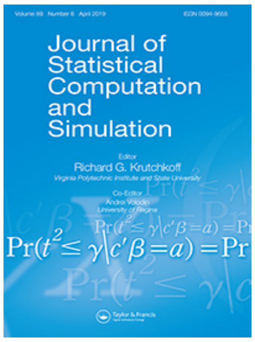

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: https://www.tandfonline.com/loi/gscs20

Multiple comparisons of mean vectors with large dimension under general conditions

M. Rauf Ahmad

To cite this article: M. Rauf Ahmad (2019) Multiple comparisons of mean vectors with large dimension under general conditions, Journal of Statistical Computation and Simulation, 89:6, 1044-1059, DOI: 10.1080/00949655.2019.1572147

To link to this article: https://doi.org/10.1080/00949655.2019.1572147

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION
2019, VOL. 89, NO. 6, 1044–1059
https://doi.org/10.1080/00949655.2019.1572147

Multiple comparisons of mean vectors with large dimension under general conditions

M. Rauf Ahmad

Department of Statistics, Uppsala University, Uppsala, Sweden

ABSTRACT
Multiple comparisons for two or more mean vectors are considered when the dimension of the vectors may exceed the sample size, the design may be unbalanced, populations need not be normal, and the true covariance matrices may be unequal. Pairwise comparisons, including comparisons with a control, and their linear combinations are considered. Under fairly general conditions, the asymptotic multivariate distribution of the vector of test statistics is derived whose quantiles can be used in multiple testing. Simulations are used to show the accuracy of the tests. Real data applications are also demonstrated.

ARTICLE HISTORY
Received 25 June 2018
Accepted 16 January 2019

KEYWORDS
Simultaneous inference; high-dimensional testing; pairwise comparisons

1. Introduction

The objective of this work is to present multiple comparisons for mean vectors in a multi-sample problem where the populations need not necessarily be normal, sample sizes and covariance matrices may be unequal, and the dimension of the vectors may exceed the sample sizes. Precisely, let $\mathbf{X}_{ik} = (X_{ik1}, \dots, X_{ikp})' \sim \mathcal{F}_i$, $k = 1, \dots, n_i$, be iid random vectors with $\mathrm{E}(\mathbf{X}_{ik}) = \boldsymbol{\mu}_i \in \mathbb{R}^p$, $\mathrm{Cov}(\mathbf{X}_{ik}) = \boldsymbol{\Sigma}_i \in \mathbb{R}^{p\times p}_{>0}$, $i = 1, \dots, g \geq 2$, where $\mathbb{R}^{p\times p}_{>0}$ denotes the space of real, symmetric, positive-definite, $p \times p$ matrices and $\mathcal{F}_i$ denotes the distribution function for ith population.

We are interested to develop multiple comparison procedures (MCP) or, correspondingly, simultaneous confidence intervals (SCI), for difference of mean vectors, by relaxing the usual linear model assumptions, e.g. normality and homoscedasticity. Thus, $\mathcal{F}_i$ may be non-normal and $\boldsymbol{\Sigma}_i$ may be unequal which, along with n_i also allowed to be unequal (unbalanced design), implies a complete multi-sample Behrens-Fisher problem. Further, we allow p to be large, even $p \gg n_i$. These comparisons are of interest as a first *post hoc* investigation after a global MANOVA hypothesis of equality of all mean vectors is rejected; see Seber [1] or Johnson and Wichern [2].

The multivariate theory offers a number of solutions to this problem for the classical case, $p < n_i$, particularly assuming normality and homoscedasticity. The global MANOVA hypotheses are mostly tested by the likelihood-ratio criterion such as Wikls' Λ and its

CONTACT M. Rauf Ahmad rauf.ahmad@statistik.uu.se Department of Statistics, Uppsala University, Uppsala, Sweden

rejection follows by finding out the mean vectors responsible for the global rejection. It commonly begins with a general strategy for a set of comparisons defined as linear combination, $\mathbf{a}'\boldsymbol{\delta}_{ij}$, $\mathbf{a} \in \mathbb{R}^p$, where $\boldsymbol{\delta}_{ij} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_j$, $i \neq j$. A case of particular interest is of pairwise differences $\boldsymbol{\delta}_{ij}$ themselves which includes all possible differences as well as special cases such as comparisons with a control.

The classical case of such comparisons has been extensively investigated; see e.g. Krishnaiah [3,4], Wijsman [5], Kropf [6], Kropf and Läuter [7], Westfall et al. [8], Läuter et al. [9], Conneely and Boehnke [10], Westfall and Troendle [11], Bretz et al. [12], Dickhaus [13], Goeman and Finos [14], Goeman and Solari [15], Guilbaud [16,17], where Dickhaus [18] is a modern, comprehensive book length reference with exhaustive bibliography.

The classical methods for MCP or SCI do not work when $p \gg n_i$ and need to be modified. The recent wave of high-dimensional data has motivated a thorough inquiry into new avenues for simultaneous inference which, already complicated enough as compared to global testing, is further exacerbated by the largeness of dimensionality. Of particular concern are the fields like genetics, microarray, agriculture, fMRI, psychology where analysing umpteen amounts of data has become a norm rather than exception; see e.g. Nichols and Hayasaka [19] and Dickhaus [18].

The multiple comparisons introduced in this paper are applicable for such high-dimensional data which, additionally, do not depend on usual assumptions such as normality and homoscedasticity. In fact, concerning normality, the tests can be used for any distribution with finite fourth moment across p-dimensional vector. A distinguishing feature of the proposed tests is that we exclusively derive asymptotic joint distribution of the entire vector of preliminary tests whose quantiles can be directly used to test any number of comparisons of g man vectors. Under a few, mild assumptions, the asymptotic covariance matrix turns out be of very simple form and particularly sparse, not only making the derivation of the limit distribution convenient but also enhancing the applicability of the proposed tests under fairly general conditions.

We begin in the next section with a concise notational set up, to be used throughout the paper, followed by the main tests and their properties. A simulation based evaluation is given in Section 3 and applications are given in Section 4. Section 5 summarizes the main points.

2. Test statistics and their properties

2.1. Notations and preliminary set up

Let the vectors $\mathbf{X}_{ik} \in \mathbb{R}^p$, $k \in \{1, \ldots, n_i\}$, $i \in \{1, \ldots, g\}$, as defined above, be generated by a probability space $(\mathcal{X}, \mathcal{A}, \mathcal{P}_\theta)$ where the probability measure $\mathcal{P}_\theta$ is indexed with parameter $\theta \in \Theta$ and Θ is the parameter space, not necessarily finite. Then $\mathbf{X}_i = (\mathbf{X}_1' \cdots \mathbf{X}_{n_i}') \in \mathbb{R}^{n_i \times p}$ is the data matrix for ith sample and $\mathbf{X} = (\mathbf{X}_1', \ldots, \mathbf{X}_{n_g}')' \in \mathbb{R}^{n \times p}$, $n = \sum_{i=1}^g n_i$, with parameter space $\{\boldsymbol{\Gamma}, \boldsymbol{\Xi}\}$, where $\boldsymbol{\Gamma} = \mathrm{E}(\mathbf{X}) = (\mathbf{1}_{n_1}' \otimes \boldsymbol{\mu}_1 | \cdots | \mathbf{1}_{n_g}' \otimes \boldsymbol{\mu}_g)'$, $\boldsymbol{\Xi} = \mathrm{Cov}(\mathbf{X}) = \oplus_{i=1}^g (\mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_i)$ with $\mathrm{Cov}(\mathbf{X}_i) = \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_i$ using $\mathrm{Cov}(\mathbf{X}_{ik}) = \boldsymbol{\Sigma}_i$ $\forall$ i, where $\oplus$ and $\otimes$ are the Kronecker sum and Kronecker product, respectively. Let $\bar{\mathbf{X}}_i = \sum_{k=1}^n \mathbf{X}_{ik}/n_i$ and $\hat{\boldsymbol{\Sigma}}_i = \sum_{k=1}^n \tilde{\mathbf{X}}_{ik}\tilde{\mathbf{X}}_{jk}'/(n_i - 1)$ be the usual unbiased estimators of $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ with $\tilde{\mathbf{X}}_{ik} = \mathbf{X}_{ik} - \bar{\mathbf{X}}_i$, or, using the ith data matrix, $i = 1, \ldots, g$,

$$\bar{\mathbf{X}}_i = \frac{1}{n_i}\mathbf{X}_i'\mathbf{1}_{n_i}, \quad \hat{\mathbf{\Sigma}}_i = \frac{1}{n_i - 1}\mathbf{X}_i'\mathbf{C}_{n_i}\mathbf{X}_i, \tag{1}$$

where $\mathbf{C}_{n_i} = \mathbf{I}_{n_i} - \mathbf{J}_{n_i}/n_i$ is centering matrix, $\mathbf{I}$ is identity matrix, $\mathbf{J} = \mathbf{1}\mathbf{1}'$ and $\mathbf{1}$ a vector of 1 s.

Let $\mathcal{H} = \{H_I : I \in \mathcal{I}\}$ be a family of hypotheses, finite or infinite, with card$\{\mathcal{I}\} = G$, corresponding to families of distributions $\{\mathcal{P}_\theta : \theta \in \Theta_I\}$ with parameter space Θ_I bifurcated into $\Theta_{0,I}$ and $\Theta_{1,I} = \Theta_I \backslash \Theta_{0,I}$, according to H_I being null ($(H_{0,I})$ or alternative $(H_{1,I})$ hypothesis, where $\Theta_0 \cup \Theta_1 = \Theta$, $\Theta_0 \cap \Theta_1 = \emptyset$. A (non-randomized) test for each H_I is carried out using a test statistic T_I with its space $\mathcal{T}_I$, which similarly bifurcates the sample space into $\mathcal{X}_{0,I}$ and $\mathcal{X}_{1,I}$, with a binary decision ϕ: $\mathcal{T}_I \to \{0, 1\}$ where $\phi = 1$ (0) when H_I is rejected (accepted).

