

Nonparametric Methods for Unbalanced Multivariate Data and Many Factor Levels

Solomon W. Harrar ^{*}
and Arne C. Bathke ^{**}

Abstract

We propose different nonparametric tests for multivariate data and derive their asymptotic distribution for unbalanced designs in which the number of factor levels tends to infinity (large a , small n_i case). Quasi gratis, some new parametric multivariate tests suitable for the large a asymptotic case are also obtained. Finite sample performances are investigated and compared in a simulation study. The nonparametric tests are based on separate rankings for the different variables. In the presence of outliers, the proposed nonparametric methods have better power than their parametric counterparts. Application of the new tests is demonstrated using data from plant pathology.

Key words: Multivariate Analysis of Variance, Nonnormality, Nonparametric Model, Ordinal Data, Rank statistic, Unbalanced Design.

Mathematics Subject Classification 2000: 62G10, 62G20, 62H10, 62H15, 62J10.

1 INTRODUCTION

Multivariate data occurs naturally in several scientific fields, as for example in agriculture, biology, medicine, and the social sciences. In many situations, it is not reasonable to assume that the observations follow a Gaussian distribution, in particular when the responses are scores on an ordinal scale, and therefore application of normal theory methods is not appropriate, and a nonparametric approach is desirable. Recently, Munzel and Brunner [1,2] (see also [3]) have proposed different tests for multivariate data that are completely nonparametric and allow for arbitrary distributions of the response variables, including discrete distributions, and for arbitrary dependence between the different variables. Munzel and Brunner's asymptotic theory is aimed at the situation where the sample sizes are large, compared

^{*}Solomon W. Harrar is Assistant Professor, Department of Mathematical Sciences, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA; Email: harrar@mso.umt.edu

^{**}Arne C. Bathke is Assistant Professor, Department of Statistics, University of Kentucky, 875 Patterson Office Tower, Lexington, KY 40506-0027, USA; Email: arne@ms.uky.edu

to the number of samples (large n_i , small a). The focus of this manuscript is the situation in which the number of samples is larger than the sample sizes (large a , small n_i). This is often the case, for example, in agricultural screening trials (see, e.g., [4,5]), or in survey data with large number of strata and few observations per stratum. Bathke and Harrar [6,7] have proposed and evaluated different types of multivariate nonparametric tests for balanced data. Balanced data facilitates a rather elegant theory in the derivation of tests for multivariate factorial designs, and for relatively simple limiting distributions of the test statistics. However, unfortunately, many real data sets are not balanced. In the present manuscript, we propose different nonparametric tests for unbalanced multivariate data and derive their asymptotic distribution as $a \rightarrow \infty$, (whereas n_i is assumed bounded). As already the case in a simple linear model, unbalanced multivariate designs can easily lead to formidable algebraic expressions, and to controversies about which types of sums of squares are appropriate. Here, we investigate three different ways to define “sums of squares”. For each way, we derive the asymptotic distribution of different types of statistics, namely Dempster-ANOVA-Type, Lawley-Hotelling-Type, and Bartlett-Nanda-Pillai-Type. In general, none of these three types is uniformly better than the other two. In fact, which test is preferable depends, among other things, on the – usually unknown – correlation structure between the different variables that form the multivariate observations. This is also confirmed in the simulation study in the present manuscript.

In our theoretical derivation of the asymptotic distributions, we make use of the fact that the three types of statistics under consideration can be expressed in similar ways. For each combination of sums of squares, we prove one general theorem regarding the asymptotic distribution of classes of multivariate rank statistics, and the results about the individual types of tests then follow as simple corollaries.

Munzel and Brunner’s, as well as our approach to define multivariate nonparametric tests is based on separate rankings of the different variables, as opposed to Thompson [8,9] who proposed multivariate tests based on overall rankings across the variables. Using separate rankings has some important advantages. First, tests based on separate rankings are invariant under separate monotone transformations of the response variables. Consider, for example, the variables *height* and *weight*: It should not matter whether they are measured in centimeters and grams, or in inches and pounds. Separate monotone transformations of the individual variables should not affect the results of the tests. This can only be achieved in separate rankings. Furthermore, it is commonly the case that the different variables are measured on completely different scales for which an overall ranking is not sensible. The separate ranking approach can even be applied when the different variables are measured on different types of scales (e.g., ordinal and quantitative). In addition, in the case that measurements of different variables are independent, separate rankings preserve the independence across the variables, whereas an overall ranking induces dependence across all variables.

Early work on multivariate nonparametric methods includes Puri and Sen [10,11] who also used separate rankings for the different variables. They considered a Wald-type statistic, assuming semiparametric location models with continuous population distributions. In Munzel and Brunner [1,2] as well as in the present manuscript, the model is completely nonparametric, and the distribution functions are allowed to be arbitrary.

The large a asymptotic behavior of nonparametric multivariate tests for unbalanced data has, to our knowledge, not been addressed previously in the literature. In addition, we would like to point out that the new asymptotic results in this manuscript are applicable not only to the ranks, but also to the original observations if we assume that the fourth moments of the population distributions exist. Parametric versions of the tests described in Sections 2.5 and 2.6 have not been used in the literature before. One principal advantage of these tests is that they could be applied in situations where the covariance matrix is not constant from group to group.

Outline of the paper. In the following Section 2.1, we define the nonparametric hypotheses under consideration in this manuscript, as well as some matrix-valued quadratic forms that we will use to test these hypotheses. Furthermore, we introduce the terms “rank transform” and “asymptotic rank transform”. In the subsequent sections, we derive the asymptotic ($a \rightarrow \infty$) null distribution of several newly proposed test statistics that are based on the matrix-valued quadratic forms defined in Section 2.1. This derivation is broken up into different parts, with some preliminary work done in Sections 2.2 and 2.3. Specifically, in Section 2.2, we prove the consistency of different covariance matrix estimators, whereas in Section 2.3, we establish the asymptotic equivalence of certain expressions defined in terms of rank transforms and the corresponding expressions in terms of asymptotic rank transforms. Finally, in Sections 2.4-2.6, we derive the asymptotic (as $a \rightarrow \infty$) distribution of all test statistics under consideration, making use of the results obtained in the previous sections. We compare the finite sample behavior of the different tests in a simulation study that is described in Section 3. Furthermore, the new results are applied to a real data set in Section 4. All proofs are deferred to the appendix.

2 DEFINITION AND ASYMPTOTIC PROPERTIES

2.1 Basic Definitions

Consider an unbalanced design with a independent samples of multivariate observations. In the following, the index $i = 1, \dots, a$ denotes the group, $j = 1, \dots, n_i$ denotes the subjects per group, and $k = 1, \dots, p$ denotes the different variables measured on the same subject.

The observations are modeled by independent random vectors $\mathbf{X}_{ij} = (X_{ij}^{(1)}, \dots, X_{ij}^{(p)})'$, $i = 1, \dots, a$, $j = 1, \dots, n_i$, with possibly dependent components. The dependence structure can be arbitrary. Denote the multivariate distributions by $\mathbf{X}_{ij} \sim F_i$, and the marginal distributions by $X_{ij}^{(k)} \sim F_i^{(k)}$, $k = 1, \dots, p$. The marginal distributions $F_i^{(k)}$ are assumed to be nondegenerate. Throughout the manuscript, we use the normalized version of the cumulative distribution function, $F_i^{(k)}(x) = \frac{1}{2}P(X_{ij}^{(k)} \leq x) + \frac{1}{2}P(X_{ij}^{(k)} < x)$ [12-15]. A factorial design can be modeled within the context of this setup by imposing a structure on the index i . However, this paper focuses on the one-way layout. The more general asymptotic results can be transferred to higher-way layouts, but notation and asymptotic variance become rather involved. An exhaustive treatment of higher-way layouts would exceed scope and length of this manuscript, and it is therefore deferred to a separate treatise.

The nonparametric hypotheses can be stated either in terms of the multivariate distributions or in terms of the marginal distributions. For example, in the nonparametric one-way layout, we consider the multivariate null hypothesis $H_0 : F_1 = \dots = F_a$ and the marginal null hypothesis $\bar{H}_0 : F_1^{(k)} = \dots = F_a^{(k)}$, $k = 1, \dots, p$. The multivariate hypothesis is stronger than the marginal hypothesis in the sense that the first implies the latter. Therefore, every result that is proved under the marginal hypothesis is also true under the stronger multivariate hypothesis, but not vice versa, unless the marginal distributions are independent. Also, note that in general the nonparametric hypotheses formulated in terms of distribution functions imply the linear model hypotheses formulated in terms of the corresponding expected values. To see this, consider an arbitrary contrast defined by weights w_i , $i = 1, \dots, a$. If for the contrast in terms of distribution functions, $\sum_{i=1}^a w_i F_i = 0$ is true, then the following holds for the same contrast in terms of expectations:

$$\sum_{i=1}^a w_i \mu(F_i) = \sum_{i=1}^a w_i \int x dF_i(x) = \int x d \sum_{i=1}^a w_i F_i(x) = 0$$

In the subsequent theorems, we have strived to provide transparency as to which of the results are true under the weaker marginal hypothesis, and we have explicitly indicated when the stronger multivariate hypothesis allows for stronger results. That is, if a result is true under the weaker marginal hypothesis \bar{H}_0 , then we have stated it under \bar{H}_0 , in some cases followed by a stronger result obtained under H_0 .

The test statistics considered in this paper use separate rankings for the p different variables. This is motivated by possible applications where each variable is measured on a different scale. In fact, the tests considered here can even be used when some of the variables are ordinal, while others are measured on a numerical scale. In such a case, it would not make sense to rank observations across all variables. Also, as we have mentioned above, only separate rankings allow for invariance under monotone transformations of the different variables. In addition, in the case where the different variables are independent, the separate ranking preserves this independence, while an overall ranking would introduce dependence across the variables. Furthermore, as an anonymous referee pointed out, both H_0 and \bar{H}_0 are invariant under the group of marginal monotone transformations \mathcal{G} , say. So the invariance principle suggests basing tests on maximal invariant statistics for \mathcal{G} , which are the separate (componentwise) ranks.

Define $\mathbf{R} = (\mathbf{R}_{11}, \dots, \mathbf{R}_{1n_1}, \mathbf{R}_{21}, \dots, \mathbf{R}_{an_a})$, where $\mathbf{R}_{ij} = (R_{ij}^{(1)}, \dots, R_{ij}^{(p)})'$, and $R_{ij}^{(k)}$ is the (mid-)rank of $X_{ij}^{(k)}$ among all $N = \sum_{i=1}^a n_i$ random variables $X_{11}^{(k)}, \dots, X_{an_a}^{(k)}$. Note that \mathbf{R} is a matrix of dimension $p \times N$. Denote $\bar{\mathbf{R}}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{R}_{ij}$, $\bar{\mathbf{R}}_{..} = \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \mathbf{R}_{ij}$ the weighted average, and $\tilde{\mathbf{R}}_{..} = \frac{1}{a} \sum_{i=1}^a \bar{\mathbf{R}}_{i.}$ the unweighted average of the rank vectors. Let I_m be the m -dimensional identity matrix, $\mathbf{1}_m$ the $m \times 1$ column vector consisting of ones, $J_m = \mathbf{1}_m \cdot \mathbf{1}_m'$ the $m \times m$ matrix of ones, and $P_m = I_m - \frac{1}{m} J_m$ the so-called centering matrix. We denote by A^- the Moore–Penrose generalized inverse of A . Besides its uniqueness and existence, this inverse also defines a continuous mapping (see Schott [16], Chapter 5).

