_{bh}) \xrightarrow{\mathcal{D}} N(0,1)$ as $n, p_i \to \infty$, under Assumptions 1–2 and 6–7. In particular, under H_{0h}, $\sigma^{-1}_{\mathrm{T}_{bh0}} \mathrm{T}_{bh} \xrightarrow{\mathcal{D}} N(0,1)$. Further, the limits hold by replacing a_h with $\hat{a}_h$ in $\sigma^2_{\mathrm{T}_{bh}}$ as defined above.*

4.3. Alternative form and computational efficiency

Test statistics in Sections 2 and 3 are defined using Frobenius norm $\tau = \|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_D\|^2$ in terms of cross-covariance operator $\boldsymbol{\Sigma}_{12}$. But, since $\tau = \|\boldsymbol{\Sigma}\|^2 - \|\boldsymbol{\Sigma}_D\|^2 = \|\boldsymbol{\Sigma}\|^2 - \sum_{i=1}^{b} \|\boldsymbol{\Sigma}_{ii}\|^2$, same tests can also be defined using unbiased and consistent estimators, say E_b and E_i, of $\|\boldsymbol{\Sigma}\|^2$ and $\|\boldsymbol{\Sigma}_{ii}\|^2$, respectively. Then we can define $\mathrm{T}_b = \tilde{E}_b - \tilde{E}_0$ with $\tilde{E}_0 = \sum_{i=1}^{b} E_i/p_i^2$ and $\tilde{E}_b = E_b/p^2$ where E_i is given in Equation (7) and likewise (see Ahmad 2017a) $E_b = \sum_{\pi(k,r,l,s)} D^2_{krls}/12P(n)$ with $D^2_{krls} = A^2_{krls} + A^2_{klrs} + A^2_{ksrl}$, $A_{krls} = \mathbf{D}'_{kr}\mathbf{D}_{ls}$, $\mathbf{D}_{kr} = \mathbf{X}_k - \mathbf{X}_r$. Under H_0, both E_b and E_0 estimate $\sum_{i=1}^{b} \|\boldsymbol{\Sigma}_{ii}\|^2$ and the properties of the tests remain same as given above.

A final remark concerns the estimators. All estimators are defined as U-statistics of symmetric kernels which help us study their properties and those of the test statistics. For computational efficiency and practical use, however, same estimators can be defined in a much simpler way, as functions of sample covariance matrices, $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\Sigma}}_{ij}$. We provide these estimators for T_b in Sections 2 and 3; see also Ahmad (2017a, 2017b). Let $\bar{\mathbf{X}} = \sum_{k=1}^{n} \mathbf{X}_k/n$ and $\hat{\boldsymbol{\Sigma}} = \sum_{k=1}^{n} \tilde{\mathbf{X}}_k \tilde{\mathbf{X}}'_k/(n-1)$ be unbiased estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with $\tilde{\mathbf{X}} = \mathbf{X}_i - \bar{\mathbf{X}}$. Define $Q = \sum_{k=1}^{n} (\tilde{\mathbf{X}}'_k \tilde{\mathbf{X}}_k)^2/(n-1)$. Similarly define Q_i using $\hat{\boldsymbol{\Sigma}}_{ii}$ and $\tilde{\mathbf{X}}_{ki}$, $i = 1, \ldots, b$. Then, we can define $E_b = \eta\{2\|\hat{\boldsymbol{\Sigma}}\|^2 + (n^2 - 3n + 1)[\|\hat{\boldsymbol{\Sigma}}\|]^2 - nQ\}$ and $E_i = \eta\{2\|\hat{\boldsymbol{\Sigma}}_{ii}\|^2 + (n^2 - 3n + 1)[\|\hat{\boldsymbol{\Sigma}}_{ii}\|]^2 - nQ_i\}$, where $\eta = (n-1)/[n(n-2)(n-3)]$.

5. Simulations

We evaluate the performance of T_b under Model (15). We generate $n = \{10, 20, 50\}$ iid vectors of dimension $p = \{60, 100, 300, 500, 1000\}$ from normal, uniform and t_5

Table 1. Estimated $1 - \alpha$ of T_2 for normal distribution.

Σ_{ii}	n	$1-\alpha$	p = 60	100	300	500	1000
CS	10	0.90	0.908	0.916	0.915	0.913	0.908
		0.95	0.945	0.947	0.948	0.946	0.951
		0.99	0.983	0.984	0.981	0.980	0.983
	20	0.90	0.916	0.913	0.914	0.912	0.907
		0.95	0.958	0.953	0.955	0.951	0.948
		0.99	0.985	0.978	0.981	0.980	0.983
AR-I	10	0.90	0.918	0.920	0.916	0.915	0.913
		0.95	0.952	0.956	0.954	0.955	0.952
		0.99	0.985	0.987	0.989	0.989	0.990
	20	0.90	0.914	0.906	0.908	0.914	0.907
		0.95	0.960	0.953	0.954	0.952	0.948
		0.99	0.988	0.986	0.985	0.985	0.989
AR-II	10	0.90	0.915	0.919	0.908	0.907	0.904
		0.95	0.945	0.948	0.942	0.944	0.946
		0.99	0.981	0.981	0.980	0.978	0.984
	20	0.90	0.916	0.913	0.914	0.918	0.907
		0.95	0.955	0.957	0.958	0.958	0.955
		0.99	0.980	0.981	0.981	0.983	0.988

Table 2. Estimated $1 - \alpha$ of T_2 for uniform distribution.