As usual, the power function $\beta(\theta_I|\Theta_I) = \alpha$ (size) if $\Theta_I = \Theta_{0,I}$ and $1 - \beta$ (power) if $\Theta_I = \Theta_{1,I}$. For a sample $\mathbf{X} \in \mathcal{X}$, $p_I = \sup_{\theta \in \Theta_{0,I}} P(T_o \geq c_\alpha)$ is the p-value of T_I with observed value T_o and critical value c_α. The problem of MCP pertains to simultaneously testing a set of G hypotheses

$$H_{0,I} : \theta \in \Theta_{0,I} \text{ vs. } H_{1,I} : \theta \in \Theta_{1,I}, \quad I \in \mathcal{I}, \quad \text{card}\{\mathcal{I}\} = G.$$

For pairwise comparisons of $\boldsymbol{\mu}_i$, we have $\theta = \boldsymbol{\delta}_{ij} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_j$, $i \neq j$, with $G = \binom{g}{2} = g(g-1)/2$, and for comparisons with a control, $\theta = \boldsymbol{\delta}_{1j} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_j$ with $G = g-1$, $j = 2, \ldots, g$, assuming, without loss of generality, sample 1 as control. In either case, we essentially deal with a vector of test statistics $\mathbf{T} \in \mathbb{R}^G$ and corresponding vector of observed p-values, $\mathbf{p} \in (0, 1)^{\otimes G}$.

With several tests being carried out simultaneously, the most serious issue in multiple testing is to effectively control α, i.e. reduce the chance of false positives (FP). Let $I_0 \subset \mathcal{I}$ be the subset corresponding to the true null hypotheses, $\mathcal{H}_0 = \{H_{0,I} : I \in \mathcal{I}_0\}$, with card$\{\mathcal{I}_0\} = G_0 \leq G$, and $R \subset \mathcal{I}$ be the subset for which $H_{0,I}$ is rejected. Then $f_m =$ card$\{R \cap \mathcal{I}_0\}$ refers to the set of FPs (rejected true hypotheses or type I errors), so that $r_m =$ card$\{R \backslash \mathcal{I}_0\}$ is the index of true positives or TPs (rightly rejected null hypotheses or power of test). We, therefore, are interested to keep f_m (r_m) as small (large) as possible. Several error control procedures can be adopted, subject to research questions. For details, see e.g. Hochberg and Tamhane [20], Bretz et al. [12], Dickhaus [18], Goeman and Solari [15], Hemerik and Goeman [21].

In practice, family-wise error control (in the strong sense), FWEs, is the most desired error control and will be our main target in the sequel. It is the proportion of all FPs, i.e. $P(f_m > 0)$. The simplest way to control FWEs is through Bonferroni inequality which ensures $P(f_m > 0) \leq G_0\alpha/G \leq \alpha$, where equality holds in most cases since $G_0 = G$, i.e. each of G tests has α/G chance for FP. It offers an efficient control for small to moderate G but is obviously conservative (or has less power) as G becomes large. An alternative option is the false discovery rate, FDR $= \mathrm{E}[\{f_m/(f_m + r_m)\}\mathbf{1}_{\{f_m + r_m \geq 1\}}]$ with $\mathbf{1}_{\{\cdot\}}$ as indicator function; see e.g. Dickhaus [18, Ch. 1].

Among other notations used in the sequel, a vector $\mathbf{a} \in \mathbb{R}^p$ is a column vector with norm $\|\mathbf{a}\|^2 = \langle \mathbf{a}, \mathbf{a} \rangle$ and a matrix norm is Frobenius $\|\mathbf{A}\|^2 = \mathrm{tr}(\mathbf{A}^2)$. The test statistics are formulated as linear combinations of second-order U-statistics of symmetric (product) kernels, $h(\cdot) : \mathbb{R}^p \mapsto \mathbb{R}$, defined as bilinear forms of independent vectors. With $h(\cdot)$ a measurable, possibly degenerate, square-integrable, $\int h^2 dP < \infty$, function, the set up conforms to a Hilbert space $\mathbb{L}_2(\mathbb{H})$ equipped with inner product $\langle \cdot, \cdot \rangle : \mathbb{R}^p \to \mathbb{R}$, so that $h(\cdot)$, with an orthonormal decomposition, is a Hilbert-Schmidt kernel; see van der Vaart [22] or Lee [23]. This helps us study the properties of test statistics under flexible conditions, the subject of next section.

2.2. Test statistics and their properties

For the data set up in Section 2.1, let $\mathrm{T}_I = \mathrm{T}_{ij}$ be the test statistic for a (preliminary) hypothesis $H_{0,I} = H_{0ij} : \boldsymbol{\delta}_{ij} = \mathbf{0}$ with $\boldsymbol{\delta}_G \in \mathbb{R}^G$ the vector of all hypotheses to be simultaneously tested. Thus, for all pairwise differences, $\boldsymbol{\delta}_{ij} : \boldsymbol{\mu}_i - \boldsymbol{\mu}_j$, $i < j$, with $G = g(g-1)/2$, $\boldsymbol{\delta}_G = (\boldsymbol{\delta}_{11}, \ldots, \boldsymbol{\delta}_{g-1,g})'$ where

$$\mathbf{T}_G = (\mathbf{T}_1, \ldots, \mathbf{T}_{g-1})' = (\mathrm{T}_{12}, \ldots, \mathrm{T}_{1g}, \mathrm{T}_{23}, \ldots, \mathrm{T}_{2g}, \ldots, \mathrm{T}_{g-2,g}, \mathrm{T}_{g-1,g})', \tag{2}$$

is the vector of test statistics, a set of simultaneous tests for $H_0 : \boldsymbol{\delta}_G = \mathbf{0}$, with $\mathbf{T}_i = (\mathrm{T}_{i,i+1}, \ldots, \mathrm{T}_{ig})'$, $i = 1, \ldots, g-1$. Our strategy begins by defining T_{ij}, a test statistic for H_{0ij}, valid for $p \gg n_i$ where $\mathcal{F}_i$ may be non-normal and $\boldsymbol{\Sigma}_i$ may be unequal. The limit of T_{ij} is derived under flexible conditions since the multiple tests heavily rest on the properties of T_{ij}. Using these properties, we derive the joint distribution of $\mathbf{T}_G$ to be used for MCP for any G. The most salient feature is that the effect of high-dimensionality, $p \to \infty$, is taken care of in T_{ij}, so that the limit of $\mathbf{T}_G$ is mainly influenced by g or G. Now, to define T_{ij}, consider $Q_{ij0} = U_i + U_j - 2U_{ij}$ where

$$U_i = \frac{1}{n_i(n_i-1)} \sum_{\substack{k=1 \\ k \neq r}}^{n_i} \sum_{r=1}^{n_i} h(\mathbf{X}_{ik}, \mathbf{X}_{ir}), \quad U_{ij} = \frac{1}{n_i n_j} \sum_{k=1}^{n_i} \sum_{l=1}^{n_j} h(\mathbf{X}_{ik}, \mathbf{X}_{jl}), \tag{3}$$

are one- and two-sample U-statistics, respectively, with symmetric kernels $h(\mathbf{X}_{ik}, \mathbf{X}_{ir}) = \mathbf{X}'_{ik}\mathbf{X}_{ir}/p$, $h(\mathbf{X}_{ik}, \mathbf{X}_{jl}) = \mathbf{X}'_{ik}\mathbf{X}_{jl}/p$, $k, r = 1, \ldots, n_i$, $k \neq r$, $l = 1, \ldots, n_j$, $i, j = 1, \ldots, g$, $i \neq j$, $n_{ij} = n_i + n_j$. Now $\mathrm{E}(Q_{ij0}) = \|\boldsymbol{\delta}_{ij}\|^2 = 0$ under H_{0ij}, $\boldsymbol{\delta}_{ij} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_j$, so that Q_{ij0} can be used to test H_{0ij}. For scaling and appropriate limit, also consider $Q_{ij1} = Q_{i1} + Q_{j1}$, $Q_{i1} = (E_i - U_i)/n_i$, $E_i = \sum_{k=1}^{n_i} \mathbf{X}'_{ik}\mathbf{X}_{ik}/n_i$. Note that, $Q_{i1} = \mathrm{tr}(\hat{\boldsymbol{\Sigma}}_i)/n_i \Rightarrow Q_{ij1} = \mathrm{tr}(\hat{\boldsymbol{\Sigma}}_{ij0})$, $\hat{\boldsymbol{\Sigma}}_{ij0} = \hat{\boldsymbol{\Sigma}}_i/n_i + \hat{\boldsymbol{\Sigma}}_j/n_j$ so that $\mathrm{E}(Q_{ij1}) = \mathrm{tr}(\boldsymbol{\Sigma}_{ij0})$, which is same under H_{0ij} and H_{1ij}, where $\boldsymbol{\Sigma}_{ij0} = \boldsymbol{\Sigma}_i/n_i + \boldsymbol{\Sigma}/n_j$. Thus, writing $Q_{ij} = Q_{ij1} + Q_{ij0}$, it follows that [see also 24]

$$\mathrm{E}(Q_{ij}) = \|\boldsymbol{\delta}_{ij}\|^2 + \mathrm{tr}(\boldsymbol{\Sigma}_{ij0}) = \mathrm{tr}(\boldsymbol{\Sigma}_{ij0}) \mathrm{under} H_{0ij}.$$

We thus define the two-sample test statistic for H_{0ij} as

$$\mathrm{T}_{ij} = 1 + \frac{n_{ij} Q_{ij0}}{[n_{ij} Q_{ij1}/p]}. \tag{4}$$

T_{ij} is location-invariant so that we can assume $\boldsymbol{\mu}_i = \mathbf{0}$ $\forall i$ without loss of generality. T_{ij} is defined in Ahmad [25] as a modification of the Hotelling's two-sample T^2 statistic to test H_{0ij} for high-dimensional data under non-normality and heteroscedasticity. Recall $\mathrm{T}^2 = (n_i n_j / n_{ij}) \hat{\boldsymbol{\delta}}'_{ij} \hat{\boldsymbol{\Sigma}}_{ij}^{-1} \hat{\boldsymbol{\delta}}_{ij}$ where $\hat{\boldsymbol{\delta}}_{ij} = \bar{\mathbf{X}}_i - \bar{\mathbf{X}}_j$ and $\hat{\boldsymbol{\Sigma}} = [(n_i - 1)\hat{\boldsymbol{\Sigma}}_i + (n_j - 1)\hat{\boldsymbol{\Sigma}}_j]/(n_i +$

$n_j - 2)$ is pooled estimator of $\mathbf{\Sigma}_i = \mathbf{\Sigma}_j = \mathbf{\Sigma}$ [1, see e.g.]. The modification pertains to removing $\hat{\mathbf{\Sigma}}^{-1}$, which does not exist when $p > n_i$, and writing $\|\hat{\boldsymbol{\delta}}_{ij}\|^2 = Q_{ij1} + Q_{ij0} = Q_{ij}$ since $\|\bar{\mathbf{X}}_i\|^2 = \sum_{k,r=1}^{n_i} \mathbf{X}'_{ik}\mathbf{X}_{ir}/n_i^2 = (E_i - U_i)/n_i + U_i$. Properties of T_{ij} are studied under the following assumptions.