We have investigated rank-based versions of three different types of test statistics that are popular in parametric multivariate inference, namely Lawley-Hotelling-Type, Bartlett-Nanda-Pillay-Type, and

Dempster-ANOVA-Type. It is possible to derive a coherent theory for these three types because they all can be defined similarly in terms of the following $p \times p$ matrices that are based on the rank matrix \mathbf{R} .

$$H_1 = \frac{1}{a-1} \sum_{i=1}^a n_i (\bar{\mathbf{R}}_i - \bar{\mathbf{R}}_{..})(\bar{\mathbf{R}}_i - \bar{\mathbf{R}}_{..})' = \frac{1}{a-1} \mathbf{R} \left(\bigoplus_{i=1}^a \frac{1}{n_i} J_{n_i} - \frac{1}{N} J_N \right) \mathbf{R}' \quad (1)$$

$$H_2 = \frac{1}{a-1} \sum_{i=1}^a (\bar{\mathbf{R}}_i - \tilde{\mathbf{R}}_{..})(\bar{\mathbf{R}}_i - \tilde{\mathbf{R}}_{..})' = \frac{1}{a-1} \mathbf{R} \left[\left(\bigoplus_{i=1}^a \frac{1}{n_i} \mathbf{1}_{n_i} \right) P_a \left(\bigoplus_{i=1}^a \frac{1}{n_i} \mathbf{1}'_{n_i} \right) \right] \mathbf{R}' \quad (2)$$

$$G_1 = \frac{1}{N-a} \sum_{i=1}^a \sum_{j=1}^{n_i} (\mathbf{R}_{ij} - \bar{\mathbf{R}}_i)(\mathbf{R}_{ij} - \bar{\mathbf{R}}_i)' = \frac{1}{N-a} \mathbf{R} \left(\bigoplus_{i=1}^a P_{n_i} \right) \mathbf{R}' \quad (3)$$

$$\begin{aligned} G_2 &= \frac{1}{a-1} \sum_{i=1}^a \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i-1} \sum_{j=1}^{n_i} (\mathbf{R}_{ij} - \bar{\mathbf{R}}_i)(\mathbf{R}_{ij} - \bar{\mathbf{R}}_i)' \\ &= \frac{1}{a-1} \mathbf{R} \left[\bigoplus_{i=1}^a \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i-1} P_{n_i} \right] \mathbf{R}' \end{aligned} \quad (4)$$

$$G_3 = \frac{1}{a} \sum_{i=1}^a \frac{1}{n_i(n_i-1)} \sum_{j=1}^{n_i} (\mathbf{R}_{ij} - \bar{\mathbf{R}}_i)(\mathbf{R}_{ij} - \bar{\mathbf{R}}_i)' = \frac{1}{a} \mathbf{R} \left(\bigoplus_{i=1}^a \frac{1}{n_i(n_i-1)} P_{n_i} \right) \mathbf{R}' \quad (5)$$

Each of the different test statistics proposed below will be based on one of the three pairs (H_1, G_1) , (H_1, G_2) , or (H_2, G_3) . Note that in a balanced design with $n_i \equiv n$, $i = 1, \dots, a$, the matrices G_1 and G_2 are identical, and furthermore, $H_2 = n^{-1} \cdot H_1$ and $G_3 = n^{-1} \cdot G_1 = n^{-1} \cdot G_2$. Because of these relations, in a balanced design, each of the three pairs will lead to the same test statistic. While the first pair (H_1, G_1) represents a straightforward extension of the univariate one-way analysis of variance (see, e.g., Scheffé [17], p.59), the multivariate test statistics based on (H_1, G_1) have a rather complicated asymptotic ($a \rightarrow \infty$) variance, which led us to consider other matrix pairs. The matrices H_2 , G_2 and G_3 can be motivated from the MANOVA approach for constructing tests. For example, in multivariate mixed models, it is a common practice to construct tests for fixed and random effects by comparing sums of squares and cross-product matrices under the null hypothesis. Let $H_2(\mathbf{X})$, $G_2(\mathbf{X})$ and $G_3(\mathbf{X})$ be defined in the same way as H_2 , G_2 and G_3 , respectively, but based on the original observations matrix \mathbf{X} . Consider the null hypothesis $\bar{H}_0 : F_1^{(k)} = \dots = F_a^{(k)}$, $k = 1, \dots, p$. Let $\text{Var}(\mathbf{X}_{ij}) = \Omega_i$. Then,

$$E_{H_0}[H_1(\mathbf{X})] = E[G_2(\mathbf{X})] = \frac{1}{a-1} \sum_{i=1}^a \left(1 - \frac{n_i}{N}\right) \Omega_i \quad \text{and}$$

$$E_{H_0}[H_2(\mathbf{X})] = E[G_3(\mathbf{X})] = \frac{1}{a} \sum_{i=1}^a \frac{1}{n_i} \Omega_i.$$

Therefore, it makes sense to compare the matrices in each of the pairs $(H_1(\mathbf{X}), G_2(\mathbf{X}))$ and $(H_2(\mathbf{X}), G_3(\mathbf{X}))$ to construct valid tests for H_0 . For univariate designs, these sums of squares have been investigated recently [18,19]. Furthermore, the ANOVA-Type and Lawley-Hotelling-Type test statistics proposed in

[1,2] in the context of the large n asymptotic setting are closely related to those in the present manuscript that are based on the matrix pair (H_2, G_3) . However, due to the different asymptotic context, the variance estimators derived in [1,2] would not be consistent in the present manuscript (large a asymptotics). A caveat in using tests based on any of these pairs is that the tests may not be invariant to affine transformations. That is, let $\mathbf{R}(\mathbf{X})$ denote the matrix of componentwise ranks of \mathbf{X} , and C be a $p \times p$ positive definite matrix. In general, $\mathbf{R}(C\mathbf{X}) \neq C\mathbf{R}(X)$ unless C is a diagonal matrix. Therefore the test statistics considered in this paper will not be affine invariant. Nonparametric affine invariant sign and signed rank tests for the multi-sample location problem have been considered in the work of Oja and colleagues [20,21,22]. For a comprehensive review, see also [23]. In these manuscripts, the population distributions are assumed to belong to the class of absolutely continuous and symmetric location families of distributions.

For the mathematical derivations in the technical proofs of this manuscript, it is convenient to use the so-called ‘‘asymptotic rank transforms’’ (ART) and ‘‘rank transforms’’ (RT). They are formally introduced in the following definition. The concept of ART was proposed by Akritas [24].

Definition 1. Let $\mathbf{X}_{ij} = (X_{ij}^{(1)}, \dots, X_{ij}^{(p)})'$, $i = 1, \dots, a$, $j = 1, \dots, n_i$, be independent random vectors with possibly dependent components $X_{ij}^{(k)} \sim F_i^{(k)}$, $k = 1, \dots, p$. Let $N = \sum_{i=1}^a n_i$.

$$H^{(k)}(x) = \frac{1}{N} \sum_{i=1}^a n_i F_i^{(k)}(x) \quad \text{denotes the average cdf for variable } k; \quad (6)$$

$$\hat{H}^{(k)}(x) = \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} c(x - X_{ij}^{(k)}), \quad \text{where } c(t) = \begin{cases} 0, & t < 0 \\ 1/2, & t = 0 \\ 1, & t > 0, \end{cases} \quad (7)$$

denotes the average empirical cdf;

$\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_a)$, where $\mathbf{Y}_i = (\mathbf{Y}_{i1}, \dots, \mathbf{Y}_{in_i})$ and $\mathbf{Y}_{ij} = (Y_{ij}^{(1)}, \dots, Y_{ij}^{(p)})'$, and finally $Y_{ij}^{(k)} = H^{(k)}(X_{ij}^{(k)})$, is the $p \times N$ matrix of asymptotic rank transforms (ART). (8)

The matrix of rank transforms (RT), $\hat{\mathbf{Y}}$, is defined analogously, with elements $\hat{Y}_{ij}^{(k)} = \hat{H}^{(k)}(X_{ij}^{(k)})$. Furthermore, $M = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_a)$, $\boldsymbol{\mu}_i = (\boldsymbol{\mu}_{i1}, \dots, \boldsymbol{\mu}_{in_i})$ where $\boldsymbol{\mu}_{ij} = (\mu_{ij}^{(1)}, \dots, \mu_{ij}^{(p)})'$ is the vector of expectations of the ART vector \mathbf{Y}_{ij} , that is $\mu_{ij}^{(k)} = \mathbf{E} Y_{ij}^{(k)}$, and $\mathbf{Y}_\mu = \mathbf{Y} - M$, $\hat{\mathbf{Y}}_\mu = \hat{\mathbf{Y}} - M$.

The expression ‘‘rank transform’’ pays tribute to the fact that $\hat{Y}_{ij}^{(k)}$ is related to the (mid-)rank $R_{ij}^{(k)}$ of the random variable $X_{ij}^{(k)}$ among all N observations $X_{11}^{(k)}, \dots, X_{an_a}^{(k)}$ by $\hat{Y}_{ij}^{(k)} = N^{-1}(R_{ij}^{(k)} - \frac{1}{2})$. However, the ‘‘asymptotic rank transforms’’ are technically more tractable than the ‘‘rank transforms’’, due to the simpler covariance structure of \mathbf{Y} as compared to $\hat{\mathbf{Y}}$. Note that the ART of independent random variables are independent, but the RT are not.

We denote the ART analogs of the matrices H_i and G_i defined in (1)–(5) by \tilde{H}_i and \tilde{G}_i , respectively. In order to prove asymptotic normality results for the rank-based test statistics considered in this paper (Sections 2.4 to 2.6), we need to first establish the asymptotic equivalence of certain quadratic forms

defined in terms of H_i , G_i (based on “rank transforms”) and the corresponding quadratic forms defined in terms of \tilde{H}_i , \tilde{G}_i (based on “asymptotic rank transforms”). This will be done in Section 2.3. Before that, we prove the consistency (as a tends to infinity) of different variance estimators that we will use later.

2.2 Consistent Variance Estimation

In this section, we prove that the matrices G_i , $i = 1, 2, 3$, as well as a linear combination of H_1 and G_1 , generate consistent estimators of covariance matrices of the ART. Note that the first two theorems are proved under the stronger multivariate hypothesis, whereas for Theorem 3, the weaker marginal hypothesis is assumed.

Denote $\text{Var}(\mathbf{Y}_{i1}) = \Sigma_i$ and assume that the following limits exist:

$$\lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \Sigma_i = \Sigma \quad \text{and} \quad \lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \frac{1}{n_i} \Sigma_i = \check{\Sigma},$$

where these limits have diagonal elements σ_k^2 and $\check{\sigma}_k^2$, and off-diagonal elements σ_{kl} and $\check{\sigma}_{kl}$, respectively.

Theorem 1. *Under the multivariate null hypothesis $H_0 : F_1 = \dots = F_a$, as $a \rightarrow \infty$, n_i bounded,*

$$\frac{1}{N^2} G_1 \xrightarrow{p} \Sigma_1.$$

Theorem 2. *Under the null hypothesis and the assumptions of Theorem 1,*

$$\frac{1}{N^2(N-1)} ((a-1)H_1 + (N-a)G_1) \xrightarrow{p} \Sigma_1.$$

Theorem 3. *Under the marginal null hypothesis $\bar{H}_0 : F_1^{(k)} = \dots = F_a^{(k)}$, $k = 1, \dots, p$, as $a \rightarrow \infty$, n_i bounded,*

$$\frac{1}{N^2} G_2 \xrightarrow{p} \Sigma \quad \text{and} \quad \frac{1}{N^2} G_3 \xrightarrow{p} \check{\Sigma},$$

$$\text{where} \quad \Sigma = \lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \Sigma_i, \quad \check{\Sigma} = \lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \frac{1}{n_i} \Sigma_i, \quad \text{and} \quad \Sigma_i = \text{Var}(\mathbf{Y}_{i1}),$$

assuming that the limits exist.

2.3 Asymptotic Equivalence

In the following theorems, we prepare the derivations of asymptotic distributions that will follow in the subsequent sections by first establishing asymptotic equivalence results of certain expressions defined in terms of rank transforms and the corresponding expressions in terms of asymptotic rank transforms.