Σ_{ii}	n	$1-\alpha$	p = 60	100	300	500	1000
CS	10	0.90	0.916	0.914	0.914	0.916	0.909
		0.95	0.947	0.949	0.948	0.945	0.944
		0.99	0.982	0.982	0.980	0.979	0.985
	20	0.90	0.0.912	0.916	0.911	0.913	0.910
		0.95	0.957	0.952	0.954	0.957	0.954
		0.99	0.982	0.980	0.980	0.982	0.988
AR-I	10	0.90	0.916	0.912	0.914	0.911	0.907
		0.95	0.954	0.954	0.951	0.952	0.955
		0.99	0.989	0.986	0.988	0.987	0.985
	20	0.90	0.916	0.915	0.913	0.908	0.906
		0.95	0.961	0.958	0.962	0.956	0.952
		0.99	0.987	0.986	0.985	0.982	0.991
AR-II	10	0.90	0.916	0.907	0.913	0.912	0.910
		0.95	0.946	0.940	0.945	0.946	0.949
		0.99	0.976	0.973	0.980	0.982	0.981
	20	0.90	0.916	0.911	0.914	0.914	0.907
		0.95	0.954	0.957	0.955	0.956	0.955
		0.99	0.981	0.980	0.980	0.981	0.984

distributions, assuming $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma}$ as CS and AR(1) with equal and unequal p_i. Under H_0, $\boldsymbol{\Sigma} = \oplus_{i=1}^{2}\boldsymbol{\Sigma}_{ii}$ with CS and AR block diagonals defined as $\boldsymbol{\Sigma}_{ii} = (1-\rho)\mathbf{I}_{p_i} + \mathbf{J}_{p_i}$ (Section 2.1) and $\boldsymbol{\Sigma}_{ii} = \mathbf{BAB}$ with $\mathbf{A} = \rho^{|k-l|^{1/5}}$ and $\mathbf{B}$ a diagonal matrix with entries square roots of $\rho + (1:p)/p$. For unequal p_i, $\mathbf{B}_i$ has elements $\rho + (1:p_i)/p_i$. Under H_1, same structures are imposed on $\boldsymbol{\Sigma}$; for example, $\boldsymbol{\Sigma} = (1-\rho)\mathbf{I}_p + \mathbf{J}_p$ with $\rho = 0.3$, $\boldsymbol{\Sigma} = \mathbf{B}_i\mathbf{A}\mathbf{B}_j + 0.3\mathbf{J}_{p_i \times p_j}$, $\mathbf{J}_{p_i \times p_j} = \mathbf{1}_{p_i}\mathbf{1}'_{p_j}$, $i, j = 1, 2$.

The size and power are estimated as averages, over 5000 simulation runs, of $\alpha = P(T_z \geq Z|H_0)$ and $1 - \beta = P(T_z \geq Z|H_1)$, where T_z denotes standardized T_b. We use $\alpha = \{0.01, 0.05, 0.10\}$ for size and $\alpha = 0.05$ for power. Tables 1–3 report estimated test sizes where Table 4 reports estimated power, where AR-I and AR-II denote AR(1) structures with equal and unequal blocks, respectively.

Table 3. Estimated $1-\alpha$ of T_2 for t_5-distribution.

			p				
Σ_{ii}	n	$1-\alpha$	60	100	300	500	1000
CS	10	0.90	0.905	0.901	0.907	0.902	0.905
		0.95	0.940	0.941	0.946	0.948	0.949
		0.99	0.978	0.979	0.978	0.973	0.980
	20	0.90	0.919	0.914	0.909	0.912	0.908
		0.95	0.946	0.944	0.940	0.942	0.948
		0.99	0.975	0.973	0.976	0.980	0.982
AR-I	10	0.90	0.908	0.904	0.906	0.898	0.901
		0.95	0.946	0.944	0.945	0.944	0.950
		0.99	0.984	0.983	0.982	0.983	0.986
	20	0.90	0.914	0.910	0.907	0.910	0.904
		0.95	0.948	0.943	0.940	0.954	0.941
		0.99	0.978	0.977	0.986	0.981	0.982
AR-II	10	0.90	0.901	0.908	0.901	0.900	0.903
		0.95	0.940	0.941	0.944	0.947	0.949
		0.99	0.979	0.980	0.978	0.983	0.981
	20	0.90	0.912	0.911	0.908	0.906	0.904
		0.95	0.944	0.941	0.940	0.943	0.945
		0.99	0.978	0.983	0.981	0.979	0.982

Table 4. Estimated $1-\beta$ of T_b: all distributions.

			p				
$\mathcal{F}$	Σ	n	60	100	300	500	1000
Normal	CS	10	0.654	0.852	0.970	0.990	0.997
		20	0.960	0.996	1.000	1.000	1.000
	AR-I	10	0.754	0.876	0.971	0.987	1.000
		20	0.984	0.996	1.000	1.000	1.000
	AR-II	10	0.214	0.244	0.274	0.277	0.302
		20	0.405	0.434	0.468	0.501	0.513
		50	0.835	0.845	0.890	0.908	0.921
Uniform	CS	10	0.650	0.845	0.970	0.990	1.000
		20	0.952	0.995	1.000	1.000	1.000
	AR-I	10	0.757	0.870	0.976	0.987	1.000
		20	0.981	0.998	1.000	1.000	1.000
	AR-II	10	0.209	0.228	0.259	0.262	0.280
		20	0.373	0.446	0.512	0.528	0.531
		50	0.799	0.828	0.856	0.887	0.908
t_5	CS	10	0.655	0.823	0.965	0.984	1.000
		20	0.945	0.992	1.000	1.000	1.000
	AR-I	10	0.744	0.850	0.962	0.982	0.999
		20	0.972	0.995	1.000	1.000	1.000
	AR-II	10	0.235	0.248	0.275	0.280	0.295
		20	0.424	0.459	0.485	0.497	0.511
		50	0.802	0.806	0.873	0.877	0.898

We observe an accurate size control for all distributions under all parameters, particularly for moderate n and increasing p_i. We also notice strong robustness of T_b to normality. Moreover, the power increases not only with increasing n but also with increasing p_i for all other parameters. We also investigated the test for other ρ values, for example, 0.2 or 0.85, and found that the power improves slightly for higher ρ and drops slightly for smaller ρ. For unequal p_i in AR, we notice a slight decline and relatively slow growth in power. For this, we add results for $n = 50$ which indicates an increasing power with p_i, particularly improving with increasing n.