Assumption 2.1: $\mathrm{E}(X_{iks}^4) \le \gamma < \infty, i = 1, \dots, g, \forall\, s = 1, \dots, p, \gamma \in \mathbb{R}^+$.

Assumption 2.2: As $n_i \to \infty, n_i/n \to \rho_i \in (0, \infty), i = 1, \dots, g$.

Assumption 2.3: As $p \to \infty, \mathrm{tr}(\mathbf{\Sigma}_i)/p = \kappa_i = O(1), i = 1, \dots, g$.

Assumption 2.4: As $p \to \infty, \boldsymbol{\mu}_i'\mathbf{\Sigma}_k\boldsymbol{\mu}_j/p^2 = \psi_{ij}, 0 < \psi_{ij} < \infty, i = 1, \dots, g, k = i$ or $k = j$.

The assumptions are stated for g samples for their further use in the sequel. Note that, by Assumption 2.3, $\|\mathbf{\Sigma}_i\|^2/p^2 = O(1)$. If we let $\lambda_i \in \mathbb{R}^+$ be the eigenvalues of $\mathbf{\Sigma}_i$, so that ν_i be those of $\mathbf{\Sigma}_i/p$, $i \in \{1, \dots, g\}$, then Assumption 2.3 and its consequence uniformly bound the first two moments of ν_i. Assumption 2.1 is inevitably needed to compute moments of bilinear forms when normality is relaxed. Assumption 2.4 is only needed for distribution under the alternative.

Assumptions 2.2 and 2.3 are mild and frequently used in high-dimensional testing problems. In particular, Assumption 2.3 holds for many commonly used covariance structures. Consider, e.g. $\mathbf{\Sigma}$ as compound symmetric (CS), $\mathbf{\Sigma} = (1 - \rho)\mathbf{I} + \rho\mathbf{J}$ with $\mathbf{I}$ as identity matrix, $\mathbf{J} = \mathbf{1}\mathbf{1}'$, $\mathbf{1}$ a vector of 1s, $-1/(p-1) \le \rho \le 1$. Then $\mathrm{tr}(\mathbf{\Sigma}^r) = O(p^r)$, $r = 1, 2$. Note that, unlike common practice in the literature, we need not assume similar bound for higher moments of the eigenvalues of $\mathbf{\Sigma}$, e.g. $\mathrm{tr}(\mathbf{\Sigma}^2)/p = O(1)$ which may collapse for many useful structures, including CS. Note also that CS belongs to spiked structures where a few eigenvalues dominate the rest, so that the proposed procedures hold for such structures as well. See also discussion after Assumption 2.6 below.

Under these assumptions, the limit of T_{ij}, for $n,_i, p \to \infty$, is given in Ahmad [25]. First, $n_{ij}Q_{ij1}/p \xrightarrow{\mathcal{P}} \rho_i^{-1}\kappa_i + \rho_j^{-1}\kappa_j = \sum_{s=1}^{\infty}(\rho_i^{-1}\nu_{si} + \rho_j^{-1}\nu_{sj}) = K_{ij}$, as $n_i, p \to \infty$. The limit obviously approximates $\mathrm{E}(Q_{ij1}) = \mathrm{tr}(\mathbf{\Sigma}_{ij0})$ and holds both under H_{0ij} and H_{1ij}. As $\mathrm{E}(Q_{ij0}) = \|\boldsymbol{\delta}_{ij}\|^2 = 0$ under H_{0ij}, the kernels of U_i and U_{ij} are degenerate, so that [22] $n_iU_i \xrightarrow{\mathcal{D}} \sum_{s=1}^{\infty} \nu_{is}(z_{is}^2 - 1)$, $\sqrt{n_in_j}U_{ij} \xrightarrow{\mathcal{D}} \sum_{s=1}^{\infty} \nu_{is}z_{is}z_{js}$, where $z_{is} \sim N(0, 1)$, iid. Then $n_{ij}Q_{ij0} \xrightarrow{\mathcal{D}} \sum_{s=1}^{\infty}(\rho_i^{-1}\nu_{is}z_{is}^2 + \rho_j^{-1}\nu_{is}z_{jm}^2 - 2\rho_i^{-1/2}\rho_j^{-1/2}\nu_{ijs}z_{is}z_{js}) - K_{ij}$ and, by Slutsky's lemma,

$$\mathrm{T}_{ij} \xrightarrow{\mathcal{D}} \frac{1}{K_{ij}} \sum_{m=1}^{\infty} (\rho_i^{-1/2}\nu_{im}^{1/2} z_{im} - \rho_j^{-1/2}\nu_{jm}^{1/2} z_{jm})^2, \tag{5}$$

where the limiting moments, $\mathrm{E}(\mathrm{T}_{ij}) \approx 1$, $\mathrm{Var}(\mathrm{T}_{ij}) \approx 2\sum_{m=1}^{\infty}(\rho_1^{-1}\nu_{1m} + \rho_2^{-1}\nu_{2m})^2/K_{ij}^2$, approximate the first two moments of $\chi^2_{f_{ij}}/f_{ij}$, $f_{ij} = [\mathrm{tr}(\mathbf{\Omega}_{0ij})]^2/\mathrm{tr}(\mathbf{\Omega}_{0ij}^2)$, $\mathbf{\Omega}_{0ij} = n\mathbf{\Sigma}_{ij0}/p$. Thus $\mathrm{T}_{ij} \xrightarrow{\mathcal{D}} \chi^2_{f_{ij}}/f_{ij}$. The normal limit follows by an application of Hájek-Šidák Lemma [26, p. 183]. The limit under H_{1ij} follows by the projection theory of U-statistics. We estimate $\mathrm{Var}(\mathrm{T}_{ij}) = \sigma^2_{\mathrm{T}_{ij}}$ by using unbiased, consistent estimators of traces in f_{ij}, i.e.

$\text{tr}(\boldsymbol{\Sigma}_i^2)$, $[\text{tr}(\boldsymbol{\Sigma}_i)]^2$, $\text{tr}(\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j)$, defined as $E_{2i} = \eta_i\{(n_i-1)(n_i-2)\text{tr}(\hat{\boldsymbol{\Sigma}}_i^2) + [\text{tr}(\hat{\boldsymbol{\Sigma}}_i)]^2 - n_iQ_i\}$, $E_{3i} = \eta_i\{2\text{tr}(\hat{\boldsymbol{\Sigma}}_i^2) + (n_i^2-3n_i+1)[\text{tr}(\hat{\boldsymbol{\Sigma}}_i)]^2 - n_iQ_i\}$ and $\text{tr}(\hat{\boldsymbol{\Sigma}}_1\hat{\boldsymbol{\Sigma}}_2)$, where $Q_i = \sum_{k=1}^{n_i}(\tilde{\mathbf{X}}_{ik}'\tilde{\mathbf{X}}_{ik})^2/(n_i-1)$, $\tilde{\mathbf{X}}_i = \mathbf{X}_{ik} - \bar{\mathbf{X}}_i$, $\eta_i = (n_i-1)/[n_i(n_i-2)(n_i-3)]$. The consistent estimator $\hat{\text{Var}}(\text{T}_{ij})$ can replace $\text{Var}(\text{T}_{ij})$ in T_{ij}. Following theorem summarizes the limit. For proof and an extension to multi-sample case, see Ahmad [25].

Theorem 2.5: *For* T_{ij} *in Equation* (4), $(\text{T}_{ij} - \text{E}(\text{T}_{ij}))/\sigma_{\text{T}_{ij}} \xrightarrow{\mathcal{D}} N(0,1)$, $n_i, n_j, p \to \infty$, *under Assumptions* 2.1–2.4. *The limit remains valid by replacing* $\sigma_{T_{ij}}^2$ *with its consistent estimator defined above.*