Our final goal is to derive asymptotic results for rank-based test statistics. However, it is technically easier to derive those results for the asymptotic rank transforms that have a much simpler covariance structure. The theorems in this section justify this approach by providing the connection between rank transforms and asymptotic rank transforms.

Theorem 4. Under the multivariate null hypothesis $H_0 : F_1 = \dots = F_a$, as $a \rightarrow \infty$, n_i bounded,

$$\sqrt{a} \left(\frac{1}{N^2} (H_1 - G_1) - (\tilde{H}_1 - \tilde{G}_1) \right) \xrightarrow{p} 0.$$

Theorem 5. Under the marginal null hypothesis $\bar{H}_0 : F_1^{(k)} = \dots = F_a^{(k)}$, $k = 1, \dots, p$, as $a \rightarrow \infty$, n_i bounded,

$$\sqrt{a} \left(\frac{1}{N^2} (H_1 - G_2) - (\tilde{H}_1 - \tilde{G}_2) \right) \xrightarrow{p} 0.$$

Theorem 6. Under the marginal null hypothesis $\bar{H}_0 : F_1^{(k)} = \dots = F_a^{(k)}$, $k = 1, \dots, p$, as $a \rightarrow \infty$, n_i bounded,

$$\sqrt{a} \left(\frac{1}{N^2} (H_2 - G_3) - (\tilde{H}_2 - \tilde{G}_3) \right) \xrightarrow{p} 0.$$

In Theorems 4–6, we have shown that in order to obtain the asymptotic ($a \rightarrow \infty$) distributions of $\sqrt{a} \frac{1}{N^2} (H_1 - G_1)$, $\sqrt{a} \frac{1}{N^2} (H_1 - G_2)$, and $\sqrt{a} \frac{1}{N^2} (H_2 - G_3)$, it is sufficient to derive the asymptotic distribution of the corresponding expression $\sqrt{a} \frac{1}{N^2} (\tilde{H}_i - \tilde{G}_j)$. This is accomplished in the following theorems. More specifically, the next theorems provide the asymptotic distribution of the traces of the random matrices $(\tilde{H}_1 - \tilde{G}_1)A$, $(\tilde{H}_1 - \tilde{G}_2)A$, and $(\tilde{H}_2 - \tilde{G}_3)A$, where A is a fixed non-negative definite matrix. Different choices of A will result in the different multivariate tests under consideration (Lawley-Hotelling, Bartlett-Nanda-Pillai, Dempster-ANOVA-Type).

2.4 Asymptotic Distribution of Tests Based on H_1 and G_1

In this and the following sections, we derive concrete tests that are based on the three pairs of $p \times p$ -matrices (H_1, G_1) , (H_1, G_2) , and (H_2, G_3) . In each case, the asymptotic distribution is first derived in general theorems covering a class of tests, and then the results for individual multivariate tests follow in corollaries.

Theorem 7. Let A be an arbitrary fixed non-negative definite matrix, and assume that, as $a \rightarrow \infty$, n_i bounded,

$$\lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a n_i = \bar{n} \quad \text{and} \quad \lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \frac{1}{n_i} = \underline{n}.$$

Then, under the multivariate null hypothesis $H_0 : F_1 = \dots = F_a$, as $a \rightarrow \infty$, n_i bounded,

$$\sqrt{a} \left(\frac{1}{N^2} \text{tr}(H_1 - G_1)A \right) \xrightarrow{\mathcal{L}} N(0, \tau(A))$$

$$\text{where } \tau(A) = \frac{2\bar{n}}{\bar{n} - 1} \text{tr}(\Sigma_1 A)^2 + \frac{\bar{n}(\bar{n}\underline{n} - 1)}{(\bar{n} - 1)^2} (\mu_4(A) - 2\text{tr}(\Sigma_1 A)^2 - (\text{tr}\Sigma_1 A)^2)$$

$$\Sigma_1 = \text{Var}(\mathbf{Y}_{11}) \text{ and } \mu_4(A) = E[(\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})' A (\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})]^2.$$

In order to be able to apply the asymptotic result from Theorem 7, we need to be able to consistently estimate the variance $\tau(A)$. This is accomplished in the following Theorem 8.

Theorem 8. *Let the null hypothesis and the assumptions be as in Theorem 7, and let $\tau(A)$ be defined as in Theorem 7. Let*

$$\hat{\tau}(A) = \frac{2\bar{n}}{\bar{n}-1} \text{tr}(G_1 A)^2 + \frac{\bar{n}(\bar{n}\underline{n}-1)}{(\bar{n}-1)^2} (\hat{\mu}_4(A) - 2\text{tr}(G_1 A)^2 - (\text{tr}G_1 A)^2) ,$$

$$\text{where } \hat{\mu}_4(A) = \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \left[(\mathbf{R}_{ij} - \frac{N+1}{2} \mathbf{1})' A (\mathbf{R}_{ij} - \frac{N+1}{2} \mathbf{1}) \right]^2 , \text{ and } G_1 \text{ is defined in (3).}$$

Then, $\frac{1}{N^4} \hat{\tau}(A) \xrightarrow{p} \tau(A)$ as $a \rightarrow \infty$, n_i bounded.

The asymptotic distributions of the different multivariate nonparametric tests can now be obtained as corollaries of the previous theorems by appropriate choices of the fixed non-negative definite matrix A . In particular, the Dempster-ANOVA-Type statistic corresponds to $A = (\text{tr}\Sigma_1)^{-1} I_p$, while Lawley-Hotelling's trace and the Bartlett-Nanda-Pillai criterion correspond to $A = \Sigma_1^-$.

Corollary 1. *Let the null hypothesis and the assumptions be as in Theorem 7. Denote the Dempster-ANOVA-Type statistic by $T_D = \text{tr}(H_1)/\text{tr}(G_1)$. Then,*

$$\sqrt{a} (T_D - 1) \xrightarrow{\mathcal{L}} N(0, \tau)$$

$$\text{where } \tau = \frac{1}{(\text{tr}\Sigma_1)^2} \left(\frac{2\bar{n}}{\bar{n}-1} \text{tr}(\Sigma_1^2) + \frac{\bar{n}(\bar{n}\underline{n}-1)}{(\bar{n}-1)^2} (\mu_4 - 2\text{tr}(\Sigma_1^2) - (\text{tr}\Sigma_1)^2) \right) ,$$

$$\Sigma_1 = \text{var}(\mathbf{Y}_{11}) , \quad \text{and} \quad \mu_4 = E[(\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})'(\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})]^2 .$$

Furthermore, the estimator

$$\hat{\tau} = \frac{1}{(\text{tr}G_1)^2} \left(\frac{2\bar{n}}{\bar{n}-1} \text{tr}(G_1^2) + \frac{\bar{n}(\bar{n}\underline{n}-1)}{(\bar{n}-1)^2} (\hat{\mu}_4 - 2\text{tr}(G_1^2) - (\text{tr}G_1)^2) \right) ,$$

$$\text{where } \hat{\mu}_4 = \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \left[(\mathbf{R}_{ij} - \frac{N+1}{2} \mathbf{1})' (\mathbf{R}_{ij} - \frac{N+1}{2} \mathbf{1}) \right]^2 , \text{ and } G_1 \text{ is defined in (3),}$$

is consistent for τ in the sense that $\hat{\tau} \xrightarrow{p} \tau$ as $a \rightarrow \infty$, n_i bounded.

Corollary 2. *Let the null hypothesis and the assumptions be as in Theorem 7. Let r_1 be the rank of G_1 . Denote the Lawley-Hotelling-Type statistic by $T_{LH} = \frac{a-1}{N-a} \text{tr}(H_1 G_1^-)$. Then,*

$$\sqrt{a} \left(\left(\frac{N-a}{a-1} \right) T_{LH} - r_1 \right) \xrightarrow{\mathcal{L}} N(0, \tau) ,$$

$$\text{where } \tau = \frac{2\bar{n}\rho_1}{\bar{n}-1} + \frac{\bar{n}(\bar{n}\underline{n}-1)}{(\bar{n}-1)^2} (\mu_4 - 2\rho_1 - \rho_1^2) ,$$

$$\rho_1 = \text{rank}(\Sigma_1) , \quad \Sigma_1 = \text{var}(\mathbf{Y}_{11}) , \quad \text{and} \quad \mu_4 = E[(\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})' \Sigma_1^- (\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})]^2 .$$

Furthermore, as $a \rightarrow \infty$, n_i bounded, the following estimator $\hat{\tau}$ is consistent for τ .

$$\hat{\tau} = \frac{2\bar{n}r_1}{\bar{n} - 1} + \frac{\bar{n}(\bar{n}\underline{n} - 1)}{(\bar{n} - 1)^2}(\hat{\mu}_4 - 2r_1 - r_1^2),$$

$$\text{where } \hat{\mu}_4 = \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \left[\left(\mathbf{R}_{ij} - \frac{N+1}{2} \mathbf{1} \right)' \left(\frac{1}{N^2} G_1 \right)^- \left(\mathbf{R}_{ij} - \frac{N+1}{2} \mathbf{1} \right) \right]^2, \text{ and } G_1 \text{ is defined in (3).}$$

Corollary 3. Let the null hypothesis and the assumptions be as in Theorem 7. Let r_2 be the rank of $(a-1)H_1 + (N-a)G_1$. Denote the Bartlett-Nanda-Pillai-Type statistic by $T_{BNP} = (a-1)\text{tr} \left(H_1 \left((a-1)H_1 + (N-a)G_1 \right)^- \right)$. Then,

$$\sqrt{a} \left(\frac{N-1}{N-a} \right) \left(\left(\frac{N-1}{a-1} \right) T_{BNP} - r_2 \right) \xrightarrow{\mathcal{L}} N(0, \tau),$$

$$\text{where } \tau = \frac{2\bar{n}\rho_1}{\bar{n} - 1} + \frac{\bar{n}(\bar{n}\underline{n} - 1)}{(\bar{n} - 1)^2}(\mu_4 - 2\rho_1 - \rho_1^2),$$

$$\rho_1 = \text{rank}(\Sigma_1), \Sigma_1 = \text{var}(\mathbf{Y}_{11}), \text{ and } \mu_4 = E[(\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})' \Sigma_1^- (\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})]^2.$$

Notice that the expression for τ in Corollary 3 is the same as in Corollary 2. Hence, the estimator $\hat{\tau}$ defined in Corollary 2 works for the Bartlett-Nanda-Pillai-Type statistic, as well.

2.5 Asymptotic Distribution of Tests Based on H_1 and G_2

In this section we prove asymptotic ($a \rightarrow \infty$) normality results for Dempster's ANOVA-Type and the Lawley-Hotelling statistics. It is not clear how to get a weighted mixture of H_1 and G_2 to obtain a consistent and asymptotically unbiased estimator of $E_{H_0}(H_1)$. Therefore, there is no straightforward way to define a Bartlett-Nanda-Pillai type statistic based on H_1 and G_2 .