6. Applications

To demonstrate an application of the proposed test, we use the well-known COMBO galaxy data set. The data set provides classification of $n = 3438$ astronomical objects as galaxies, based on the information on 29 variables measured on each object. The variables are grouped into two vectors, with $p_1 = 23$ and $p_2 = 6$ dimensions. We denote the vectors as $\mathbf{X}_{k1} \in \mathbb{R}^{p_1}$ and $\mathbf{X}_{k2} \in \mathbb{R}^{p_2}$, respectively, so that $\mathbf{X}_k = (\mathbf{X}'_{k1},\ \mathbf{X}'_{k2})' \in \mathbb{R}^p$, $p = p_1 + p_2 = 29$. Since the objects are independent, we once use all 3438 objects and once take a random sample of 10 objects whence $n \ll p$, in order to show the application of the test statistic for high-dimensional case.

For full data, we compute the test statistic (see Section 2.1) as 236.4 with p-value virtually 0, seriously rejecting the null hypothesis. For the subset of the data, with $n = 10$, the test statistic is 0.364 with p-value 0.3579, providing no sufficient evidence to reject the null hypothesis.

7. Discussion and conclusions

Correlation tests for two or more vectors are proposed when the data are high-dimensional and the distribution may not be normal. Properties of the tests are studies and their asymptotic distributions are derived under certain mild assumptions. Some subsequent tests are also discussed. All tests are defined as functions of U-statistics based estimators of the Frobenius norm between the null and alternative hypotheses. Simulation results are used to show the accuracy of the proposed tests for a general multivariate class of distributions including multivariate normal.

Acknowledgements

The authors are thankful to the editor and the referees for their comments which helped improve the original version of the manuscript.

Funding

Prof. S. Ejaz Ahmed's research is supported by the Natural Sciences and the Engineering Research Council of Canada (NSERC).

ORCID

M. Rauf Ahmad http://orcid.org/0000-0002-5362-5835

References

Ahmad, M. R. 2017a. Location-invariant multi-sample U-tests for covariance matrices with large dimension. Scandinavian *Journal of Statistics.* 44 (2):500–23.

Ahmad, M. R. 2017b. Location-invariant tests of homogeneity of large dimensional covariance matrices. *Journal of Statistical Theory and Practice* 11 (4):731–45. doi:10.1080/15598608.2017.1308895.

Ahmad, M. R. 2019. A significance test of the RV coefficient for large dimensions. *Computational Statistics and Data Analysis*131:116–30.
Albert, M., Y. Bouret, M. Fromont, and P. Reynaud-Bouret. 2015. Bootstrap and permutation tests of independence for pont processes. *Annals of Statistics* 43:2537–64.
Anderson, T. W. 2003. *An introduction to multivariate statistical analysis*, 3rd ed. NY: Wiley.
Anderson, T. W. 1999. Asymptotic theory for canonical correlation analysis. *Journal of Multivariate Analysis* 70 (1):1–29. doi:10.1006/jmva.1999.1810.
Anderson, N. H., P. Hall, and D. M. Titterington. 1994. Two-sample test statistics for measuring discrepencies between two multivariate probability density functions using kernel based density estimates. *Journal of Multivariate Analysis* 50 (1):41–54. doi:10.1006/jmva.1994.1033.
Bao, Z., J. Hu, G. Pan, and W. Zhou. 2017. Test of independence for high-dimensional random vectors based on freeness in block correlation matrices. *Electronic Journal of Statistics* 11: 1527–48.
Eaton, M. L., and D. Tyler. 1994. The asymptotic distribution of singular values with applications to canonical correlations and correspondence analysis. *Journal of Multivariate Analysis* 50 (2): 238–64. doi:10.1006/jmva.1994.1041.
Gretton, A., and L. Györfi. 2010. Consistent nonparametric tests of independence. *Journal of Machine Learning Research* 11:1391–423.
Harville, D. A. 2008. *Matrix algebra from a statistician's perspective*. NY: Springer.
Horváth, L., M. Hušková, and G. Rice. 2013. Test of independence for functional data. *Journal of Multivariate Analysis* 117:100–19. doi:10.1016/j.jmva.2013.02.005.
Koroljuk, V. S., and Y. V. Borovskich. 1994. *Theory of U-statistics*. Dordrecht: Kluwer Press.
Lehmann, E. L. 1999. *Elements of large-sample theory*. NY: Springer.
Mao, G. 2020. On high-dimensional tests for mutual independence based on Pearson's correlation coefficient. *Communications in Statistics - Theory and Methods* 49 (14):3572–84. doi:10.1080/03610926.2019.1593459.
Muirhead, R. J. 2005. *Aspects of multivariate statistical theory*. NY: Wiley.
Muirhead, R. J., and C. M. Waternaux. 1980. Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika* 67 (1): 31–43. doi:10.1093/biomet/67.1.31.
Nkiet, G. M. 2017. Asymptotic Theory of Multiple-Set Linear Canonical Analysis. *Mathematical Methods of Statistics* 26 (3):196–211. doi:10.3103/S1066530717030036.
Pfister, N., P. Bühlmann, B. Schülkopf, and J. Peters. 2018. Kernel-based tests for joint independence. *JRSS B* 80 (1):5–31. doi:10.1111/rssb.12235.
Schott, J. R. 2005. Testing for complete independence in high dimensions. *Biometrika* 92 (4): 951–56. doi:10.1093/biomet/92.4.951.
Schott, J. R. 2008. A test for independence of two sets of variables when the number of variables is large relative to the sample size. *Statistics & Probability Letters* 78 (17):3096–102. doi:10.1016/j.spl.2008.05.031.
Searle, S. R. 1971. *Linear models*. NY: Wiley.
Serfling, R. J. 1980. *Approximation theorems of mathematical statistics*. NY: Wiley.
Srivastava, M. S., T. Kollo, and D. von Rosen. 2011. Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis* 102 (6):1090–103. doi:10.1016/j.jmva.2011.03.003.
van der Vaart, A. W. 1998. *Asymptotic statistics*. Cambridge: Cambridge University Press.
Xu, K. 2017. Testing diagonality of high-dimensional covariance matrix under non-normality. *Journal of Statistical Computation and Simulation* 87 (16):3208–24. doi:10.1080/00949655.2017.1362405.
Yang, Y., and G. Pan. 2015. Independence test for high dimensional data based on regularized canonical correlation coefficients. *Annals of Statistics* 43:467–500.