A few remarks concerning Theorem 2.5 will help us proceed further. First, the limit of T_{ij} holds for any distribution with finite fourth moment. Second, the composition of T_{ij} in terms of U-statistics helps us relax normality and obtain the limit conveniently as the kernels are simple bilinear forms of independent components. The accuracy of T_{ij} for small or moderate n_i and large p is shown through simulations in Ahmad [25]. This also implies that the dimension p is taken care of in the limit of T_{ij}, so that the extension to multiple comparisons will not be much influenced by p. Finally, as Q_{ij1} converges to $\text{E}(Q_{ij1}) = \text{tr}(\boldsymbol{\Sigma}_{ij0})$ in probability, the limit of T_{ij} mainly follows from Q_{ij0}. Thus, in extending the limit to $\mathbf{T}_G$, we mainly focus on Q_{ij0}. For this, note that

$$\text{Var}(Q_{ij0}) = 2\|\boldsymbol{\Sigma}_{ij0}\|^2 + 4\boldsymbol{\delta}_{ij}'\boldsymbol{\Sigma}_{ij0}\boldsymbol{\delta}_{ij} \tag{6}$$

$$\text{Cov}(Q_{ij0}, Q_{ij'0}) = \frac{2}{n_i^2}\|\boldsymbol{\Sigma}_i\|^2 + \frac{4}{n_i}\boldsymbol{\delta}_{ij}'\boldsymbol{\Sigma}_i\boldsymbol{\delta}_{ij'} \tag{7}$$

$$\text{Cov}(Q_{ij0}, Q_{i'j0}) = \frac{2}{n_j^2}\|\boldsymbol{\Sigma}_j\|^2 + \frac{4}{n_j}\boldsymbol{\delta}_{ij}'\boldsymbol{\Sigma}_j\boldsymbol{\delta}_{i'j} \tag{8}$$

with $\text{Cov}(Q_{ij0}, Q_{i'j'0}) = 0$ for $i \neq i', j \neq j'$ (see Appendix) where, under H_{0ij},

$$\text{Var}(Q_{ij0}) = 2\|\boldsymbol{\Sigma}_{ij0}\|^2, \quad \text{Cov}(Q_{ij0}, Q_{ij'0}) = \frac{2}{n_i^2}\|\boldsymbol{\Sigma}_i\|^2, \quad \text{Cov}(Q_{ij0}, Q_{i'j0}) = \frac{2}{n_j^2}\|\boldsymbol{\Sigma}_j\|^2, \tag{9}$$

independent of $\boldsymbol{\mu}_i$. Now, with $\mathbf{Q}_{i0} = (Q_{ij0}, \ldots, Q_{ig0})'$, $i = 1, \ldots, g-1$, consider the vector

$$\mathbf{Q}_0 = (\mathbf{Q}_{10}', \ldots, \mathbf{Q}_{g-1,0}')', \tag{10}$$

where $\text{E}(\mathbf{Q}_0) = \mathbf{0}$, $\text{Cov}(\mathbf{Q}_0) = \boldsymbol{\Lambda} = 2(\boldsymbol{\Lambda}_{ij}/p^2)_{i,j=1}^G \in \mathbb{R}^{G\times G}$, a partitioned matrix with diagonal and off-diagonal blocks $\text{Cov}(\mathbf{Q}_{i0}) = \boldsymbol{\Lambda}_{ii}/p^2 \in \mathbb{R}^{(g-i)\times(g-i)}$, $\text{Cov}(\mathbf{Q}_{i0}, \mathbf{Q}_{j0}) = \boldsymbol{\Lambda}_{ij}/p^2 \in \mathbb{R}^{(g-i)\times(g-j)}$, i.e.

$$\boldsymbol{\Lambda}_{ii} = \frac{1}{n_i^2}\|\boldsymbol{\Sigma}_i\|^2(\mathbf{J}-\mathbf{I})_{g-i} + \oplus_{j=i+1}^g\|\boldsymbol{\Sigma}_{ij0}\|^2, \quad \boldsymbol{\Lambda}_{ij} = \mathbf{0}' \; \frac{1}{n_i^2}\|\boldsymbol{\Sigma}_i\|^2\mathbf{1}_{g-i}' \; \frac{1}{n_j^2} \oplus_{j=i+2}^g \|\boldsymbol{\Sigma}_j\|^2 \tag{11}$$

$i = 1, \ldots, g-1$, $j = i+1, \ldots, g$, $\mathbf{1}$ is vector of 1s, $\mathbf{J} = \mathbf{11}'$, $\mathbf{I}$ is identity matrix, $\oplus$ is Kronecker sum and $\mathbf{0}$ in $\boldsymbol{\Lambda}_{ij}$ is of order $(j-i-1)\times(g-j)$ with no zero row if $j{-}i{-}1 = 0$.

A closer look at the structure of $\boldsymbol{\Lambda}$ reveals several aspects which will simplify the computations that follow. Ignoring p^2 for simplicity, and denoting $a_i = \|\boldsymbol{\Sigma}_i\|^2/n_i^2$, $a_{ij} = \|\boldsymbol{\Sigma}_{ij0}\|^2$, we can write

$$\boldsymbol{\Lambda}_{ii} = \begin{pmatrix} a_{i,i+1} & a_i & \dots & a_i \\ a_i & a_{i,i+2} & \dots & a_i \\ \vdots & \vdots & \ddots & \vdots \\ a_i & a_i & \dots & a_{i,g} \end{pmatrix}. \tag{12}$$

For any given i, $\boldsymbol{\Lambda}_{ii}$ has same off-diagonal element, a_i, with diagonal elements a_{ij}, where $\boldsymbol{\Sigma}_{ij0} = \boldsymbol{\Sigma}_i/n_i + \boldsymbol{\Sigma}_j/n_j = \text{Cov}(\hat{\boldsymbol{\delta}}_{ij}), j = i+1$. For off-diagonal blocks $\boldsymbol{\Lambda}_{ij}$,

$$\boldsymbol{\Lambda}_{12} = \begin{pmatrix} a_2 & a_2 & \dots & a_2 \\ a_3 & 0 & \dots & 0 \\ 0 & a_4 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_g \end{pmatrix}, \quad \boldsymbol{\Lambda}_{13} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ a_3 & a_3 & \dots & a_3 \\ a_4 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_g \end{pmatrix},$$

$$\boldsymbol{\Lambda}_{1,g-2} = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ a_{g-2} & a_{g-2} \\ a_{g-1} & 0 \\ 0 & a_g \end{pmatrix}, \quad \boldsymbol{\Lambda}_{1,g-1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{g-1} \\ a_g \end{pmatrix}$$

$$\boldsymbol{\Lambda}_{23} = \begin{pmatrix} a_3 & a_3 & \dots & a_3 \\ a_4 & 0 & \dots & 0 \\ 0 & a_5 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_g \end{pmatrix}, \quad \boldsymbol{\Lambda}_{2,g-2} = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ a_{g-2} & a_{g-2} \\ a_{g-1} & 0 \\ 0 & a_g \end{pmatrix},$$

$$\boldsymbol{\Lambda}_{2,g-1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{g-1} \\ a_g \end{pmatrix}, \quad \boldsymbol{\Lambda}_{g-2,g-1} = \begin{pmatrix} a_{g-1} & 0 \\ 0 & a_g \end{pmatrix}$$

The off-diagonal elements in $\boldsymbol{\Lambda}_{ij}$ are mostly 0 and the number of (rows with) zeros increases with increasing j for every i, making $\boldsymbol{\Lambda}$ an increasingly sparse matrix. However, the distinct non-zero elements in $\boldsymbol{\Lambda}$ consist of a much smaller set

$$\left\{\text{tr}(\boldsymbol{\Sigma}_i^2),\ \text{tr}(\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j),\ i,\ j = 1, \dots, g,\ i < j\right\}, \tag{13}$$

with cardinality $C_e = g(g+1)/2$. Thus, for any g, we only need to estimate C_e out of $C_T = G(G+1)/2$ elements in order to estimate $\boldsymbol{\Lambda}$. For example, for $g = 6, 9, 12, 15, 20$ samples, $C_e = 21, 66, 78, 120, 210$ whereas $C_T = 120, 1540, 2211, 5565, 18145$, respectively. The consistent estimators of these traces are given before Theorem 2.5. Used as plug-in estimators, they lead to a consistent estimator, $\hat{\boldsymbol{\Lambda}}$, of $\boldsymbol{\Lambda}$. A further simplification follows from weak (mostly zero) off-diagonal elements as compared to diagonal ones, so that the following assumption holds trivially.

Assumption 2.6: $\lim_{p\to\infty} \|\boldsymbol{\Sigma}_i\|^2/[\{\text{tr}(\boldsymbol{\Sigma}_i + \boldsymbol{\Sigma}_j)\}\{\text{tr}(\boldsymbol{\Sigma}_i + \boldsymbol{\Sigma}_k)\}] \to \gamma \in [0,1)$, $i \neq j \neq k = 1,\ldots,g$.

Although, Assumption 2.6 is kept flexible to adjust many covariance structures, it can be shown that the ratio indeed vanishes for most covariance structures, so that Assumption 2.6 encompasses many practical cases, including trivial ones e.g. $\boldsymbol{\Sigma} \propto \mathbf{I}$; see also Section 4. For the distribution of $\mathbf{T}_G$, consider the moments of Q_{ij0} in Equations (6)–(9). Using the projection theory of U-statistics (Appendix), the projection of Q_{ij0} can be shown as

$$\hat{Q}_{ij0} = 2\boldsymbol{\delta}_{ij}'\left\{(\bar{\mathbf{X}}_i - \boldsymbol{\mu}_i) - (\bar{\mathbf{X}}_j - \boldsymbol{\mu}_j)\right\}/p = 2\boldsymbol{\delta}_{ij}'\left\{(\bar{\mathbf{X}}_i - \bar{\mathbf{X}}_j) - \boldsymbol{\delta}_{ij}\right\}/p,$$

see [25, Appendix B.2]. As $\hat{Q}_{ij0}$ is composed of independent components and holds for any pair (i,j), the projection of $\mathbf{Q}_{i0}$, hence of $\mathbf{Q}_0$, consists of sums of these independent components. Further, with Q_{ij1} converging to a constant in probability, the limit for $\mathbf{T}_G$ follows conveniently by the Cramér-Wold device and Slutsky's lemma [22]. Finally, using the plug-in consistent estimators of the elements of $\boldsymbol{\Lambda}$, the limit also extends to $\hat{\boldsymbol{\Lambda}}$. We have the following theorem.

Theorem 2.7: *For $\mathbf{T}_G$, the limit in Equation (14) holds under Assumptions 2.1–2.6, as $n_i, p \to \infty$. Further, the limit remains valid by replacing $\boldsymbol{\Lambda}$ with its consistent estimator defined above.*

As mentioned above, the off-diagonal elements in $\boldsymbol{\Lambda}$ vanish under Assumption 2.6 for most covariance matrices, leaving $\boldsymbol{\Lambda}$ diagonal. This makes the limit in Theorem 2.7 much easier to prove and simpler to use. In particular, with $\hat{f}_{ij}$ as the estimator of f_{ij}, as discussed after Equation (5), we can use the Chi-square limit with $\text{Cov}(\mathbf{T}_G) \approx \text{diag}(2/f_{12},\ldots,2/f_{g-1,g})$ with f_{ij} estimated as $\hat{f}_{ij}$. Alternatively, the corresponding normal limit may be used. In fact, given the structure of the test statistics, and also because the normal limit follows through Chi-square limit, it has been observed that the Chi-square approximation mostly performs relatively better that the normal limit, and is thus strongly recommended for practical applications.