Theorem 9. Assume that the following limit exists.

$$\lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2 \text{tr}(\Sigma_i A)^2 = \tau_2(A).$$

Then, under the (marginal) null hypothesis $\bar{H}_0 : F_1^{(k)} = F_2^{(k)} = \dots = F_a^{(k)}$ for all $k = 1, \dots, p$, as $a \rightarrow \infty$, n_i fixed and any $A_{p \times p} \geq 0$ fixed,

$$\sqrt{a} \cdot \text{tr}(H_1 - G_2)A \xrightarrow{\mathcal{L}} N(0, \tau_2(A)).$$

Corollary 4. Assume that the following limit exists.

$$\frac{1}{(\text{tr}\Sigma)^2} \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2 \text{tr}(\Sigma_i)^2 = \tau_2,$$

where Σ is as defined in Theorem 3. Denote the Dempster–ANOVA-Type statistic based on H_1 and G_2 by $T_D^{(2)} = \text{tr}(H_1)/\text{tr}(G_2)$. Then, under the marginal null hypothesis as stated in Theorem 9,

$$\sqrt{a} \left(T_D^{(2)} - 1 \right) \xrightarrow{\mathcal{L}} N(0, \tau_2).$$

Under the multivariate null hypothesis, as stated in Theorem 7, τ_2 simplifies to

$$\tau_2 = \frac{\text{tr}(\Sigma^2)}{(\text{tr} \Sigma)^2} \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2,$$

which can be consistently estimated by

$$\hat{\tau}_2 = \frac{\text{tr}(G_2^2)}{(\text{tr} G_2)^2} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2.$$

Corollary 5. Assume that the following limit exists.

$$\lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2 \text{tr}(\Sigma_i \Sigma^{-1}) = \tau_2,$$

where Σ is as defined in Theorem 3. Let $r_1^{(2)}$ be the rank of G_2 . Denote the Lawley-Hotelling-Type statistic based on H_1 and G_2 by $T_{LH}^{(2)} = \text{tr}(H_1 G_2^{-1})$. Then, under the marginal null hypothesis as stated in Theorem 9,

$$\sqrt{a} \left(T_{LH}^{(2)} - r_1^{(2)} \right) \xrightarrow{\mathcal{L}} N(0, \tau_2).$$

Under the multivariate null hypothesis, as stated in Theorem 7, τ_2 simplifies to

$$\tau_2 = \text{rank}(\Sigma) \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2,$$

which can be consistently estimated by

$$\hat{\tau}_2 = \text{rank}(G_2) \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N} \right)^2.$$

It is clear from Theorem 5 that the asymptotic variance of the test statistics depends on the variances and covariances of the ART. Since the columns of the ART matrix are independent, the results derived based on the ART matrix can be applied directly to obtain the null distribution of the same test statistics based on the original observations, whose columns are also independent. However, the entries of the ART matrix are uniformly bounded, whereas it will be necessary to assume existence of fourth moments for the parent populations when basing the test statistics on the original observations.

Unlike the tests based on H_1 and G_1 , the asymptotic null distributions presented in Corollaries 4 and 5 do not depend on the fourth order moments of the asymptotic transforms. Hence, when the statistics based on these corollaries are applied to the original observations, the sizes of the test will be asymptotically invariant with regard to the distribution of the populations from which the samples are coming.

2.6 Asymptotic Distribution of Tests Based on H_2 and G_3

Here, we calculate the asymptotic ($a \rightarrow \infty$) distributions of Dempster's ANOVA-Type and the Lawley-Hotelling statistics, based on H_2 and G_3 . As in the previous section, there is again no straightforward way to define a Bartlett-Nanda-Pillai type statistic based on H_2 and G_3 .

Theorem 10. *Assume that the following limit exists.*

$$\lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)} \text{tr}(\Sigma_i A)^2 = \tau_3(A).$$

Then, under the (marginal) null hypothesis $\bar{H}_0 : F_1^{(k)} = F_2^{(k)} = \dots = F_a^{(k)}$ for all $k = 1, \dots, p$, as $a \rightarrow \infty$, n_i fixed and any $A_{p \times p} \geq 0$ fixed,

$$\sqrt{a} \cdot \text{tr}(H_2 - G_3)A \xrightarrow{\mathcal{L}} N(0, \tau_3(A)).$$

Corollary 6. *Assume that the following limit exists.*

$$\frac{1}{(\text{tr} \check{\Sigma})^2} \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)} \text{tr}(\Sigma_i)^2 = \tau_3$$

where $\check{\Sigma}$ is as defined in Theorem 3. Denote the Dempster-ANOVA-Type statistic based on H_2 and G_3 by $T_D^{(3)} = \text{tr}(H_2)/\text{tr}(G_3)$. Then, under the marginal null hypothesis as stated in Theorem 10,

$$\sqrt{a} \left(T_D^{(3)} - 1 \right) \xrightarrow{\mathcal{L}} N(0, \tau_3).$$

Under the multivariate null hypothesis, as stated in Theorem 7, τ_3 simplifies to

$$\tau_3 = \frac{\text{tr} \Sigma^2}{(\text{tr} \check{\Sigma})^2} \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)},$$

which can be consistently estimated by

$$\hat{\tau}_3 = \frac{\text{tr} G_2^2}{(\text{tr} G_3)^2} \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)}.$$

Corollary 7. *Assume that the following limit exists.*

$$\lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)} \text{tr}(\Sigma_i \check{\Sigma}^-)^2 = \tau_3$$

where $\check{\Sigma}$ is as defined in Theorem 3. Let $r_1^{(3)}$ be the rank of G_3 . Denote the Lawley-Hotelling-Type statistic based on H_2 and G_3 by $T_{LH}^{(3)} = \text{tr}(H_2 G_3^-)$. Then, under the marginal null hypothesis as stated in Theorem 10,

$$\sqrt{a} \left(T_{LH}^{(3)} - r_1^{(3)} \right) \xrightarrow{\mathcal{L}} N(0, \tau_3).$$

Under the multivariate null hypothesis, as stated in Theorem 7, τ_3 simplifies to

$$\tau_3 = \text{tr} (\Sigma \check{\Sigma}^-)^2 \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)},$$

which can be consistently estimated by

$$\hat{\tau}_3 = \text{tr} (G_2 G_3^-)^2 \frac{2}{a} \sum_{i=1}^a \frac{1}{n_i(n_i - 1)}.$$

The asymptotic distributions of tests based on H_2 and G_3 also do not depend on the fourth order moments of the asymptotic transforms. Therefore, as in the previous section, the corresponding procedures applied to the original observations will be asymptotically invariant with regard to the population distributions.

3 SIMULATION STUDY

We investigated the finite sample performance of the proposed nonparametric procedures as well as its parametric counterparts through computer simulations using SAS IML (SAS 9.1). Power functions were plotted with R [25]. For the simulations, we assumed a one-way layout, and we considered the following nonparametric (rank-based) multivariate tests. Dempster–ANOVA type and Lawley-Hotelling type statistic based on (H_1, G_1) , (H_1, G_2) , and (H_2, G_3) , and Bartlett-Nanda-Pillai type statistic based on (H_1, G_1) . As parametric competitors, we used the corresponding statistics based on the original observations. Note that the parametric tests based on (H_1, G_2) and (H_2, G_3) are newly introduced in this manuscript.

The number of simulations was always 10,000. Sample sizes per group were chosen in the following ways. (1) Half of the samples were of size $n_i = 4$, the other half had $n_i = 6$. (2) Half of the samples were of size $n_i = 4$, the other half had $n_i = 8$. (3) Ten percent of the samples were of size $n_i = 4$, the remaining 90 percent were of size $n_i = 6$ or (4) $n_i = 8$. We only report the results from setting (1) in the tables below, because the other simulations had very similar outcomes.

First, we simulated the actual α -level of all seven test statistics in a one-way layout under the null hypothesis that the multivariate distributions are the same for each factor level. We set the number of factor levels to $a = 10, 20, 50, 100$, and 200 , while the number of replications per factor level followed one of the patterns described above, and the number of variables was $p = 2$ (Table 1) and $p = 4$ (Table 2).

We have used different underlying population distributions, namely multivariate normal with positive, negative, and zero correlation between the variables, as well as with and without contamination through 10% outliers (coming from a $N(10,1)$ distribution), and multivariate distributions based on Student's t with different degrees of freedom and different correlation. The results for some selected configurations are summarized in Tables 1 and 2 below.

Next, we have investigated the power of the proposed tests under location alternatives. We are conducting the power comparison based on a fairly small value of a . We selected those four tests whose simulated α -levels were in general closest to the nominal 5%. These are the Dempster–ANOVA type statistic based on (H_1, G_1) and (H_2, G_3) , as well as Lawley-Hotelling and Bartlett-Nanda-Pillai type statistics based on (H_1, G_1) . However, for large a , the other tests will arguably have comparable powers. For half of the factor levels, the mean is shifted by x where $x \in [0, 2]$ for each variable. In general, alternatives for power simulations can be chosen in many ways. One of the intentions of the simulation study is to compare the nonparametric tests to each other and also to their parametric competitors. The parametric tests are naturally defined in the framework of a location model. Therefore, we have also chosen a location model for this simulation. For the power simulations, we have set the number of variables to $p = 2$. The number of factor levels is $a = 20$, and the number of replications per factor level is as described in setting (1) above. Plots of different simulated power functions are given in Figures 1 and 2.

The α -level simulations in Tables 1 and 2 suggest that the Lawley-Hotelling type statistics based on (H_1, G_2) and (H_2, G_3) should not be recommended unless a is very large since they are by far exceeding the nominal α . All tests are slightly liberal for small to moderate a . Most of the tests based on (H_1, G_1) are within two percentage points from the nominal α of 5% when the number of levels is $a = 20$ or larger, and within one percentage point when $a = 50$. The best convergence results are achieved by the Bartlett-Nanda-Pillai statistic. In general, contamination or correlation do not seem to affect the simulated α -levels. For $p = 4$, convergence to the nominal α is usually faster than for $p = 2$.

As expected, nonparametric and parametric tests perform very similarly under null hypothesis. Under a simulated location alternative, they also perform similarly when the normal model assumptions are met. Figure 1 intends to demonstrate the dramatic power differences between parametric and nonparametric tests when the data contains about 10% outliers or is from a more heavy-tailed distribution such as Student's t with $df = 3$. From this figure, it can also be seen that in general the tests that exceed the nominal α -level by the farthest are also the ones that keep the highest power in general. That is, higher power is in general achieved through a higher type I error rate.

Generally, the nonparametric tests based on (H_1, G_1) are recommended when the simulated type of alternative is suspected. The Dempster–ANOVA type statistic based on (H_2, G_3) has only slightly lower power than the one based on (H_1, G_1) , but it has the practical advantage that it can easily be calculated using SAS standard procedures (see the data example below).

In case of positive correlation between the variables, the Dempster–ANOVA type statistic has higher power than the other two types, whereas for negatively correlated variables, the Lawley-Hotelling and Bartlett-Nanda-Pillai statistics achieve higher simulated power. This is shown in Figure 2. In Figure 1, the variables are not correlated, and the Dempster–ANOVA type statistic fares best among those tests that also keep the nominal level well.