Appendix
A. Miscellaneous results and proofs

A.1. quadratic and bilinear forms

For the following moments, see for example, (Searle 1971, Chap 2).

Theorem 7. *Let* $\mathbf{u}=(\mathbf{u}_1',\ldots,\mathbf{u}_4')'\sim\mathcal{N}(\mathbf{0},\boldsymbol{\Sigma})$, $\mathbf{u}_i\in\mathbb{R}^{p_i}$, $\boldsymbol{\Sigma}=(\boldsymbol{\Sigma}_{ij})_{i,j=1}^{4}$, $\mathrm{Cov}(\mathbf{u}_i,\mathbf{u}_j)=\boldsymbol{\Sigma}_{ij}$. *For symmetric* $\boldsymbol{A}_p$, $\boldsymbol{B}_p$, *let* $Q_i=\mathbf{u}_i'\mathbf{A}\mathbf{u}_i$, $B_{ij}=\mathbf{u}_i'\mathbf{A}\mathbf{u}_j$. *Then* $\mathrm{E}(Q_1)=tr(\mathbf{A}\boldsymbol{\Sigma}_{11})$, $\mathrm{Var}(Q_1)=2tr(\mathbf{A}\boldsymbol{\Sigma}_{11})^2$, $\mathrm{E}(B_{12})=\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{21})$, $\mathrm{Var}(B_{12})=\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{21})^2+\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{22}\mathbf{A}'\boldsymbol{\Sigma}_{11})$, $\mathrm{Cov}(B_{12},B_{34})=\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{23}\mathbf{B}\boldsymbol{\Sigma}_{41})+\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{24}\mathbf{B}'\boldsymbol{\Sigma}_{31})$, $\mathrm{Cov}(Q_1,Q_2)=\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{12}\mathbf{B}'\boldsymbol{\Sigma}_{21})$.

The results of Theorem 7 can be used to derive extended moments of quadratic and bilinear forms, as given in the following lemma.

Lemma 1. *Let* $\mathbf{a}_{ti}\sim\mathcal{N}_p(\mathbf{0},\boldsymbol{\Sigma}_{ii})$, $\mathrm{Cov}(\mathbf{a}_{ti},\mathbf{a}_{sj})=\boldsymbol{\Sigma}_{ij}$ *($t=s$) or 0 ($t\neq s$). Then*

$$\mathrm{Cov}(\mathbf{D}_{uvi}'\boldsymbol{\Sigma}_{ij}\mathbf{D}_{uvj},\mathbf{a}_{ui}'\boldsymbol{\Sigma}_{ij}\mathbf{a}_{vj})=\omega_{iijj}+\|\boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{ji}\|^2 \tag{21}$$

$$\mathrm{Cov}(\mathbf{D}_{uvi}'\boldsymbol{\Sigma}_{ij}\mathbf{D}_{uvj},\mathbf{a}_{ui}'\mathbf{a}_{iv}\mathbf{a}_{vj}'\mathbf{a}_{vj})=\omega_{iijj}+2\|\boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{ji}\|^2 \tag{22}$$

$$\mathrm{Cov}(\mathbf{D}_{tui}'\boldsymbol{\Sigma}_{ij}\mathbf{D}_{tuj},\mathbf{a}_{ti}'\mathbf{D}_{uvj}\mathbf{D}_{uvj}'\mathbf{a}_{tj})=2\omega_{iijj}+3\|\boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{ji}\|^2 \tag{23}$$

$$\mathrm{Var}(\mathbf{a}_{ui}'\mathbf{a}_{vi}\mathbf{a}_{vj}'\mathbf{a}_{uj})=4\omega_{iijj}+2\|\boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{ji}\|^2+\left[\|\boldsymbol{\Sigma}_{ij}\|^2\right]^2+\|\boldsymbol{\Sigma}_{ii}\|^2\|\boldsymbol{\Sigma}_{jj}\|^2. \tag{24}$$

(21) *also holds for* $\mathrm{Cov}(\mathbf{a}_{ui}'\boldsymbol{\Sigma}_{ij}\mathbf{a}_{uj},\mathbf{a}_{ui}'\mathbf{a}_{vi}\mathbf{a}_{uj}'\mathbf{a}_{vj})$, $\mathrm{Cov}(\mathbf{a}_{ti}'\mathbf{a}_{ui}\mathbf{a}_{tj}'\mathbf{a}_{uj},\mathbf{a}_{ui}'\mathbf{a}_{vi}\mathbf{a}_{uj}'\mathbf{a}_{vj})$. *Covariances like* $\mathrm{Cov}(\mathbf{a}_{ui}'\boldsymbol{\Sigma}_{ij}\mathbf{a}_{vj},\ \mathbf{a}_{ui}'\mathbf{a}_{vi}\mathbf{a}_{uj}'\mathbf{a}_{vj})$, $\mathrm{Cov}(\mathbf{a}_{ti}'\mathbf{a}_{ui}\mathbf{a}_{tj}'\mathbf{a}_{uj},\ \mathbf{a}_{ui}'\mathbf{a}_{vi}\mathbf{a}_{tj}'\mathbf{a}_{vj})$ *vanish.*