Note that, Theorem 2.7 implies that the limit also holds for any linear combination $\mathbf{a}'\mathbf{T}_G$, $\mathbf{a} \in \mathbb{R}^G\backslash\{\mathbf{0}\}$. With $\text{E}(\mathbf{T}_G) \approx \mathbf{1}_G$, we have, for $n_i, p \to \infty$,

$$\mathbf{a}'\mathbf{T}_G \xrightarrow{\mathcal{D}} N(\mathbf{a}'\mathbf{1}, \mathbf{a}'\boldsymbol{\Lambda}\mathbf{a}), \tag{14}$$

so that we can also test any linear combination $H_0 : \mathbf{a}'\boldsymbol{\delta}_G = 0$, particularly including any single $\boldsymbol{\delta}_{ij} = \mathbf{0}$, using $\sqrt{2/f_{ij}}(\text{T}_{ij} - 1) \xrightarrow{\mathcal{D}} N(0,1)$. The corresponding 100(1 - α)% simultaneous confidence interval (SCI) for $\mathbf{a}'\boldsymbol{\delta}_G$ follows as

$$\mathbf{a}'\hat{\mathbf{T}}_G \mp z_{\alpha/2}\sqrt{\mathbf{a}'\hat{\boldsymbol{\Lambda}}\mathbf{a}}, \tag{15}$$

where $z_{\alpha/2}$ is 100(α/2)% quantile of $N(0,1)$-distribution. Note that, the observed length of this confidence interval is $\hat{L} = 2z_{\alpha/2}(\mathbf{a}'\hat{\boldsymbol{\Lambda}}\mathbf{a})^{1/2}$. By the consistency of $\hat{\boldsymbol{\Lambda}}$ (Theorems

2.5-2.7) and the continuous mapping theorem, $\mathrm{E}(\hat{L})$ converges to $\mathbf{a}'\boldsymbol{\Lambda}\mathbf{a}$ which, under the assumptions, is a finite value, assuming $\|\mathbf{a}\|^2 < \infty$ which holds conveniently.

The comparison of treatments with a control is a special case of all pairwise comparisons presented above. Let Sample 1 be treated as control, and the interest is to test it against all other samples, i.e. $H_{i0} : \boldsymbol{\delta}_{1i} = \mathbf{0}$, $\boldsymbol{\delta}_{1i} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_i$, $i = 2, \ldots g$. The vector of tests is

$$\mathbf{T}_1 = (\mathrm{T}_{12}, \ldots, T_{1g})', \tag{16}$$

which is the first sub-vector of $\mathbf{T}_G$ in Equation (2). Using the related computations, we get $\mathrm{E}(\mathbf{Q}_{01}) = \mathbf{0}_{g-1}$, $\mathrm{Cov}(\mathbf{Q}_{01}) = \boldsymbol{\Lambda}_{11}$, the first diagonal block of $\boldsymbol{\Lambda}$, so that under the assumptions, $\mathrm{E}(\mathbf{T}_1) \approx \mathbf{1}_{g-1}$ and, assuming zero off-diagonals, $\mathrm{Cov}(\mathbf{T}_1) = \mathrm{diag}(2/f_{12}, \ldots, 2/f_{1g})$. The multiple tests and corresponding confidence intervals follows from those given for $\mathbf{T}_G$ above, without much changes.

3. Simulations

We do a simulation study to assess the performance of the proposed tests, in terms of their size control and power, and also their robustness to the violation of assumptions. We consider $g = 3$ and 6 samples and generate p-dimensional iid vectors from normal, uniform and exponential distributions. For $g = 3$, we use $(n_1, n_2, n_3) = (10, 15, 20), (20, 30, 40)$, (10, 30, 60) and (50, 75, 100), with $p \in \{50, 300, 500, 1000\}$, where the last sample size triplet corresponds to large samples and penultimate triplet amounts to very unbalanced design. The other two triplets are used to show the accuracy of the tests for small to moderate sample sizes. We use three covariance structures, Compound Symmetry (CS), Autoregressive of order 1, AR(1), and unstructured (UN), defined, respectively, as $\kappa\mathbf{I} + \rho\mathbf{J}$, $\mathrm{Cov}(X_i, X_j) = \kappa\rho^{|i-j|}$, $\forall\, k, l$ and $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1}^{d}$ with $\sigma_{ij} = 1(1)d\ (i=j)$, $\rho_{ij} = (i-1)/d\ (i>j)$, where $\mathbf{I}$ is identity matrix and $\mathbf{J} = \mathbf{1}\mathbf{1}'$ is matrix of 1 s.

To include violation of homoscedasticity assumption, we combine the structures as (CS, AR(1, 0.5), AR(1, 0.7)), (AR(1, 0.5), AR(1, 0.7), UN), where 0.5 and 0.7 are ρ values used. We use $\kappa = 1$ for all cases. For $g = 6$, we use $(n_1, n_2, \ldots, n_6) = (10, 10, 10, 20, 20, 20)$, (30, 40, 50, 30, 40, 50), (30, 40, 50, 60, 70, 80), with same covariance matrix combinations as used for $g = 3$, repeated for first three and next three populations. Due to the close similarity of the results, we restrict the presentation of power to (CS, AR, AR) combination for $g = 3$ and to normal and exponential distributions, with first two sample size sextuples, for $g = 6$.

For both size and power, we use $\alpha = 0.05$. For $g = 3$, we test all (three) pairwise hypotheses $\boldsymbol{\delta}_{ij} = \mathbf{0}$, $i < j$, $i,j = 1, 2, 3$, where for $g = 6$, we do comparisons with (sample 1 as) control, that is, $H_0 : \boldsymbol{\delta}_{1j} = \mathbf{0}$, $j = 2, \ldots, 6$. Moreover, for power, we add non-centrality parameter, defined as $\boldsymbol{\vartheta} = 0.2(0.2)1\mathbf{q}$ with $\mathbf{q} = (1/p, \ldots, p/p)$, to population 1 for both $g = 3$ and 6. This, for $g = 3$, affects tests for $\boldsymbol{\delta}_{12}$ and $\boldsymbol{\delta}_{13}$, whereas for $g = 6$ and comparisons with control, it affects all tests. The p-values and power are estimated using the asymptotic distribution in Theorem 2.7, averaged over 1000 simulation runs.

For comparison, we also compute, under the same set up, size and power for the most commonly used multiple test procedure, namely max test, $\mathrm{T}_{\max}$, with Bonferroni error control. We thus compute $\mathrm{T}_{\max} = \max\{T_{ij} : i, j = 1, \ldots G,\ i < j\}$ and use α/G as nominal

Table 1. Estimated size of pairwise comparisons for $g = 3$: all distributions.

			ND				UD				ED			
$\boldsymbol{\Sigma}_i$	n_1, n_2, n_3	p	T_{12}	T_{13}	T_{23}	$T_{\max}$	T_{12}	T_{13}	T_{23}	$T_{\max}$	T_{12}	T_{13}	T_{23}	$T_{\max}$
CS,AR,AR	10,15,20	50	0.945	0.939	0.945	0.934	0.942	0.939	0.954	0.921	0.944	0.945	0.953	0.923
		300	0.957	0.956	0.946	0.952	0.954	0.942	0.945	0.940	0.946	0.939	0.945	0.927
		500	0.946	0.940	0.948	0.933	0.945	0.938	0.942	0.928	0.940	0.945	0.944	0.939
		1000	0.940	0.951	0.947	0.947	0.942	0.939	0.944	0.931	0.938	0.945	0.947	0.925
	20,30,40	50	0.945	0.944	0.951	0.920	0.943	0.949	0.948	0.930	0.946	0.949	0.951	0.927
		300	0.953	0.954	0.959	0.945	0.962	0.960	0.958	0.954	0.950	0.949	0.955	0.933
		500	0.953	0.947	0.955	0.953	0.945	0.944	0.949	0.944	0.955	0.957	0.950	0.956
		1000	0.943	0.948	0.954	0.951	0.947	0.951	0.956	0.940	0.951	0.954	0.943	0.940
	10,30,60	50	0.950	0.945	0.955	0.915	0.946	0.946	0.951	0.929	0.943	0.938	0.941	0.908
		300	0.941	0.945	0.947	0.937	0.948	0.947	0.941	0.935	0.944	0.947	0.942	0.918
		500	0.949	0.951	0.948	0.942	0.953	0.943	0.947	0.943	0.944	0.942	0.958	0.931
		1000	0.953	0.945	0.948	0.937	0.955	0.949	0.946	0.938	0.959	0.953	0.944	0.912
	50,75,100	50	0.949	0.944	0.948	0.925	0.950	0.952	0.954	0.928	0.943	0.944	0.947	0.926
		300	0.942	0.948	0.951	0.932	0.958	0.954	0.945	0.949	0.947	0.952	0.962	0.949
		500	0.940	0.944	0.957	0.952	0.952	0.948	0.943	0.951	0.947	0.949	0.941	0.926
		1000	0.953	0.945	0.947	0.937	0.945	0.948	0.950	0.942	0.949	0.952	0.940	0.930
AR,AR,UN	10,15,20	50	0.959	0.940	0.941	0.921	0.948	0.944	0.949	0.920	0.946	0.935	0.940	0.917
		300	0.944	0.948	0.943	0.928	0.947	0.940	0.949	0.932	0.948	0.943	0.953	0.935
		500	0.943	0.941	0.946	0.927	0.950	0.940	0.960	0.939	0.938	0.940	0.944	0.925
		1000	0.948	0.949	0.953	0.925	0.941	0.956	0.954	0.936	0.940	0.942	0.951	0.938
	20,30,40	50	0.953	0.946	0.944	0.920	0.950	0.960	0.943	0.937	0.949	0.945	0.956	0.943
		300	0.947	0.942	0.941	0.932	0.940	0.956	0.949	0.943	0.946	0.961	0.944	0.958
		500	0.954	0.952	0.950	0.943	0.942	0.944	0.951	0.937	0.943	0.962	0.951	0.943
		1000	0.951	0.950	0.947	0.940	0.944	0.940	0.953	0.939	0.945	0.955	0.946	0.941
	10,30,60	50	0.940	0.941	,.947	0.922	0.944	0.940	0.951	0.924	0.953	0.944	0.942	0.923
		300	0.951	0.946	0.946	0.948	0.940	0.939	0.942	0.933	0.948	0.951	0.958	0.943
		500	0.945	0.951	0.954	0.944	0.946	0.951	0.954	0.946	0.945	0.953	0.957	0.943
		1000	0.950	0.956	0.948	0.947	0.945	0.942	0.956	0.936	0.951	0.947	0.960	0.941
	50,75,100	50	0.942	0.945	0.948	0.924	0.951	0.957	0.946	0.936	0.947	0.957	0.956	0.942
		300	0.950	0.963	0.956	0.952	0.948	0.953	0.944	0.943	0.944	0.951	0.953	0.926
		500	0.957	0.945	0.947	0.948	0.956	0.948	0.947	0.933	0.952	0.942	0.945	0.938
		1000	0.955	0.949	0.946	0.940	0.951	0.941	0.944	0.938	0.943	0.947	0.950	0.941

level to exercise Bonferroni control. Note that, both types of error control pertain to family-wise in the strong sense (FWEs); see Section 1. The estimated quantiles, $1 - \hat{\alpha}$ and power, $1 - \hat{\beta}$, are reported in Tables 1–4, respectively, for $g = 3$ and 6.