Underlying Distribution	Test Statistic	Number of Variables $p = 2$				
		$a = 10$	$a = 20$	$a = 50$	$a = 100$	$a = 200$
multivariate normal with correlation=0	ANOVA type (H_1, G_1)	7.9 (7.8)	7.2 (7.2)	6.4 (6.4)	6.2 (6.4)	5.5 (5.7)
	ANOVA type (H_1, G_2)	9.6 (10.1)	8.3 (8.5)	6.5 (6.6)	6.4 (6.5)	5.7 (5.8)
	ANOVA type (H_2, G_3)	8.1 (8.1)	7.3 (7.2)	6.2 (6.2)	6.1 (6.2)	5.8 (5.8)
	LH type (H_1, G_1)	6.8 (6.5)	6.3 (6.3)	6.0 (5.8)	5.6 (5.9)	5.3 (5.3)
	LH type (H_1, G_2)	14.1 (13.9)	11.0 (10.8)	7.8 (7.7)	7.1 (7.1)	6.1 (6.2)
	LH type (H_2, G_3)	12.3 (11.9)	9.9 (9.5)	7.6 (7.3)	6.8 (7.0)	6.2 (6.2)
	BNP type (H_1, G_1)	6.2 (5.8)	6.0 (5.9)	5.6 (5.5)	5.4 (5.8)	5.2 (5.2)
multivariate normal with correlation = 0.5 and 10% outliers	ANOVA type (H_1, G_1)	8.9 (7.6)	7.3 (7.2)	7.2 (6.8)	6.4 (6.1)	5.5 (5.6)
	ANOVA type (H_1, G_2)	10.3 (9.1)	8.1 (8.2)	7.1 (7.2)	6.5 (6.3)	5.6 (5.8)
	ANOVA type (H_2, G_3)	8.5 (7.3)	7.4 (7.2)	6.7 (6.8)	6.4 (6.3)	6.0 (5.7)
	LH type (H_1, G_1)	7.2 (6.4)	6.0 (6.5)	6.2 (6.2)	5.8 (5.6)	5.2 (5.3)
	LH type (H_1, G_2)	15.0 (12.0)	10.3 (10.5)	7.9 (8.6)	7.4 (6.9)	6.0 (6.4)
	LH type (H_2, G_3)	12.9 (9.7)	9.4 (9.2)	7.6 (8.2)	6.9 (6.8)	6.2 (6.2)
	BNP type (H_1, G_1)	6.5 (5.7)	5.7 (5.9)	6.0 (5.9)	5.6 (5.4)	5.1 (5.1)
multivariate normal with correlation = -0.5 and 10% outliers	ANOVA type (H_1, G_1)	8.3 (7.3)	7.8 (7.0)	6.4 (6.5)	6.1 (6.1)	5.9 (5.8)
	ANOVA type (H_1, G_2)	10.5 (9.0)	8.3 (8.5)	6.9 (7.1)	6.4 (6.4)	6.1 (6.2)
	ANOVA type (H_2, G_3)	8.7 (7.2)	7.5 (7.4)	6.6 (6.6)	6.0 (6.4)	6.1 (6.2)
	LH type (H_1, G_1)	6.8 (6.1)	6.8 (6.4)	5.6 (5.9)	5.8 (5.9)	5.5 (5.5)
	LH type (H_1, G_2)	14.4 (12.4)	10.5 (10.9)	7.7 (8.3)	7.3 (7.1)	6.2 (6.7)
	LH type (H_2, G_3)	12.5 (10.2)	9.8 (9.7)	7.4 (7.9)	7.1 (7.0)	6.2 (6.8)
	BNP type (H_1, G_1)	6.3 (5.3)	6.3 (5.8)	5.4 (5.5)	5.6 (5.6)	5.4 (5.4)

Table 1: Simulated α -levels [in percent] for the proposed nonparametric multivariate tests (in parentheses: their respective parametric counterparts) of the Dempster-ANOVA, Lawley-Hotelling, and Bartlett-Nanda-Pillai type. Nominal α is 5%. Number of simulations is 10,000 (standard error 0.43). Number of variables is $p = 2$. Varying numbers of factor levels between $a = 10$ and $a = 200$. In each case, half of the samples are of size $n_i = 4$, the other half is of size $n_i = 6$.

Underlying Distribution	Test Statistic	Number of Variables $p = 4$				
		$a = 10$	$a = 20$	$a = 50$	$a = 100$	$a = 200$
multivariate normal with correlation=0	ANOVA type (H_1, G_1)	7.0 (6.7)	6.9 (6.6)	5.7 (5.9)	5.5 (5.7)	5.3 (5.5)
	ANOVA type (H_1, G_2)	9.0 (8.6)	7.8 (7.4)	6.1 (6.3)	5.6 (6.0)	5.5 (5.7)
	ANOVA type (H_2, G_3)	7.1 (6.7)	6.9 (6.5)	5.8 (6.0)	5.7 (5.6)	5.5 (5.4)
	LH type (H_1, G_1)	10.2 (9.8)	8.6 (8.3)	6.3 (6.8)	5.8 (6.0)	5.6 (5.7)
	LH type (H_1, G_2)	22.9 (22.7)	15.5 (15.2)	10.4 (10.4)	7.7 (8.2)	7.0 (7.1)
	LH type (H_2, G_3)	20.9 (20.5)	14.8 (14.4)	10.3 (10.3)	7.9 (8.1)	7.1 (7.2)
	BNP type (H_1, G_1)	5.7 (5.5)	5.9 (5.8)	4.9 (5.4)	5.1 (5.1)	5.1 (5.2)
multivariate normal with correlation = 0.5 and 10% outliers	ANOVA type (H_1, G_1)	7.9 (6.7)	6.7 (6.5)	6.3 (5.6)	5.9 (5.7)	5.5 (5.5)
	ANOVA type (H_1, G_2)	9.7 (8.4)	7.8 (7.5)	6.7 (6.4)	6.1 (6.1)	5.7 (5.6)
	ANOVA type (H_2, G_3)	7.7 (6.7)	7.1 (6.5)	6.3 (6.0)	6.0 (5.8)	5.7 (5.7)
	LH type (H_1, G_1)	10.5 (9.3)	7.9 (8.1)	6.3 (6.4)	5.7 (6.2)	5.7 (5.8)
	LH type (H_1, G_2)	23.0 (20.8)	15.1 (14.7)	9.8 (10.3)	7.8 (8.3)	7.2 (7.3)
	LH type (H_2, G_3)	20.6 (17.6)	14.2 (13.7)	9.6 (9.9)	7.7 (8.2)	7.4 (7.6)
	BNP type (H_1, G_1)	5.7 (4.7)	5.4 (5.3)	5.0 (5.0)	5.0 (5.2)	5.2 (5.3)
multivariate normal with correlation = -0.3	ANOVA type (H_1, G_1)	7.2 (7.1)	6.8 (6.6)	5.7 (5.8)	5.5 (5.6)	5.7 (5.7)
	ANOVA type (H_1, G_2)	9.2 (9.0)	7.7 (7.5)	6.3 (6.2)	5.6 (5.5)	6.0 (5.8)
	ANOVA type (H_2, G_3)	7.4 (7.5)	6.8 (6.6)	5.9 (5.8)	5.3 (5.4)	5.9 (5.7)
	LH type (H_1, G_1)	9.8 (9.7)	8.4 (7.9)	6.5 (6.2)	6.1 (6.0)	5.7 (5.7)
	LH type (H_1, G_2)	22.6 (22.8)	15.7 (15.1)	9.8 (9.8)	8.1 (8.1)	7.2 (7.1)
	LH type (H_2, G_3)	20.7 (20.5)	14.8 (14.4)	10.0 (9.8)	8.2 (8.1)	7.2 (7.1)
	BNP type (H_1, G_1)	5.1 (5.3)	5.6 (5.6)	5.0 (5.0)	5.2 (5.3)	5.1 (5.0)

Table 2: Simulated α -levels [in percent] for the proposed nonparametric multivariate tests (in parentheses: their respective parametric counterparts). Nominal α is 5%. Number of simulations is 10,000 (standard error 0.43). Number of variables is $p = 4$. Varying numbers of factor levels between $a = 10$ and $a = 200$. In each case, half of the samples are of size $n_i = 4$, the other half is of size $n_i = 6$.

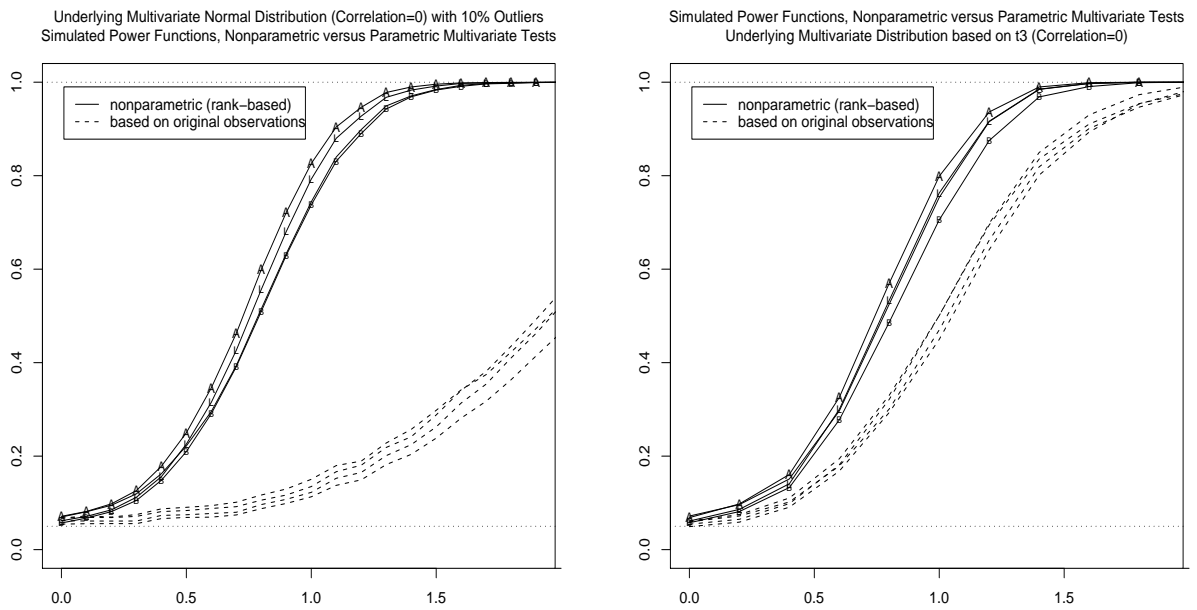


Figure 1: Simulated power of proposed nonparametric (on top) and corresponding parametric versions of four of the multivariate tests under investigation. $a = 20$ levels, $p = 2$ variables. Half of the samples are of size $n_i = 4$, the other half is of size $n_i = 6$. Underlying distribution is bivariate normal with correlation=0, and 10% outliers (left) or bivariate Student's t with $df = 3$, correlation=0, and no outliers (right). The letters "A", "L", and "B" denote ANOVA-Type, Lawley-Hotelling, and Bartlett-Nanda-Pillai, respectively, all based on (H_1, G_1) .