For $\mathbf{Y}_{ik}$ in Equation (15), let $A_{ki}=\mathbf{Y}_{ki}'\mathbf{Y}_{ki}$, $A_{kli}=\mathbf{Y}_{ki}'\mathbf{Y}_{li}$, $k\neq l$, $B_k=\mathbf{Y}_{k1}'\boldsymbol{\Sigma}_{12}\mathbf{Y}_{k2}$, $B_{kl}=\mathbf{Y}_{k1}'\boldsymbol{\Sigma}_{12}\mathbf{Y}_{l2}$. From Equation (16), $\kappa_{ii}=\mathrm{E}(A_{ki}^2)-2\mathrm{tr}(\boldsymbol{\Sigma}_{ii}^2)-[\mathrm{tr}(\boldsymbol{\Sigma}_{ii})]^2=0$ under normality. Theorem 8 extends the moments further under Model (15), relaxing normality, with constant K involving only κ_{12} and κ_{ii}.

Theorem 8. $\mathrm{E}(A_{ki})=\mathrm{tr}(\boldsymbol{\Sigma}_{ii})$, $\mathrm{E}(A_{kri})=0$, $\mathrm{E}(A_{kli}^2)=\|\boldsymbol{\Sigma}_{ii}\|^2$, $\mathrm{E}(B_k)=\|\boldsymbol{\Sigma}_{12}\|^2$, $\mathrm{E}(B_{kl})=0$, $\mathrm{Var}(B_k)=K+\omega_{1122}$, $\mathrm{Var}(B_{kl})=\omega_{1122}$, $\mathrm{Var}(A_{kl1}A_{lk2})=K+2\omega_{1122}+\|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2$.

A.2. U-Statistics

Here we collect some basic results of U-statistics; for details, see Koroljuk and Borovskich (1994) or Serfling (1980). For iid X_i, let $h(X_1,\ldots,X_m):\mathbb{R}^m\to\mathbb{R}$ denote the kernel of an mth order U-statistic, U_n, with $\mathrm{E}(U_n)=\theta=\mathrm{E}[h(\cdot)]$ with its projection $h_c(x_1,\ldots,x_c)=\mathrm{E}[h(\cdot)|x_1,\ldots,x_c]$, $h_m(\cdot)=h(\cdot)$ and $\xi_c=\mathrm{Var}[h_c(\cdot)$, $c=1,\ldots,m$, so that $\mathrm{Var}(U_n)=\sum_{c=1}^m\binom{m}{c}\binom{n-m}{m-c}\xi_c/\binom{n}{m}$. If $0<\xi_c<\infty$ $\forall\ c$, then $(U_n-\mathrm{E}(U_n))/\sqrt{\mathrm{Var}(U_n)}\xrightarrow{\mathcal{D}}N(0,\,1)$. For two U-statistics, U_{ni}, of order m_i, kernels $h_i(\cdot)$, projections $h_{ic}(\cdot)$, $i=1,\,2$, let $\xi_{cc}=\mathrm{Cov}[h_{1c}(\cdot),h_{2c}(\cdot)]$, $c=1,\ldots,m_1\leq m_2$. Then $\mathrm{Cov}(U_{n1},U_{n2})=\sum_{c=1}^{m_1}\binom{m_2}{c}\binom{n-m_2}{m_1-c}\xi_{cc}/\binom{n}{m_1}$. Let $U_{n_1n_2}$ be a U-statistic of two independent samples, with kernel $h(X_{11},\ldots,X_{1m_1},X_{21},\ldots,X_{2m_2})$, symmetric in each sample, projection $h_{c_1c_2}=\mathrm{E}[h(\cdot)|X_{11},\ldots,X_{1c_1};X_{21},\ldots,X_{2c_2}]$, $\xi_{c_1c_2}=\mathrm{Cov}[h(\cdot),h_{c_1c_2}(\cdot)]$, $\xi_{00}=0$, $c_i=0,1,\ldots,m_i$. If $0\leq$

$n_i/n \leq 1$, $n = n_1 + n_2$, $0 < \xi_{c_1c_2} < \infty \;\; \forall \; c_i$, then $(U_{n_1n_2} - \mathrm{E}(U_{n_1n_2}))/\sqrt{\mathrm{Var}(U_{n_1n_2})} \xrightarrow{\mathcal{D}} N(0,\; 1)$ where $\mathrm{Var}(U_{n_1n_2}) = \sum_{c_1=0}^{m_1}\sum_{c_2=0}^{m_2}\binom{m_1}{c_1}\binom{n_1-m_1}{m_1-c_1}\binom{m_2}{c_2}\binom{n_2-m_2}{m_2-c_2}\xi_{c_1c_2}/\binom{n_1}{m_1}\binom{n_2}{m_2}$.

A.3. Proof of Theorem 1

Consider T_2 in Equation (3) with kernel $h(\cdot) = B_{lskr12}/p_1p_2$, degenerate under H_0. By the asymptotic theory of U-statistics (van der Vaart 1998), $n^{c/2}U_n$ has a non-degenerate limit with variance $c!\xi_c$, with c the least value for which $h(\cdot)$ is non-degenerate. In our case $c = 2$ so that nU_n has a limit. Further, $h(\cdot)$ varies with n (and p_i through n). Many authors have considered U-statistics with varying kernels; see Anderson, Hall, and Titterington (1994) or Koroljuk and Borovskich (1994).