We observe an accurate size control by the proposed tests for both 3 and 6 samples, under all covariance structures and for all populations. The accuracy for exponential distribution as a serious non-normal case is particularly noticeable. Likewise is the case for the covariance structures involving CS, being highly spiked covariance matrix, with only two distinct eigenvalues. These results depict strong robustness of the tests against several violations of usual assumptions. Similar situation appears for power which steadily increases not only for increasing sample sizes but also for increasing dimension. Note the power converging quickly to 1 for sample sizes as small as 10 or 20, even for exponential distribution. Due to this, we reduce ϑ values for each p as soon as the power approaches its maximum value. For example, for $p = 500$, power was already observed 1 for $\vartheta = 0.4$, hence not reported. We also note, in comparison, that $T_{\max}$ often moves between being conservative to liberal and looses its stability, although it generally shows nice power.

To conclude, the proposed tests can be generally considered for most of practically used distributions and covariance structures, where the dimension may far exceed the sample size, and for a moderate number of independent samples. Note that, theoretically, the

Table 2. Estimated size of comparisons with a control for $g = 6$: All distributions.

			Σ_i: CS, AR, AR						Σ_i: AR, AR, UN					
$\mathcal{F}$	$n_1, \ldots, n_6$	p	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{max}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{max}
ND	(10,10,10	50	0.939	0.947	0.946	0.935	0.946	0.945	0.941	0.942	0.952	0.942	0.944	0.960
	20,20,20)	300	0.938	0.936	0.945	0.941	0.947	0.966	0.947	0.941	0.940	0.944	0.942	0.963
		500	0.941	0.943	0.957	0.944	0.946	0.974	0.953	0.940	0.944	0.940	0.945	0.971
		1000	0.946	0.954	0.952	0.953	0.948	0.965	0.940	0.949	0.950	0.954	0.944	0.970
	(30,40,50	50	0.944	0.943	0.943	0.939	0.946	0.947	0.940	0.943	0.942	0.946	0.947	0.953
	30,40,50)	300	0.946	0.940	0.942	0.942	0.946	0.965	0.945	0.943	0.959	0.947	0.944	0.975
		500	0.944	0.950	0.945	0.952	0.945	0.980	0.945	0.947	0.951	0.947	0.948	0.981
		1000	0.951	0.946	0.953	0.949	0.941	0.961	0.947	0.949	0.950	0.961	0.945	0.973
	(30,40,50	50	0.960	0.944	0.951	0.949	0.948	0.953	0.946	0.947	0.946	0.942	0.943	0.957
	60,70,80)	300	0.953	0.954	0.955	0.940	0.943	0.981	0.954	0.948	0.963	0.948	0.940	0.975
		500	0.940	0.944	0.942	0.948	0.945	0.976	0.941	0.944	0.946	0.942	0.946	0.965
		100	0.947	0.948	0.958	0.951	0.944	0.948	0.945	0.951	0.941	0.949	0.947	0.941
UD	(10,10,10	50	0.938	0.942	0.941	0.940	0.945	0.940	0.943	0.937	0.943	0.945	0.952	0.953
	20,20,20)	300	0.943	0.944	0.942	0.952	0.945	0.974	0.943	0.940	0.946	0.943	0.950	0.981
		500	0.941	0.944	0.942	0.943	0.942	0.962	0.941	0.947	0.944	0.949	0.940	0.973
		1000	0.944	0.944	0.950	0.959	0.945	0.972	0.942	0.948	0.951	0.960	0.941	0.975
	(30,40,50	50	0.952	0.946	0.947	0.944	0.949	0.963	0.961	0.947	0.941	0.948	0.948	0.971
	30,40,50)	300	0.949	0.943	0.946	0.950	0.950	0.974	0.949	0.954	0.949	0.941	0.951	0.977
		500	0.950	0.946	0.953	0.940	0.944	0.975	0.951	0.951	0.942	0.958	0.955	0.978
		1000	0.948	0.959	0.952	0.944	0.941	0.991	0.942	0.950	0.956	0.942	0.955	0.988
	(30,40,50	50	0.928	0.943	0.951	0.960	0.941	0.968	0.949	0.959	0.940	0.956	0.948	0.969
	60,70,80)	300	0.953	0.942	0.949	0.945	0.950	0.982	0.940	0.956	0.954	0.949	0.944	0.973
		500	0.953	0.950	0.950	0.948	0.944	0.973	0.948	0.940	0.952	0.940	0.946	0.985
		1000	0.957	0.958	0.949	0.951	0.948	0.938	0.955	0.952	0.941	0.943	0.940	0.939
ED	(10,10,10	50	0.943	0.945	0.931	0.936	0.938	0.958	0.950	0.956	0.949	0.960	0.958	0.957
	20,20,20)	300	0.954	0.946	0.941	0.942	0.950	0.968	0.948	0.945	0.947	0.949	0.946	0.964
		500	0.947	0.950	0.956	0.946	0.961	0.977	0.947	0.948	0.949	0.945	0.954	0.975
		1000	0.953	0.946	0.953	0.959	0.955	0.972	0.950	0.948	0.958	0.958	0.946	0.977
	(30,40,50	50	0.948	0.951	0.960	0.947	0.948	0.965	0.945	0.950	0.952	0.947	0.955	0.965
	30,40,50)	300	0.943	0.943	0.945	0.946	0.951	0.972	0.946	0.963	0.942	0.956	0.946	0.974
		500	0.956	0.950	0.955	0.947	0.956	0.980	0.949	0.946	0.941	0.948	0.945	0.973
		1000	0.942	0.0.935	0.942	0.949	0.943	0.975	0.944	0.938	0.946	0.945	0.952	0.981
	(30,40,50	50	0.951	0.946	0.945	0.949	0.952	0.963	0.952	0.946	0.945	0.941	0.945	0.958
	60,70,80)	300	0.946	0.941	0.946	0.955	0.943	0.968	0.941	0.962	0.945	0.953	0.954	0.956
		500	0.948	0.941	0.945	0.944	0.951	0.978	0.946	0.952	0.949	0.947	0.954	0.978
		1000	0.956	0.948	0.961	0.953	0.950	0.985	0.954	0.951	0.953	0.947	0.944	0.982

asymptotic covariance matrix of the vector of tests, $\boldsymbol{\Lambda}$, holds for any g, hence any G, but a large g is practically a rare phenomenon. In most cases, g is a moderate values like $g \leq 6$ or 7, as compared to p which may run into thousands. In this context, the tests may find applicability in a wide array of practical problems. On the other hand, the largeness of g may, at least in a few special contexts, be of interest and is therefore being considered for a future work. It indeed needs a different sort of asymptotics to allow for $g \to \infty$ simultaneously with $p \to \infty$.

4. Applications

We apply the proposed procedures to two data sets, heretofore called SRBCT and Species data, with $g = 4$ and 5 samples, respectively. The first data set consists of small, round blue cell tumors (SRBCT) observed over four independent groups, including a normal group, with sizes $n_1 = 29$, $n_2 = 25$, $n_3 = 11$, $n_4 = 18$, with dimension $p = 2308$ gene expressions. The second, species, data set consists of $p = 809$ species counts of macrobenthos, observed

Table 3. Estimated power of pairwise comparisons for $g = 3$: All distributions.

				ND			UD			ED		
$\boldsymbol{\Sigma}_i$	n_1, n_2, n_3	p	ϑ	T_{12}	T_{13}	$T_{\max}$	T_{12}	T_{13}	$T_{\max}$	T_{12}	T_{13}	$T_{\max}$
CS,AR,AR	10,15,20	50	0.2	0.149	0.141	0.173	0.134	0.139	0.134	0.143	0.137	0.158
			0.4	0.493	0.511	0.532	0.490	0.505	0.537	0.474	0.505	0.514
			0.6	0.905	0.947	0.949	0.900	0.928	0.943	0.902	0.943	0.963
			0.8	0.997	1.000	0.999	0.995	0.998	1.000	0.997	1.000	1.000
			1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.167	0.175	0.165	0.162	0.190	0.196	0.149	0.175	0.167
			0.4	0.679	0.742	0.745	0.658	0.726	0.755	0.668	0.711	0.742
			0.6	0.991	0.998	0.999	0.955	0.986	0.997	0.992	0.998	1.000
			0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		300	0.2	0.262	0.288	0.272	0.260	0.283	0.263	0.258	0.279	0.270
			0.4	0.961	0.982	0.981	0.962	0.981	0.976	0.959	0.980	0.983
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10,30,60	50	0.2	0.138	0.179	0.165	0.157	0.168	0.155	0.132	0.168	0.162
			0.4	0.596	0.675	0.617	0.614	0.670	0.619	0.587	0.648	0.613
			0.6	0.968	0.975	0.970	0.967	0.990	0.984	0.970	0.974	0.992
			0.8	0.999	1.000	1.000	0.998	0.999	0.999	0.999	0.998	1.000
			1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.195	0.212	0.179	0.185	0.215	0.186	0.178	0.208	0.203
			0.4	0.811	0.875	0.835	0.809	0.879	0.848	0.800	0.896	0.816
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		300	0.2	0.339	0.406	0.367	0.335	0.409	0.371	0.337	0.410	0.368
			0.4	0.996	0.996	0.999	0.995	0.997	0.998	0.996	0.998	0.999
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	50,75,100	50	0.2	0.570	0.667	0.657	0.568	0.664	0.671	0.576	0.647	0.656
			0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.812	0.863	0.884	0.808	0.868	0.863	0.812	0.876	0.869
			0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

from $n = 101$ independent sites in five different regions, with sample sizes $n_1 = 16$, $n_2 = 21$, $n_3 = 25$, $n_4 = 19$, $n_5 = 20$, along a long transact of the Norwegian continental shelf.