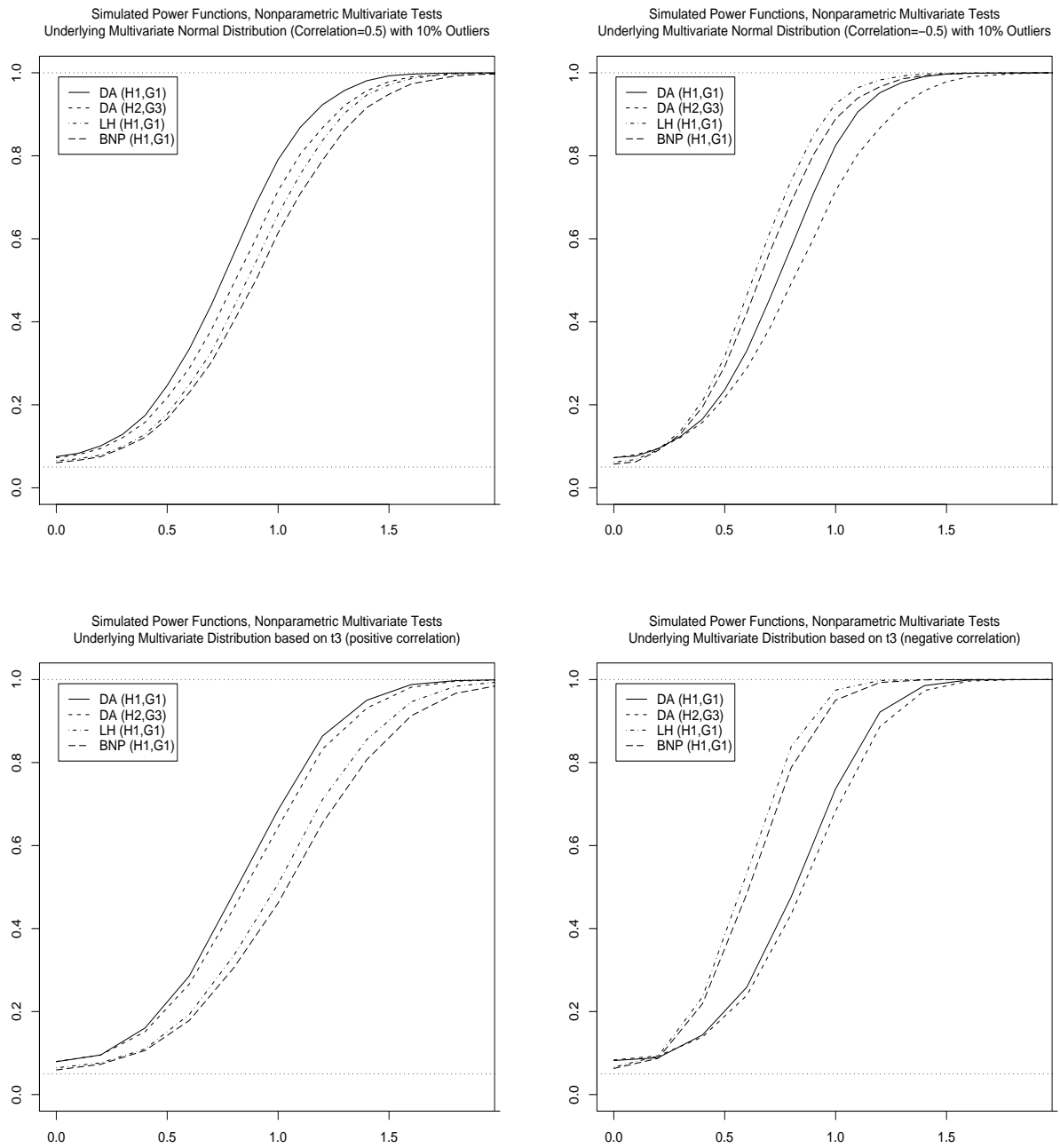


Figure 2: Simulated power of three nonparametric versions of the multivariate tests under investigation. $\alpha = 20$ levels, $p = 2$ variables. Half of the samples are of size $n_i = 4$, the other half is of size $n_i = 6$. First row: Underlying distribution is bivariate normal with positive correlation (0.5, left) and negative correlation (-0.5, right), and 10% outliers. Second row: Underlying bivariate distribution is based on Student's t with $df = 3$, positive (left) and negative correlation (right), and no outliers.

4 APPLICATION

We applied the methods derived in the present manuscript to a plant pathology data set. Chatfield et al. [4] evaluated $a = 63$ varieties of crabapples for disease resistance against apple scab at $p = 4$ times during the growing season, with $n_i = 3$ to 5 replicates of each variety. Apple scab is a major fungal disease problem that can severely affect marketability of crabapples. The best method for control is through the use of resistant crabapple selections [4]. One of the goals in the described trial is to find out whether the 63 varieties differ significantly with regard to their disease resistance. The response variable is ordinal, each crabapple tree was rated on a scale from zero to five. Therefore, parametric normal theory techniques are not appropriate for the analysis of this type of data, and nonparametric methods should be used.

We tested the multivariate null hypothesis that there is no difference in disease resistance between the 63 varieties. According to the simulation results reported in the last section, the nonparametric version of the Dempster–ANOVA type test is recommended for this type of data since we would expect positive correlation between the measurements at different time points. This was confirmed by a simulation of α -level and power for a design imitating the crabapple experiment as closely as possible (simulating high correlation of 0.9 between the variables, and discretizing the response to a six-point ordinal scale). However, it actually turns out that for this data set, each of the parametric and nonparametric versions of the multivariate tests under consideration led to the same conclusion of high significance ($p < 0.0001$). That is, the different varieties of crabapples show significantly different disease resistance. As a result, it is indeed possible to reduce the impact of apple scab through the choice of more resistant varieties.

We have performed the analysis using code written in SAS IML that we can email on request. It is possible to roughly approximate some of the test statistics derived in this manuscript using SAS standard procedures, though. For example, the terms T_{LH} in Corollary 2 as well as T_{BNP} in Corollary 3 can be obtained using the `manova` option in `Proc Glim`, after individually ranking each of the variables. After `Proc Rank` outputs the ranks of the four scores into the variables `r1–r4`, the appropriate SAS code for this data set is then:

```
proc glm data=a; class variety; model r1-r4=variety;
manova h=variety; run;
```

The terms T_{LH} and T_{BNP} appear in the SAS output as “Hotelling-Lawley Trace” and “Pillai’s Trace”, respectively. As Dr. Larry Madden pointed out to us, the term $T_D^{(3)}$ in Corollary 6 can also be reproduced, using `Proc Mixed` along with the `anovaf` option on the ranked data. To this end, the separate ranks of all four variables have to be stacked into one new rank variable (here called `rr`), and two more variables representing “time” and individual “subject” have to be created. Then, the effect of variety is tested as a simple factor effect of “variety|time”, using the following code:

```
proc mixed data=a2 anovaf method=mivque0;
class time variety subject;
model rr = time time*variety / chisq noint;
repeated / group=variety sub=subject type=un; run;
```

However, calculation of the standardized test statistics that can be compared to quantiles from a standard normal distribution still involves some programming. When the design is not too far from being balanced, a rough approximation is possible by substituting the asymptotic variance τ in Corollaries 2 and 3 by $2\bar{n}r_1/(\bar{n} - 1)$.

A Some Useful Results

In this section, we restate some results from Bathke and Harrar [6] (wherein the proof can be found). These results are used in the present manuscript. They facilitate determining asymptotic equivalence of rank transforms and asymptotic rank transforms, as well as calculating the variance of the limiting distribution..

Theorem A.1. *Let $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N)$ be a matrix that consists of independent random vectors $\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(p)})'$ with multivariate distribution $\mathbf{X}_i \sim F_i, i = 1, \dots, N$. Let \mathbf{Y} and $\hat{\mathbf{Y}}$ be the corresponding matrices of same dimension whose components are the asymptotic rank transforms and rank transforms defined in Section 2. Let $C = (c_{ik})_{1 \leq i \leq N; 1 \leq k \leq N}$ be a symmetric matrix, and let*

$$\begin{aligned}\Sigma_C &= \sum_{i=1}^N \sum_{k=1}^N |c_{ik}|, \\ S_C &= \sum_{i=1}^N \sum_{k=1}^N \sum_{k'=1}^N |c_{ik}c_{ik'}| + \sum_{i \neq k} (|c_{ii}c_{kk}| + c_{ik}^2).\end{aligned}$$

Furthermore, let $T_N = \hat{\mathbf{Y}}_\mu \mathbf{C} \hat{\mathbf{Y}}_\mu'$ and $V_N = \mathbf{Y}_\mu \mathbf{C} \mathbf{Y}_\mu'$ be two $(p \times p)$ -matrices of quadratic forms generated by the matrix C .

Then,

$$T_N - V_N = O_P(\Sigma_C^2/N^2) + O_P(S_C/N).$$

Corollary A.1. *Suppose C is such that all its entries are nonnegative. Then,*

$$\Sigma_C = \mathbf{1}'_N \mathbf{C} \mathbf{1}_N.$$

If in addition all the diagonal entries of C are equal, then

$$S_C = \mathbf{1}'_N \mathbf{C}^2 \mathbf{1}_N + (1 - \frac{2}{N})(\text{tr} \mathbf{C})^2 + \text{tr} \mathbf{C}^2.$$

Lemma A.1. *Suppose $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ is a $p \times n$ random matrix whose columns $\mathbf{Y}_i, i = 1, \dots, n$, are independently distributed with mean $\mathbf{0}$ and covariance Σ_i . Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times n$ symmetric matrices. Then,*

$$\begin{aligned}\text{Cov}(\text{vec}(\mathbf{Y} \mathbf{A} \mathbf{Y}'), \text{vec}(\mathbf{Y} \mathbf{B} \mathbf{Y}')) &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij} (I_{p^2} + K_{p,p})(\Sigma_i \otimes \Sigma_j) \\ &\quad + \sum_{i=1}^n a_{ii} b_{ii} K_4(\mathbf{Y}_i)\end{aligned}$$

where $K_4(\mathbf{Y}_i) = \mathbb{E}(\text{vec}(\mathbf{Y}_i \mathbf{Y}_i') \text{vec}(\mathbf{Y}_i \mathbf{Y}_i')') - (I_{p^2} + K_{p,p})(\Sigma_i \otimes \Sigma_i) - \text{vec}(\Sigma_i) \text{vec}(\Sigma_i)'$.

B Proofs

Proof of Theorem 1 Note that under the null hypothesis, $\Sigma_i \equiv \Sigma$, $i = 1, \dots, a$. Define $C = \frac{1}{N-a} \bigoplus_{i=1}^a P_{n_i}$. Then, $\frac{1}{N^2} G_1 = \hat{\mathbf{Y}} C \hat{\mathbf{Y}}'$. Theorem A.1 and Corollary A.1 in Appendix A yield $\hat{\mathbf{Y}} C \hat{\mathbf{Y}}' - \mathbf{Y} C \mathbf{Y}' = O_p(\Sigma_C^2/N^2) + O_p(S_C/N)$ where $\Sigma_C = \mathbf{1}'_N C_* \mathbf{1}_N$ and $S_C = \mathbf{1}'_N C_*^2 \mathbf{1}_N + (1 - \frac{2}{N})(\text{tr } C_*)^2 + \text{tr } C_*^2$. Here, $C_* = \frac{1}{N-a} \bigoplus_{i=1}^a \left(\frac{1}{n_i} J_{n_i} + (1 - \frac{2}{n_i}) I_{n_i} \right)$ is the matrix whose elements are the absolute values of the elements of C . It immediately follows that $\Sigma_C = 2$ and $S_C = O(1)$. Therefore, $\hat{\mathbf{Y}} C \hat{\mathbf{Y}}' - \mathbf{Y} C \mathbf{Y}' = O_p(\frac{1}{a}) \xrightarrow{p} 0$ as $a \rightarrow \infty$. Now, consider an arbitrary diagonal element of $\mathbf{Y} C \mathbf{Y}'$, $\tilde{\sigma}_k^2 = \frac{1}{N-a} \sum_{i=1}^a \sum_{j=1}^{n_i} (Y_{ij}^{(k)} - \bar{Y}_{i.}^{(k)})^2 = \frac{1}{N-a} \sum_{i=1}^a (n_i - 1) \tilde{\sigma}_{k:i}^2$ where $\mathbb{E}(\tilde{\sigma}_{k:i}^2) = \sigma_k^2$ and $\text{Var}(\tilde{\sigma}_{k:i}^2) < M$ since $\mathbf{Y}_{ij}^{(k)}$ is a bounded random variable. Thus, it follows that $\tilde{\sigma}_k^2$ converges almost surely to σ_k^2 . In the same manner, the off-diagonal elements $\tilde{\sigma}_{kl}$ converge almost surely to σ_{kl} , which finishes the proof. \square