The key point in the limit of nU_n is the behavior of ξ_c. For H_1, all ξ_c, $c = 1, \ldots, 4$, are bounded under the assumptions; see Section A.5. Then, by Equation (6), the second term is bounded by the first which in turn vanishes as $p_i \to \infty$, under A1-A2. Thus $\mathrm{Var}[n(T_\eta - \|\boldsymbol{\Sigma}_{12}\|^2)/\varphi] = O(1)$ with $\varphi^2 = \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2/p_1^2p_2^2$. The limit follows from (Lehmann 1999, Theorem 6.1.2). Under H_0, $\xi_1 = 0$ and $h(\cdot)$ is first-order degenerate. From Section A.2 and Equation 5, $\varphi^2 = \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2/p_1^2p_2^2$ keeps $\xi_2 > 0 \Rightarrow \mathrm{Var}(\mathrm{T}_2) > 0$. With $h(\cdot) : \mathbb{R}^{p_1 \times p_2} \to \mathbb{R}$ square-integrable function, composed of inner products, we have $h(\cdot) \in \mathcal{L}_2(\mathcal{H})$, where $\mathcal{H}$ is the Hilbert space and $\mathcal{L}_2(\cdot)$ is the space of square-integrable random variables. We write

$$A_{lskr12} = \left[\left(\mathbf{D}'_{ls1} \otimes \mathbf{D}'_{kr1}\right)\left(\mathbf{D}_{kr2} \otimes \mathbf{D}_{ls2}\right)\right]/12p_1p_2 \;\; = \;\; \mathrm{tr}(\mathbf{D}_{kr1}\mathbf{D}'_{ls1} \otimes \mathbf{D}_{ls2}\mathbf{D}'_{kr2})/p_1p_2,$$

similarly other parts of $h(\cdot)$, where the components are independent under H_0 with variance ξ_2. Further, by the properties of Kronecker product (Harville 2008, Ch. 16), $\gamma_{s12} = \nu_{s1}\nu_{s2}$ are the eigenvalues of $\boldsymbol{\Gamma}_{12} = (\boldsymbol{\Sigma}_{11} \otimes \boldsymbol{\Sigma}_{22})/p_1p_2$, where $\nu_{si} = \lambda_{si}/p_i$ are the eigenvalues of $\boldsymbol{\Sigma}_{ii}/p_i$. Thus ξ_2 corresponds to $\sum_{s1=1}^{p_1}\nu_{s1}^2\sum_{s2=1}^{p_2}\nu_{s2}^2$ and the eigenvalues of the kernel correspond to those of $\boldsymbol{\Gamma}_{12}$ since $\sum_s \gamma_{s12}^2 = \mathrm{tr}(\boldsymbol{\Gamma}_{12}^2) = \sum_{s1=1}^{p_1}\nu_{s1}^2\sum_{s2=1}^{p_2}\nu_{s2}^2$. Under this set up, $h(\cdot)$ is a Hilbert-Schmidt (product) kernel (Serfling 1980) with an orthonormal decomposition, where the weak convergence of such a kernel is given as

$$n(\mathrm{T}_2 - \mathrm{E}(\mathrm{T}_2)) \xrightarrow{\mathcal{D}} \sum_{s1=1}^{\infty}\sum_{s2=1}^{\infty}\nu_{s1}\nu_{s2}(Z_{s1s2}^2 - 1) \tag{25}$$

with Z_{s1s2} a sequence of independent $N(0,\;1)$ variables. The normal limit follows by an application of Lindeberg-Feller CLT for triangular arrays (see also Ahmad 2017a).

A.4. Proof of Corollary 1

The proof follows by showing the consistency of $\widehat{\|\boldsymbol{\Sigma}_{ii}\|^2} = E_i$ (Section 4.3) under Model (15). With κ_{12} as in Equation (16), this immediately follows from Ahmad (2017a)

$$\mathrm{Var}(E_i) = \frac{4}{P(n)}\left[(2n^3 - 9n^2 + 9n - 16)\|\boldsymbol{\Sigma}_{ii}^2\|^2 + (n^2 - 3n + 8)\left[\|\boldsymbol{\Sigma}_{ii}^2\|\right]^2 + \kappa_{12}O(n^2)\right].$$

under Assumptions 1–3, as $n, p_i \to \infty$. This proves Equation (8) and the corollary.

A.5. Proof of Theorem 2

From Equation (3) and Section A.2, $h(\mathbf{a}_k, \mathbf{a}_r, \mathbf{a}_l, \mathbf{a}_s) = A_{lskr1}A_{krls2} + A_{lksr1}A_{srlk2} + A_{lrks1}A_{kslr2}$ with projection $h_c(\cdot) = \mathrm{E}[h(\cdot)|\mathbf{a}_l, \ldots]$, $c = 1, \ldots, 4$, so that

$$h_1(\cdot) = 6\left[(\mathbf{a}'_{1l}\boldsymbol{\Sigma}_{12}\mathbf{a}_{l2}) + \|\boldsymbol{\Sigma}_{12}\|^2\right]$$
$$h_2(\cdot) = 2\left[\mathbf{d}'_{ls1}\boldsymbol{\Sigma}_{12}\mathbf{d}_{ls2} + \mathbf{a}'_{l1}\boldsymbol{\Sigma}_{12}\mathbf{a}_{l2} + \mathbf{a}'_{s1}\boldsymbol{\Sigma}_{12}\mathbf{a}_{l2} + \mathbf{a}'_{l1}\mathbf{a}_{s1}\mathbf{a}'_{s2}\mathbf{a}_{l2} + \mathrm{tr}(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21})\right]$$
$$h_3(\cdot) = \mathbf{d}'_{ls1}\left[\mathbf{a}_{k1}\mathbf{a}'_{k2} + \boldsymbol{\Sigma}_{12}\right]\mathbf{d}_{ls2} + \mathbf{d}'_{lk1}\left[\mathbf{a}_{s1}\mathbf{a}'_{s2} + \boldsymbol{\Sigma}_{12}\right]\mathbf{d}_{kl2} + \mathbf{d}_{ks1})'\left[\mathbf{a}_{l1}\mathbf{a}'_{l2} + \boldsymbol{\Sigma}_{12}\right]\mathbf{a}_{ks2}$$