We have $\mathbf{X} = (\mathbf{X}_1', \ldots, \mathbf{X}_5')' \in \mathbb{R}^{n \times p}$ as complete data matrix with $\mathbf{X}_i = (\mathbf{X}_{i1}', \ldots, \mathbf{X}_{in_i}')' \in \mathbb{R}^{n_i \times p}$ for ith sample, where n_i and p are given above. Both data sets represent unbalanced one-way MANOVA designs with $g = 4$ and 5 independent samples, with dimensions $p = 2308$ and 809, and total sample size $n = \sum_{i=1}^{5} n_i = 83$ and 101, respectively.

We begin by testing global hypotheses, i.e. $H_{0g} : \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_g$ vs $H_{15} : \boldsymbol{\mu}_i \neq \boldsymbol{\mu}_j$ for at least one pair $i \neq j$, $i, j = 1, \ldots, g$, with $g = 4$ and 5, respectively. We use MANOVA test statistic proposed, under identical general conditions as used here, in Ahmad [25]. The observed values of the test statistic, T_g (see Equation 8 in the reference), for SRBCT data are 378.1604 and 45.7850, respectively, for Chi-square and normal approximations, with p-value virtually zero in each case. A detailed analysis of species data is already provided in Ahmad [25, Sec. 5], by which $T_g = 180.4$ and 40.61 for Chi-square and normal approximations, respectively, with p-values again zero. With global hypotheses strongly rejected, we expect to find vectors responsible for this rejection.

For multiple comparisons, we consider sample 1 as control and compare it with the remaining samples for Species data, i.e. we test $H_{01j} : \boldsymbol{\delta}_{1j} = \mathbf{0}$, $j = 2, 3, 4, 5$, with $G = 4$, whereas we do all $G = 6$ pairwise comparisons for SRBCT data, i.e. $H_{ij0} : \boldsymbol{\delta}_{ij} = \mathbf{0}$, $i,j = 1, 2, 3, 4$, $i < j$. The vectors of test statistics for Species and SRBCT data, respectively, are computed as

$$\mathbf{T}_5 = (5.15, 12.24, 10.98, 10.36)', \quad \mathbf{T}_6 = (5.17, 3.76, 5.32, 6.25, 5.43, 5.07)',$$

Table 4. Estimated power of comparisons with a control for $g = 6$: All distributions.

				$\mathbf{\Sigma}_i$: CS, AR, AR						$\mathbf{\Sigma}_i$: AR, AR, UN					
$\mathcal{F}$	$n_1,\ldots,n_6$	p	ϑ	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	$T_{\max}$	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	$T_{\max}$
ND	(10,10,10	50	0.2	0.134	0.149	0.145	0.157	0.167	0.127	0.118	0.122	0.149	0.157	0.155	0.128
	20,20,20)		0.4	0.425	0.416	0.536	0.571	0.545	0.538	0.396	0.420	0.533	0.542	0.540	0.521
			0.6	0.844	0.848	0.936	0.950	0.932	0.955	0.815	0.837	0.947	0.943	0.950	0.952
			0.8	0.989	0.989	0.999	1.000	0.999	1.000	0.993	0.991	1.000	0.999	0.999	1.000
			1.0	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.144	0.144	0.182	0.213	0.183	0.123	0.145	0.161	0.184	0.187	0.198	0.132
			0.4	0.581	0.567	0.759	0.765	0.758	0.750	0.573	0.565	0.720	0.738	0.743	0.730
			0.6	0.974	0.972	0.999	0.997	0.996	0.998	0.969	0.966	0.997	0.997	0.998	0.995
			0.8	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		300	0.2	0.24	0.248	0.289	0.305	0.312	0.217	0.221	0.224	0.315	0.334	0.301	0.226
			0.4	0.913	0.909	0.990	0.989	0.991	0.990	0.899	0.909	0.984	0.988	0.989	0.989
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	(30,40,50	50	0.2	0.340	0.383	0.303	0.327	0.360	0.341	0.340	0.381	0.308	0.342	0.367	0.341
	30,40,50)		0.4	0.981	0.986	0.949	0.973	0.984	0.995	0.960	0.984	0.945	0.976	0.982	0.988
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.473	0.546	0.400	0.496	0.527	0.502	0.496	0.543	0.425	0.494	0.562	0.519
			0.4	1.000	1.000	0.998	1.000	1.000	1.000	0.998	1.000	0.999	1.000	1.000	1.000
ED	(10,10,10	50	0.2	0.126	0.116	0.140	0.132	0.124	0.103	0.121	0.130	0.140	0.144	0.132	0.095
	20,20,20)		0.4	0.404	0.373	0.521	0.535	0.540	0.496	0.399	0.403	0.536	0.512	0.513	0.512
			0.6	0.851	0.826	0.951	0.945	0.941	0.974	0.823	0.815	0.942	0.950	0.956	0.974
			0.8	0.989	0.985	1.000	1.000	0.999	1.000	0.989	0.990	1.000	0.999	1.000	1.000
			1.0	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.141	0.149	0.189	0.172	0.155	0.119	0.151	0.148	0.166	0.169	0.195	0.113
			0.4	0.557	0.562	0.746	0.742	0.744	0.729	0.558	0.585	0.736	0.746	0.747	0.755
			0.6	0.965	0.976	0.997	0.998	1.000	0.999	0.961	0.965	0.998	0.998	0.999	0.999
			0.8	0.999	1.000	1.000	1.000	1.000	1.000	0.999	0.999	1.000	1.000	1.000	1.000
		300	0.2	0.220	0.222	0.287	0.287	0.322	0.218	0.210	0.204	0.304	0.289	0.283	0.202
			0.4	0.902	0.906	0.986	0.985	0.988	0.991	0.913	0.907	0.992	0.986	0.991	0.996
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	v	1.000	1.000	1.000
	(30,40,50	50	0.2	0.339	0.366	0.310	0.335	0.339	0.344	0.351	0.365	0.275	0.331	0.369	0.334
	30,40,50)		0.4	0.975	0.987	0.954	0.977	0.995	0.999	0.977	0.979	0.935	0.978	0.987	0.989
			0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		100	0.2	0.453	0.545	0.402	0.502	0.528	0.500	0.489	0.541	0.413	0.516	0.511	0.494
			0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	1.000	1.000	1.000

with the corresponding vectors of p-values $(0.004, 0, 0, 0)'$ and $\mathbf{0}_6$. The results indicate all means, statistically, discernably different from each other at any reasonable nominal size. For further assessment, we also compute the $\mathbf{\Lambda}$ matrix (see Equation 11) for the two data sets, respectively, of order 4×4 and 5×5, shown in Equations (17) and (19). It may be mentioned that the analysis reported above is based on Chi-square approximation which, as already discussed, has relatively better performance than the normal one, and the ratio in Assumption 2.6 is assumed to vanish, so that $\hat{\mathbf{\Lambda}}$ are used as diagonal matrices. This can be easily witnessed from the matrices computed for the two data sets. It is clear that ignoring the off-diagonal elements does not cause much loss of information concerning the comparisons.

To expand more on this, and to highlight an additional important property of the proposed tests, $\hat{\mathbf{\Lambda}}^{-1}$ is also reported in each case; Equations (18) and (20). First, we observe that, estimated $\mathbf{\Lambda}$ is a non-singular matrix, hence can be inverted, something that in fact can be shown for $\hat{\mathbf{\Lambda}}_G$ in general. Second, this in turn implies that the tests can be defined as affine-invariant, using $\hat{\mathbf{\Lambda}}^{-1}$. As we have not proved this inverse for the general case explicitly, it is left for a later work. Finally, we notice that the off-diagonal elements virtually vanish in the inverses. Thus, in affine-invariant form, the tests may be used even more safely under Assumption 2.6.

$$\hat{\mathbf{\Lambda}} = \begin{pmatrix} 2.326 & 0.032 & 0.068 & 0.055 \\ 0.032 & 3.459 & 0.163 & 0.131 \\ 0.068 & 0.163 & 2.846 & 0.275 \\ 0.055 & 0.131 & 0.275 & 4.177 \end{pmatrix} \tag{17}$$

$$\hat{\mathbf{\Lambda}}^{-1} = \begin{pmatrix} 0.430 & -0.003 & -0.009 & -0.005 \\ -0.003 & 0.290 & -0.016 & -0.008 \\ -0.009 & -0.016 & 0.355 & -0.023 \\ -0.005 & -0.008 & -0.023 & 0.241 \end{pmatrix} \tag{18}$$

$$\hat{\mathbf{\Lambda}} = \begin{pmatrix} 15.559 & 0.014 & 0.022 & 0.019 & 0.028 & 0.000 \\ 0.014 & 7.272 & 0.016 & 0.090 & 0.000 & 0.096 \\ 0.022 & 0.016 & 5.748 & 0.000 & 0.036 & 0.026 \\ 0.019 & 0.090 & 0.000 & 8.358 & 0.018 & 0.014 \\ 0.028 & 0.000 & 0.036 & 0.018 & 6.695 & 0.023 \\ 0.000 & 0.096 & 0.026 & 0.014 & 0.023 & 5.655 \end{pmatrix} \tag{19}$$

$$\hat{\mathbf{\Lambda}}^{-1} = \begin{pmatrix} 0.064 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.138 & 0.000 & -0.002 & 0.000 & -0.002 \\ 0.000 & 0.000 & 0.174 & 0.000 & -0.001 & -0.001 \\ 0.000 & -0.002 & 0.000 & 0.120 & 0.000 & 0.000 \\ 0.000 & 0.000 & -0.001 & 0.000 & 0.149 & -0.001 \\ 0.000 & -0.002 & -0.001 & 0.000 & -0.001 & 0.177 \end{pmatrix} \tag{20}$$

5. Discussion and conclusions

In the context of multi-sample multivariate problem, multiple comparisons of mean vectors with very large dimension, possibly much larger than the number of vectors in any sample, are considered. The case is of frequent interest, for example, as a first *post hoc* assessment of mean vectors after a global MANOVA hypothesis is rejected. All possible pairwise differences and comparisons with a control are treated. In particular, the joint asymptotic distribution, under $n_i, p \to \infty$, is derived whose tail probabilities can be directly used to carry out the multiple tests. Simulations results are used to show the accuracy of the tests, and a comparison with max test is also given.