Proof of Theorem 2 Notice that under the null hypothesis $\mu_{ij} = \mu$. Then, observe that

$$\begin{aligned} & \frac{1}{N^2(N-1)} ((a-1)H_1 + (N-a)G_1) \\ &= \frac{1}{N-1} \sum_{i=1}^a \sum_{j=1}^{n_i} (\hat{\mathbf{Y}}_{ij} - \mu)(\hat{\mathbf{Y}}_{ij} - \mu)' - \frac{N}{N-1} (\hat{\mathbf{Y}}_{..} - \mu)(\hat{\mathbf{Y}}_{..} - \mu)' \\ &= \hat{\mathbf{Y}}_{\mu} \left(\frac{1}{N-1} I_N \right) \hat{\mathbf{Y}}_{\mu}' - \hat{\mathbf{Y}}_{\mu} \left(\frac{1}{N(N-1)} J_N \right) \hat{\mathbf{Y}}_{\mu}' \\ &= \hat{\mathbf{Y}}_{\mu} C_1 \hat{\mathbf{Y}}_{\mu}' - \hat{\mathbf{Y}}_{\mu} C_2 \hat{\mathbf{Y}}_{\mu}' \end{aligned}$$

where

$$C_1 = \frac{1}{N-1} I_N \quad \text{and} \quad C_2 = \frac{1}{N(N-1)} J_N.$$

It can be verified that

$$\begin{aligned} \Sigma_{C_1} &= \frac{N}{N-1} = O(1), \quad S_{C_1} = \frac{N}{(N-1)^2} + \left(1 - \frac{2}{N}\right) \left(\frac{N}{N-1}\right)^2 + \frac{N}{(N-1)^2} = O(1) \\ \Sigma_{C_2} &= \frac{N^2}{N(N-1)} = O(1), \quad \text{and} \quad S_{C_2} = \frac{N}{(N-1)^2} + \left(1 - \frac{2}{N}\right) \left(\frac{1}{N-1}\right)^2 + \frac{1}{(N-1)^2} = O(1) \end{aligned}$$

as $N \rightarrow \infty$. Therefore, by Theorem A.1, $\hat{\mathbf{Y}}_{\mu} C_1 \hat{\mathbf{Y}}_{\mu}' - \mathbf{Y}_{\mu} C_1 \mathbf{Y}_{\mu}' = o_p(1)$ and $\hat{\mathbf{Y}}_{\mu} C_2 \hat{\mathbf{Y}}_{\mu}' - \mathbf{Y}_{\mu} C_2 \mathbf{Y}_{\mu}' = o_p(1)$. Moreover, (9) implies that $\mathbf{Y}_{\mu} C_2 \mathbf{Y}_{\mu}' = o_p(1)$. Thus,

$$\frac{1}{N^2(N-1)} ((a-1)H_1 + (N-a)G_1) - \mathbf{Y}_{\mu} C_1 \mathbf{Y}_{\mu}' = o_p(1).$$

Now,

$$\mathbf{Y}_\mu C_1 \mathbf{Y}'_\mu = \frac{1}{N-1} \sum_{i=1}^a \sum_{j=1}^{n_i} (\mathbf{Y}_{ij} - \boldsymbol{\mu})(\mathbf{Y}_{ij} - \boldsymbol{\mu})' \xrightarrow{p} E(\mathbf{Y}_{11} - \boldsymbol{\mu})(\mathbf{Y}_{11} - \boldsymbol{\mu})' = \Sigma_1$$

by the SLLN. \square

Proof of Theorem 3 We prove the result for G_2 . The proof for G_3 follows along the same lines.

Define $C = \frac{1}{a-1} \bigoplus_{i=1}^a (1 - \frac{n_i}{N}) \frac{1}{n_i-1} P_{n_i}$. Then, $\frac{1}{N^2} G_2 = \hat{\mathbf{Y}} C \hat{\mathbf{Y}}'$. Using the techniques applied in the proof of Theorem 1, we obtain $\hat{\mathbf{Y}} C \hat{\mathbf{Y}}' - \mathbf{Y} C \mathbf{Y}' \xrightarrow{p} 0$ as $a \rightarrow \infty$. Furthermore, the diagonal elements of $\mathbf{Y} C \mathbf{Y}'$ are of the form $\tilde{\sigma}_k^2 = \frac{1}{a-1} \sum_{i=1}^a (1 - \frac{n_i}{N}) \tilde{\sigma}_{k:i}^2$ where $\tilde{\sigma}_{k:i}^2 = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (Y_{ij}^{(k)} - \bar{Y}_i^{(k)})^2$, and under the marginal hypothesis, $E(\tilde{\sigma}_{k:i}^2) = \sigma_k^2$ and $\text{Var}(\tilde{\sigma}_{k:i}^2) < M$. Therefore, for bounded n_i , $\tilde{\sigma}_k^2 = \frac{1}{a} \sum_{i=1}^a \tilde{\sigma}_{k:i}^2 + o_p(1) \xrightarrow{p} \sigma_k^2$. However, note that for the off-diagonal elements $\tilde{\sigma}_{kl} = \frac{1}{a-1} \sum_{i=1}^a (1 - \frac{n_i}{N}) \tilde{\sigma}_{kl:i}$, under the marginal hypothesis, the expectation $E(\tilde{\sigma}_{kl:i}) = \sigma_{kl:i}$ still depends on i , therefore $\tilde{\sigma}_{kl} = \frac{1}{a} \sum_{i=1}^a \tilde{\sigma}_{kl:i} + o_p(1) \xrightarrow{p} \lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \sigma_{kl:i} = \sigma_{kl}$. \square

Proof of Theorem 4

$$\begin{aligned} \frac{1}{N^2} (H_1 - G_1) &= \frac{1}{a-1} \sum_{i=1}^a \sum_{j=1}^{n_i} (\hat{\mathbf{Y}}_i - \hat{\mathbf{Y}}_{..})(\hat{\mathbf{Y}}_i - \hat{\mathbf{Y}}_{..})' - \frac{1}{N-a} \sum_{i=1}^a \sum_{j=1}^{n_i} (\hat{\mathbf{Y}}_{ij} - \hat{\mathbf{Y}}_i)(\hat{\mathbf{Y}}_{ij} - \hat{\mathbf{Y}}_i)' \\ &= \hat{\mathbf{Y}}_\mu \left[\frac{1}{a-1} \left(\bigoplus_{i=1}^a \frac{1}{n_i} J_{n_i} - \frac{1}{N} J_N \right) - \frac{1}{N-a} P_{n_i} \right] \hat{\mathbf{Y}}'_\mu \\ &= \hat{\mathbf{Y}}_\mu (C_1 + C_2 + C_3) \hat{\mathbf{Y}}'_\mu \end{aligned}$$

$$\text{where } C_1 = \bigoplus_{i=1}^a \left(\frac{N-1}{(N-a)(a-1)n_i} - \frac{1}{N(a-1)} \right) (J_{n_i} - I_{n_i}),$$

$$C_2 = -\frac{1}{N(a-1)} \left(J_N - \bigoplus_{i=1}^a J_{n_i} \right),$$

$$\text{and } C_3 = \bigoplus_{i=1}^a \left(\frac{N-1}{(N-a)(a-1)n_i} - \frac{1}{N(a-1)} - \frac{1}{N-a} \right) I_{n_i}.$$

For C_1 and C_2 , we can apply Theorem A.1 and Corollary A.1 stated in the appendix to show that $\sqrt{a}(\hat{\mathbf{Y}}_\mu C_i \hat{\mathbf{Y}}'_\mu - \mathbf{Y}_\mu C_i \mathbf{Y}'_\mu) \xrightarrow{p} 0$ as $a \rightarrow \infty$, $i = 1, 2$.

In order to prove that $\sqrt{a}(\hat{\mathbf{Y}}_\mu C_3 \hat{\mathbf{Y}}'_\mu - \mathbf{Y}_\mu C_3 \mathbf{Y}'_\mu) \xrightarrow{p} 0$ as $a \rightarrow \infty$, consider first an arbitrary diagonal element $\sqrt{a} \sum_{i=1}^a \sum_{j=1}^{n_i} w_i \left((\hat{Y}_{ij}^{(k)})^2 - (Y_{ij}^{(k)})^2 \right)$ where $w_i = \left(\frac{N-1}{(N-a)(a-1)n_i} - \frac{1}{N(a-1)} - \frac{1}{N-a} \right)$. We simplify notation by collapsing the two indices (i, j) into one index $i = 1, \dots, N$ and by defining

$$\phi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) = \left(c(X_i^{(k)} - X_j^{(k)})c(X_i^{(k)} - X_l^{(k)}) - H^{(k)}(X_i^{(k)})^2 \right),$$

as well as

$$\psi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) = \phi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) - \mathbf{E}[\phi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) | X_j^{(k)}] - \mathbf{E}[\phi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) | X_l^{(k)}],$$

and we obtain

$$\begin{aligned} \sqrt{a} \sum_{i=1}^N w_i (\hat{Y}_i^2 - Y_i^2) &= \sqrt{a} \sum_{i=1}^N w_i \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \phi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) \\ &= \sqrt{a} \frac{1}{N^2} \sum_{(i,j,l) \in \mathcal{A}} w_i \psi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) + o_p(1). \end{aligned}$$

Here, \mathcal{A} denotes the set of index triples (i, j, l) consisting of three distinct numbers. The summation over the set of index triples where at least two of the three indices are equal is a term of order $o_p(1)$ because $w_i = O(1/a)$ and the expression in parentheses is bounded. Also, note that $\sum_{i=1}^N w_i = 0$, so that the sum of the conditional expectations is equal to zero, as well.

The conditional expectation $\mathbf{E}(\psi(X_i^{(k)}, X_j^{(k)}, X_l^{(k)}) | X_v^{(k)})$ equals zero for every $v = 1, \dots, N$. Therefore,

$$\begin{aligned} &\mathbf{E} \left(\sqrt{a} \sum_{i=1}^N w_i (\hat{Y}_i^2 - Y_i^2) \right) \\ &= \frac{a}{N^4} \sum_{(i_1, j_1, l_1) \in \mathcal{A}} \sum_{(i_2, j_2, l_2) \in \mathcal{A}} w_{i_1} w_{i_2} \mathbf{E} \left(\psi(X_{i_1}^{(k)}, X_{j_1}^{(k)}, X_{l_1}^{(k)}) \psi(X_{i_2}^{(k)}, X_{j_2}^{(k)}, X_{l_2}^{(k)}) \right) + o_p(1) = o_p(1), \end{aligned}$$

because if the number of different indices among $i_1, j_1, l_1, i_2, j_2, l_2$ is either six or five, the expectation will be equal to zero, and thus the first term only contributes $o_p(1)$ as $a \rightarrow \infty$.

In a similar way, it can be shown for an arbitrary off-diagonal element (k, k') that

$$\sqrt{a} \sum_{i=1}^a \sum_{j=1}^{n_i} w_i \left(\hat{Y}_{ij}^{(k)} \hat{Y}_{ij}^{(l)} - Y_{ij}^{(k)} Y_{ij}^{(l)} \right) \xrightarrow{p} 0.$$

To this end, choose

$$\phi(X_i^{(k)}, X_j^{(k)}, X_i^{(k')}, X_l^{(k')}) = \left(c(X_i^{(k)} - X_j^{(k)}) c(X_i^{(k')} - X_l^{(k')}) - H^{(k)}(X_i^{(k)}) H^{(k')}(X_i^{(k')}) \right),$$

as well as

$$\begin{aligned} \psi(X_i^{(k)}, X_j^{(k)}, X_i^{(k')}, X_l^{(k')}) &= \phi(X_i^{(k)}, X_j^{(k)}, X_i^{(k')}, X_l^{(k')}) - \mathbf{E}[\phi(X_i^{(k)}, X_j^{(k)}, X_i^{(k')}, X_l^{(k')}) | X_j^{(k)}] \\ &\quad - \mathbf{E}[\phi(X_i^{(k)}, X_j^{(k)}, X_i^{(k')}, X_l^{(k')}) | X_l^{(k')}] . \end{aligned}$$