and $h_4(\cdot) = h(\cdot)$ with $\mathbf{A}_{ls12} = \mathbf{D}_{ls1}\mathbf{D}'_{ls2}$, $\mathbf{d}_{ls1} = \mathbf{a}_{l1} - \mathbf{a}_{s1}$ etc. We need $\xi_c = \mathrm{Var}[h_c(\cdot)]$ where ξ_1 follows from Theorem 7 which, along with Lemma 1, also gives ξ_2, with several covariances like $\mathrm{Cov}(\mathbf{a}_{l1}\boldsymbol{\Sigma}_{12}\mathbf{a}_{l2}, \mathbf{a}_{s1}\boldsymbol{\Sigma}_{12}\mathbf{a}_{s2})$ vanishing by independence. Part of ξ_3 follows exactly as ξ_2 and the rest, after tedious computations, using the moments below which themselves follow from Lemma 1

$$\mathrm{Var}(\mathbf{D}_{ls2}\boldsymbol{\Sigma}_{21}\mathbf{D}_{ls1}) = 4\{\omega_{1122} + \|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2\}$$
$$\mathrm{Cov}(\mathbf{a}'_{k1}\mathbf{A}_{ls12}\mathbf{a}_{k2}, \mathbf{D}_{lk2}\boldsymbol{\Sigma}_{21}\mathbf{D}_{lk1}) = 3\{\omega_{1122} + \|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2\}$$
$$\mathrm{Cov}(\mathbf{D}_{ls2}\boldsymbol{\Sigma}_{21}\mathbf{D}_{ls1}, \mathbf{D}_{lk2}\boldsymbol{\Sigma}_{21}\mathbf{D}_{lk1}) = \omega_{1122} + \|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2$$
$$\mathrm{Cov}(\mathbf{a}'_{k1}\mathbf{A}_{ls12}\mathbf{a}_{k2}, \mathbf{a}'_{s1}\mathbf{A}_{lk12}\mathbf{a}_{s2}) = 7\omega_{1122} + 5\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + \left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2$$
$$\mathrm{Var}(\mathbf{a}'_{k1}\mathbf{A}_{ls12}\mathbf{a}_{k2}) = 4\left\{4\omega_{1122} + 2\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + \left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2\right\}$$

Consider for example, $\mathrm{Var}(\mathbf{a}'_{k1}\mathbf{A}_{ls12}\mathbf{a}_{k2})$. Using $\mathbf{A}_{ls12} = \mathbf{d}_{ls1}\mathbf{d}'_{ls2}$ we can write

$$\mathrm{Var}(\mathbf{a}'_{k1}\mathbf{a}_{l1}\mathbf{a}'_{l2}\mathbf{a}_{k2}) + \mathrm{Var}(\mathbf{a}'_{k1}\mathbf{a}_{s1}\mathbf{a}'_{s2}\mathbf{a}_{k2}) + \mathrm{Var}(\mathbf{a}'_{k1}\mathbf{a}_{l1}\mathbf{a}'_{s2}\mathbf{a}_{k2}) + \mathrm{Var}(\mathbf{a}'_{k1}\mathbf{a}_{s1}\mathbf{a}'_{l2}\mathbf{a}_{k2})$$
$$+ 2\mathrm{Cov}(\mathbf{a}'_{k1}\mathbf{a}_{l1}\mathbf{a}'_{l2}\mathbf{a}_{k2}, \mathbf{a}'_{k1}\mathbf{a}_{s1}\mathbf{a}'_{s2}\mathbf{a}_{k2}) + 2\mathrm{Cov}(\mathbf{a}'_{k1}\mathbf{a}_{l1}\mathbf{a}'_{s2}\mathbf{a}_{k2}, \mathbf{a}'_{k1}\mathbf{a}_{s1}\mathbf{a}'_{l2}\mathbf{a}_{k2})$$

with other covariances zero. First two variances give the same result, using Lemma 1. Next two reduce, by conditioning on $\mathbf{a}_k$, to $\mathrm{Var}(\mathbf{a}'_{l1}\mathbf{M}\mathbf{a}_{s2})$, $\mathbf{M} = \mathbf{a}_{k1}\mathbf{a}'_{k2}$, and also give same result. The covariances, using independence and trace properties, reduce to $\mathrm{Var}(\mathbf{a}'_{k1}\boldsymbol{\Sigma}_{12}\mathbf{a}_{k2})$ which also follow from Theorem 7. Now $\xi_4 = 3\mathrm{Var}(A_{lskr1}A_{krls2}) + 6\mathrm{Cov}(A_{lskr1}A_{krls2}, A_{lksr1}A_{srlk2})$ for which

$$\mathrm{Var}(A_{lskr1}A_{krls2}) = 16\left\{2\omega_{1122} + 4\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + \left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2\right\}$$
$$\mathrm{Cov}(\mathbf{a}'_{l1}\mathbf{A}_{kr12}\mathbf{a}_{l2}, \mathbf{a}'_{k1}\mathbf{A}_{sr12}\mathbf{a}_{k2}) = 3\{3\omega_{1122} + \|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2\} = \mathrm{Cov}(\mathbf{a}'_{s1}\mathbf{A}_{kr12}\mathbf{a}_{s2}, \mathbf{a}'_{l1}\mathbf{A}_{sr12}\mathbf{a}_{l2})$$

where $\mathrm{Cov}(\mathbf{a}'_{l1}\mathbf{A}_{kr12}\mathbf{a}_{l2}, \mathbf{a}'_{l1}\mathbf{A}_{sr12}\mathbf{a}_{l2})$ and $\mathrm{Cov}(\mathbf{a}'_{s1}\mathbf{A}_{kr12}\mathbf{a}_{s2}, \mathbf{a}'_{k1}\mathbf{A}_{sr12}\mathbf{a}_{k2})$ are also same as given above. With $A_{lskr1}A_{krls2} = \mathbf{D}_{ls1}\mathbf{D}_{kr1}\mathbf{D}'_{kr2}\mathbf{D}_{ls2}$, the variance part follows from Lemma 1, and with $\mathbf{A}_{kr12} = \mathbf{D}_{kr1}\mathbf{D}'_{kr2}$, the covariance part follows by conditioning as argued above. The other three covariances follow the same way. We thus have