Following the objectives of the present work, as stated in Section 1, the proposed tests can be used in applied problems requiring simultaneous inference for two or more large mean vectors which might have been sampled from a non-normal distribution and may have unequal covariance matrices as well as the sample sizes. Whereas the test statistics are asymptotically approximated with Chi-square and Normal distributions, it is observed that the former provides relatively better accuracy than the later and is thus highly recommended for practical use.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

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

References

[1] Seber GAF. Multivariate observations. New York (NY): Wiley.
[2] Johnson RA, DW Wichern. Applied multivariate statistical analysis. 6th ed. Upper Saddle River (NJ): Pearson Education; 2007.
[3] Krishnaiah PR. On the simultaneous ANOVA and MANOVA tests. Ann Inst Stat Math. 1965;17:35–53.
[4] Krishnaiah PR. Simultaneous test procedures under general MANOVA models. In: Krishnaiah PR, editor. Multivariate analysis. Vol II. New York (NY): Academic Press; 1969. p. 121–143.
[5] Wijsman RA. Constructing all smallest simultaneous confidence sets in a given class with applications to MANOVA. An Stat. 1979;7(5):1003–1018.
[6] Kropf S. Hochdimensionale multivariate Verfahren in der medizinischen Statistik. Aachen: Shaker; 2000.
[7] Kropf S, Läuter J. Multiple tests for different sets of variables using a data-driven ordering of hypotheses, with an application to gene expression data. Biom J. 2002;44:789–800.
[8] Westfall P, Kropf S, Finos L. Weighted FWE-controlling methods in high-dimensional situations. In: Benjamini Y, Bretz F, Sarakr SK, editors. Recent developments in multiple comparison procedures. Vol. 47, IMS Lecture Notes and Monpgraph Series; 2004. p. 143–154.
[9] Läuter J, Glimm E, Eszlinger M. Search for relevant sets of variables in a high-dimensional setup keeping the familywise error rate. Statist Neerl. 2005;59(3):298–312.
[10] Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of *p* values for multiple correlated tests. Amer J Human Genet. 2007;81:1158–1168.
[11] Westfall P, Troendle JF. Multiple testing with minimal assumptions. Biom J. 2008;50:745–755.
[12] Bretz F, Hothorn T, Westfall P. Multiple comparisons using R. Boca Raton (FL): CRC Press; 2011.
[13] Dickhaus T. Simultaneous statistical inference in dynamic factor models. Berlin: Humboldt-Universitätzu; 2012. (Discussion paper, 2012-033.).
[14] Goeman J, Finos L. The inheritance procedure: Multiple testing of tree-structured hypotheses. Stat App Genet Molec Biol. 2012;11:1–18.
[15] Goeman J, Solari A. Multiple hypothesis testing in genomics. Stat Med. 2014;33:1946–1978.
[16] Guilbaud O. Simultaneous confidence regions for closed tests, including Holms-, Hochberg-, and Hommel-related procedures. Biom J. 2012;54:317–342.
[17] Guilbaud O. Sharper Confidence Intervals for Hochberg- and Hommel-Related Multiple Tests Based On an Extended Simes Inequality. Statist Biopharm Res. 2014;6:123–136.
[18] Dickhaus T. Simultaneous statistical inference: with applications in the life sciences. New York (NY): Springer; 2014.
[19] Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res. 2003;12:419–446.
[20] Hochberg Y, Tamhane AC. Multiple comparison procedures. New York (NY): Wiley; 1987.
[21] Hemerik J, Goeman J. False discovery proportion estimation by permutations: confidence for significance analysis of microarrays. J R Statist Soc B. 2018;80:137–155.
[22] van der Vaart AW. Asymptotic statistics. Cambridge: Cambridge University Press; 1998.
[23] Lee AJ. U-statistics: theory and practice. Boca Raton (FL): CRC Press; 1990.
[24] Ahmad MR. Location-invariant multi-sample U-tests for covariance matrices with large dimension. Scand J Stat. 2017;44:500–523.
[25] Ahmad MR. A unified approach to testing mean vectors with large dimension. AStA Adv Stat Anal. 2018.
[26] Jiang J. Large sample techniques for statistics. New York (NY): Springer; 2010.
[27] Koroljuk VS, Borovskich YV. Theory of U-statistics. Dordrecht: Kluwer Academic Press; 1994.

Appendix. Some basic moments

Consider U_i with symmetric kernel $h(\mathbf{X}_{ik}, \mathbf{X}_{ir})$ and conditional expectation (projection) $h_c(\cdot) = \mathrm{E}[h(\cdot)|\mathbf{X}_{1k}, \ldots \mathbf{X}_c)$, $c = 1{,}2$ so that $h_1(\mathbf{X}_{ik}) = \mathrm{E}[h(\cdot)|\mathbf{X}_{ik}]$, $h_2(\cdot) = h(\cdot)$ with $\mathrm{Var}[h_i(\cdot)] = \xi_i$, $i = 1{,}2$. For U_{ij} with symmetric kernel $h(\mathbf{X}_{ik}, \mathbf{X}_{jl})$ with $m_1 = 1 = m_2$ and with $c_1, c_2 = 0, 1$, the conditional expectations are $h_{10}(\mathbf{X}_{ik}) = \mathrm{E}[h(\cdot)|\mathbf{X}_{ik}]$, $h_{01}(\mathbf{X}_{jl})$, $h_{11}(\cdot) = h(\cdot)$ with corresponding variances $\xi_{10}, \xi_{01}, \xi_{11}$. Here, $h(\cdot)$ is used when the arguments are evident from the context. Then, the moments of U-statistics follow as given, e.g., in Koroljuk and Borovskich [27] or van der Vaart [22]; see also Ahmad [25, Appendix A]

Using these notations, $\mathrm{E}(U_i) = \boldsymbol{\mu}_i'\boldsymbol{\mu}_i$, $\mathrm{E}(U_{ij}) = \boldsymbol{\mu}_i'\boldsymbol{\mu}_j$, with $h(\mathbf{X}_{ik}, \mathbf{X}_{ir}) = \mathbf{X}_{ik}'\mathbf{X}_{ir}$, $h_1(\mathbf{X}_{ik}) = \boldsymbol{\mu}_i'\mathbf{X}_{ik}$, $\xi_1 = \mathrm{Var}[h_i(\cdot)] = \boldsymbol{\mu}_i'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_i$ and $\xi_2 = \mathrm{Var}[h(\cdot)] = \mathrm{tr}(\boldsymbol{\Sigma}_i^2) + 2\boldsymbol{\mu}_i'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_i$. For U_{ij}, $h(\mathbf{X}_{ik}, \mathbf{X}_{jl}) = \mathbf{X}_{ik}'\mathbf{X}_{jl}$, with $h_{10} = \boldsymbol{\mu}_j'\mathbf{X}_{ik}$, $h_{01} = \boldsymbol{\mu}_i'\mathbf{X}_{jl}$, $\xi_{10} = \mathrm{Var}[h_{10}(\cdot)] = \boldsymbol{\mu}_j'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_j$, $\xi_{10} = \mathrm{Var}[h_{10}(\cdot)] = \boldsymbol{\mu}_i'\boldsymbol{\Sigma}_j\boldsymbol{\mu}_i$, $h_{11}(\cdot) = h(\cdot)$, $\xi_{11} = \mathrm{Var}[h_{11}(\cdot)] = \boldsymbol{\mu}_i'\boldsymbol{\Sigma}_j\boldsymbol{\mu}_i + \boldsymbol{\mu}_j'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_j + \mathrm{tr}(\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j)$. Now $\mathrm{Var}(U_i) = 2[2(n_i - 1)\boldsymbol{\mu}_i'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_i + \mathrm{tr}(\boldsymbol{\Sigma}_i^2)]/n_i(n_i - 1)$, $\mathrm{Var}(U_{ij}) = [n_i\boldsymbol{\mu}_i'\boldsymbol{\Sigma}_j\boldsymbol{\mu}_i + n_j\boldsymbol{\mu}_j'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_j + \mathrm{tr}(\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_j)]/n_in_j$, $\mathrm{Cov}(U_i, U_{ij}) = 2\boldsymbol{\mu}_j'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_i/n_i$, $\mathrm{Cov}(U_j, U_{ij}) = 2\boldsymbol{\mu}_i'\boldsymbol{\Sigma}_j\boldsymbol{\mu}_j/n_j$, $\mathrm{Cov}(U_{ij}, U_{ij'}) = \boldsymbol{\mu}_j'\boldsymbol{\Sigma}_i\boldsymbol{\mu}_{j'}/n_i$, $\mathrm{Cov}(U_{ij}, U_{i'j}) = \boldsymbol{\mu}_i'\boldsymbol{\Sigma}_j\boldsymbol{\mu}_{i'}/n_j$, $i \neq j$, $i \neq j'$, $i' \neq j$, where the remaining covariances vanish by independence.