$$\xi_1 = 36\{\omega_{1122} + \|\boldsymbol{\Sigma}_{12}\|^2\}$$
$$\xi_2 = 4\left\{22\omega_{1122} + 20\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + \left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + \|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2\right\}$$
$$\xi_3 = 6\left\{28\omega_{1122} + 22\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + 3\left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + 3\|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2\right\}$$
$$\xi_4 = 12\left\{26\omega_{1122} + 16\|\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\|^2 + 5\left[\|\boldsymbol{\Sigma}_{12}\|^2\right]^2 + 5\|\boldsymbol{\Sigma}_{11}\|^2\|\boldsymbol{\Sigma}_{22}\|^2\right\}.$$

Substituting in $\mathrm{Var}(U_n)$ in Section A.2 gives $\mathrm{Var}(\mathrm{T}_2)$. For $\mathrm{Var}(\mathrm{T}_b)$, $\mathrm{Var}(\mathrm{T}_2)$ gives $\mathrm{Var}(\mathrm{T}_{ij})$. For covariances we focus on $C_1 = \mathrm{Cov}(\mathrm{T}_{ij}, T_{ij'})$ where C_2, C_3 follow similarly. From Section A.2, $h_i(\cdot)$ are same as $h(\cdot)$ above, so that $h_{ic} = \mathrm{E}[h_i(\cdot)|\mathbf{a}_k, \mathbf{a}_r, ...]$ with $\xi_{cc} = \mathrm{Cov}[h_{1c}(\cdot), h_{2c}(\cdot)]$. This, after long and tedious computations, gives

$$
\begin{aligned}
\xi_{11} &= 36\mathrm{Cov}(\mathbf{a}_{li}'\boldsymbol{\Sigma}_{ij}\mathbf{a}_{lj} + \|\boldsymbol{\Sigma}_{ij}\|^2, \mathbf{a}_{li}'\boldsymbol{\Sigma}_{ij'}\mathbf{a}_{lj'} + \|\boldsymbol{\Sigma}_{ij}\|^2) \;\; = \;\; 36\omega_{ijij'} + \omega_{ijj'i} \\
\xi_{22} &= 4\mathrm{Cov}(\mathbf{D}_{lsi}\boldsymbol{\Sigma}_{ij}\mathbf{D}_{lsj} + \mathbf{a}_{li}'\mathbf{a}_{si}\mathbf{a}_{sj}\mathbf{a}_{lj} + \mathbf{a}_{li}'\boldsymbol{\Sigma}_{ij}\mathbf{a}_{lj} + \mathbf{a}_{si}'\boldsymbol{\Sigma}_{ij}\mathbf{a}_{sj} + \|\boldsymbol{\Sigma}_{ij}\|^2, \mathbf{D}_{lsi}\boldsymbol{\Sigma}_{ij'}\mathbf{D}_{lsj'} \\
&\quad + \mathbf{a}_{li}'\mathbf{a}_{si}\mathbf{a}_{sj'}\mathbf{a}_{lj'} + \mathbf{a}_{li}'\boldsymbol{\Sigma}_{ij'}\mathbf{a}_{lj'} + \mathbf{a}_{si}'\boldsymbol{\Sigma}_{ij'}\mathbf{a}_{sj'} + \|\boldsymbol{\Sigma}_{ij'}\|^2) = 56\{\omega_{ijij'} + \omega_{iij'j}\} \\
\xi_{33} &= \mathrm{Cov}(\mathbf{D}_{lsi}'(\mathbf{a}_{ki}\mathbf{a}_{kj}' + \boldsymbol{\Sigma}_{ij})\mathbf{D}_{lsj} + \mathbf{D}_{lki}'(\mathbf{a}_{si}\mathbf{a}_{sj}' + \boldsymbol{\Sigma}_{ij})\mathbf{D}_{lkj} + \mathbf{D}_{ksi}'(\mathbf{a}_{si}\mathbf{a}_{sj}' + \boldsymbol{\Sigma}_{ij})\mathbf{D}_{ksj}, \\
&\quad \mathbf{D}_{lsi}'(\mathbf{a}_{ki}\mathbf{a}_{kj'}' + \boldsymbol{\Sigma}_{ij'})\mathbf{D}_{lsj'} + \mathbf{D}_{lki}'(\mathbf{a}_{si}\mathbf{a}_{sj'}' + \boldsymbol{\Sigma}_{ij'})\mathbf{D}_{lkj'} + \mathbf{D}_{ksi}'(\mathbf{a}_{si}\mathbf{a}_{sj'}' + \boldsymbol{\Sigma}_{ij'})\mathbf{D}_{ksj'}) \\
&= 6\Big\{22\omega_{ijij'} + 31\omega_{ijj'i} + 2\|\boldsymbol{\Sigma}_{ij}\|^2\|\boldsymbol{\Sigma}_{ij'}\|^2 + 5\|\boldsymbol{\Sigma}_{jj'}\|^2\|\boldsymbol{\Sigma}_{ii}\|^2\Big\} \\
\xi_{44} &= 48\Big\{2\{\omega_{ijij'} + \omega_{iij'j} + \omega_{iijj'}\} + \|\boldsymbol{\Sigma}_{ij}\|^2\|\boldsymbol{\Sigma}_{ij'}\|^2 + \|\boldsymbol{\Sigma}_{jj'}\|^2\|\boldsymbol{\Sigma}_{ii}\|^2\Big\}.
\end{aligned}
$$

Substituting and simplifying gives the required covariance.