□

Proof of Theorem 5

$$\begin{aligned} \frac{1}{N^2}(H_1 - G_2) &= \frac{1}{a-1} \sum_{i=1}^a \left[\sum_{j=1}^{n_i} (\hat{\mathbf{Y}}_{i.} - \hat{\mathbf{Y}}_{..})(\hat{\mathbf{Y}}_{i.} - \hat{\mathbf{Y}}_{..})' - \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i-1} \sum_{j=1}^{n_i} (\hat{\mathbf{Y}}_{ij} - \hat{\mathbf{Y}}_{i.})(\hat{\mathbf{Y}}_{ij} - \hat{\mathbf{Y}}_{i.})' \right] \\ &= \hat{\mathbf{Y}}_{\mu} \left(\frac{1}{a-1} \bigoplus_{i=1}^a \left[\frac{1}{n_i} J_{n_i} - \frac{1}{N} J_N - \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i-1} P_{n_i} \right] \right) \hat{\mathbf{Y}}_{\mu}' \\ &= \hat{\mathbf{Y}}_{\mu} (C_1 + C_2) \hat{\mathbf{Y}}_{\mu}' \end{aligned}$$

$$\text{where } C_1 = \frac{1}{a-1} \bigoplus_{i=1}^a \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i-1} (J_{n_i} - I_{n_i})$$

$$\text{and } C_2 = -\frac{1}{N(a-1)} \left(J_N - \bigoplus_{i=1}^a J_{n_i} \right).$$

Applying Theorem A.1 and Corollary A.1, it follows that $\sqrt{a}(\hat{\mathbf{Y}}_{\mu} C_i \hat{\mathbf{Y}}_{\mu}' - \mathbf{Y}_{\mu} C_i \mathbf{Y}_{\mu}') \xrightarrow{p} 0$ as $a \rightarrow \infty$, $i = 1, 2$. \square

Proof of Theorem 6

$$\begin{aligned} \frac{1}{N^2}(H_2 - G_3) &= \hat{\mathbf{Y}}_{\mu} \left[\frac{1}{a-1} \left(\bigoplus_{i=1}^a \frac{1}{n_i} \mathbf{1}_{n_i} \right) P_a \left(\bigoplus_{i=1}^a \frac{1}{n_i} \mathbf{1}'_{n_i} \right) - \frac{1}{a} \bigoplus_{i=1}^a \frac{1}{n_i(n_i-1)} P_{n_i} \right] \hat{\mathbf{Y}}_{\mu}' \\ &= \hat{\mathbf{Y}}_{\mu} (C_1 + C_2) \hat{\mathbf{Y}}_{\mu}' \end{aligned}$$

$$\text{where } C_1 = \frac{1}{a} \bigoplus_{i=1}^a \frac{1}{n_i(n_i-1)} (J_{n_i} - I_{n_i})$$

$$\text{and } C_2 = -\frac{1}{a(a-1)} \left(\bigoplus_{i=1}^a \frac{1}{n_i} \mathbf{1}_{n_i} \right) (J_a - I_a) \left(\bigoplus_{i=1}^a \frac{1}{n_i} \mathbf{1}'_{n_i} \right).$$

Application of Theorem A.1 and Corollary A.1 yields that $\sqrt{a}(\hat{\mathbf{Y}}_{\mu} C_i \hat{\mathbf{Y}}_{\mu}' - \mathbf{Y}_{\mu} C_i \mathbf{Y}_{\mu}') \xrightarrow{p} 0$ as $a \rightarrow \infty$, $i = 1, 2$. \square

Proof of Theorem 7 By Theorem 4, it suffices to find the asymptotic distribution of $\sqrt{a}(\tilde{H}_1 - \tilde{G}_1)$. Now,

$$\begin{aligned} \tilde{H}_1 &= \frac{1}{a-1} \sum_{i=1}^a n_i (\bar{\mathbf{Y}}_{i.} - \bar{\mathbf{Y}}_{..})(\bar{\mathbf{Y}}_{i.} - \bar{\mathbf{Y}}_{..})' \\ &\doteq \frac{1}{a} \sum_{i=1}^a n_i (\bar{\mathbf{Y}}_{i.} - \bar{\mathbf{Y}}_{..})(\bar{\mathbf{Y}}_{i.} - \bar{\mathbf{Y}}_{..})' \\ &= \frac{1}{a} \sum_{i=1}^a n_i (\bar{\mathbf{Y}}_{i.} - \boldsymbol{\mu})(\bar{\mathbf{Y}}_{i.} - \boldsymbol{\mu}) - \frac{N}{a} (\bar{\mathbf{Y}}_{..} - \boldsymbol{\mu})(\bar{\mathbf{Y}}_{..} - \boldsymbol{\mu})'. \end{aligned}$$

Similarly,

$$\begin{aligned}\tilde{G}_1 &= \frac{1}{N-a} \sum_{i=1}^a \sum_{j=1}^{n_i} (\mathbf{Y}_{ij} - \mathbf{Y}_i)(\mathbf{Y}_{ij} - \mathbf{Y}_i)' \\ &= \frac{1}{N-a} \sum_{i=1}^a \sum_{j=1}^{n_i} [(\mathbf{Y}_{ij} - \boldsymbol{\mu}) - (\bar{\mathbf{Y}}_i - \boldsymbol{\mu})][(\mathbf{Y}_{ij} - \boldsymbol{\mu}) - (\bar{\mathbf{Y}}_i - \boldsymbol{\mu})]'.\end{aligned}$$

Then,

$$\begin{aligned}\sqrt{a}(\tilde{H}_1 - \tilde{G}_1) &\doteq \frac{1}{\sqrt{a}} \sum_{i=1}^a n_i (\bar{\mathbf{Y}}_i - \boldsymbol{\mu})(\bar{\mathbf{Y}}_i - \boldsymbol{\mu}) \\ &\quad - \frac{1}{\sqrt{a}(\bar{n}_a - 1)} \sum_{i=1}^a \sum_{j=1}^{n_i} [(\mathbf{Y}_{ij} - \boldsymbol{\mu}) - (\bar{\mathbf{Y}}_i - \boldsymbol{\mu})][(\mathbf{Y}_{ij} - \boldsymbol{\mu}) - (\bar{\mathbf{Y}}_i - \boldsymbol{\mu})]' \\ &\quad - \frac{N}{\sqrt{a}} (\bar{\mathbf{Y}}_{..} - \boldsymbol{\mu})(\bar{\mathbf{Y}}_{..} - \boldsymbol{\mu})' \\ &= \frac{1}{\sqrt{a}} \sum_{i=1}^a (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \frac{1}{n_i} J_{n_i} (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})' \\ &\quad - \frac{1}{\sqrt{a}(\bar{n}_a - 1)} \sum_{i=1}^a (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) P_{n_i} (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})' \\ &\quad - \frac{1}{\sqrt{a}} (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N) \frac{1}{N} J_N (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N)' \\ &= \frac{1}{\sqrt{a}} \sum_{i=1}^a (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \left(\frac{1}{n_i} J_{n_i} - \frac{1}{(\bar{n}_a - 1)} P_{n_i} \right) (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})' \\ &\quad - \frac{1}{\sqrt{a}} \mathbf{Y}_\mu \frac{1}{N} J_N \mathbf{Y}'_\mu\end{aligned}$$

Denote

$$Q = \frac{1}{\sqrt{a}} \mathbf{Y}_\mu \frac{1}{N} J_N \mathbf{Y}'_\mu.$$

It can easily be shown that

$$E(Q) = a^{-1/2} \Sigma_1 = o(1). \quad (9)$$

Moreover, by using Lemma A.1, we get

$$\text{var}(Q) = \frac{1}{a} (I_{p^2} + K_{p,p})(\Sigma_1 \otimes \Sigma_1) + \frac{1}{\sqrt{a}} K_4(Y_{11})$$

where

$$K_4(Y_{11}) = E[\text{vec}(Y_{11} Y'_{11}) \text{vec}(Y_{11} Y'_{11})'] - (I_{p^2} + K_{p,p})(\Sigma_1 \otimes \Sigma_1) - \text{vec}(\Sigma_1) \text{vec}(\Sigma_1)'. \quad (10)$$

Therefore $\text{var}(Q) = o(1)$. This together with (9) implies $Q \xrightarrow{p} 0$. Consequently,

$$\begin{aligned}\sqrt{a} \cdot \text{tr}(\tilde{H}_1 - \tilde{G}_1)A &\doteq \frac{1}{\sqrt{a}} \sum_{i=1}^a \text{tr} \left((\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \left(\frac{1}{n_i} J_{n_i} - \frac{1}{(\bar{n}_a - 1)} P_{n_i} \right) (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})' A \right) \\ &= \frac{1}{\sqrt{a}} \sum_{i=1}^a Z_i\end{aligned}$$

where

$$Z_i = \text{tr} \left((\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \left(\frac{1}{n_i} J_{n_i} - \frac{1}{(\bar{n}_a - 1)} P_{n_i} \right) (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})' A \right).$$

It may be noted that,

$$\frac{1}{\sqrt{a}} \sum_{i=1}^a E(Z_i) = \frac{1}{\sqrt{a}} \sum_{i=1}^a \left(1 - \frac{n_i - 1}{(\bar{n}_a - 1)} \right) \text{tr}(\Sigma_1 A) = 0$$

and

$$\frac{1}{a} \sum_{i=1}^a \text{var}(Z_i) = \frac{1}{a} \sum_{i=1}^a \text{vec}(A)' \text{var}(\mathbf{W}_i) \text{vec}(A) \quad (11)$$

where

$$\mathbf{W}_i = (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \left(\frac{1}{n_i} J_{n_i} - \frac{1}{(\bar{n}_a - 1)} P_{n_i} \right) (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})'.$$

But, applying Lemma A.1,

$$\text{var}(\mathbf{W}_i) = \left(1 + \frac{n_i - 1}{(\bar{n}_a - 1)^2} \right) (I_{p^2} + K_{p,p})(\Sigma_1 \otimes \Sigma_1) + \frac{1}{n_i} \left(\frac{\bar{n}_a - n_i}{\bar{n}_a - 1} \right)^2 K_4(\mathbf{Y}_{11}) \quad (12)$$

where $K_4(\mathbf{Y}_{11})$ is as defined in (10). Substituting (12) in (11) and making some simplifications,

$$\frac{1}{a} \sum_{i=1}^a \text{var}(Z_i) = 2 \left(1 + \frac{1}{\bar{n}_a - 1} \right) \text{tr}(\Sigma_1 A)^2 + \frac{\bar{n}_a(\bar{n}_a \underline{n}_a - 1)}{(\bar{n}_a - 1)^2} (\mu_4(A) - 2\text{tr}(\Sigma_1 A)^2 - (\text{tr} \Sigma_1 A)^2) \quad (13)$$

where $\underline{n}_a = (1/a) \sum_{i=1}^a (1/n_i)$. From (13) it follows that $\lim_{a \rightarrow \infty} (1/a) \sum_{i=1}^a \text{var}(Z_i) = \tau$. Finally, since $\sum_{i=1}^{\infty} \text{var}(Z_i) = \infty$ and Z_i is a bounded random variable, the Lindeberg condition holds. \square

Proof of Theorem 8 Let $A = (a_{ij})$ and $a_0 = \max_{i,j} |a_{ij}|$. The consistency of $\frac{1}{N^2} G_1$ for Σ_1 is proved in Theorem 1 above. Thus, we need to prove that $\frac{1}{N^4} \hat{\mu}_4(A) \xrightarrow{p} \mu_4(A)$. Note that $\frac{1}{N^4} \hat{\mu}_4 = \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} [(\hat{\mathbf{Y}}_{ij} - \frac{1}{2} \mathbf{1})' A (\hat{\mathbf{Y}}_{ij} - \frac{1}{2} \mathbf{1})]^2$. For the corresponding expression defined in terms of the ART, we have

$$\frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} [(\mathbf{Y}_{ij} - \boldsymbol{\mu}_{ij})' A (\mathbf{Y}_{ij} - \boldsymbol{\mu}_{ij})]^2 \xrightarrow{a.s.} E[(\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})' A (\mathbf{Y}_{11} - \boldsymbol{\mu}_{11})]^2,$$

due to the strong law of large numbers. Therefore, the proof will be finished by showing that

$$\frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} [(\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu})^2 - (\mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu})^2] \xrightarrow{p} 0,$$

where $\hat{\mathbf{Y}}_{ij,\mu} = \hat{\mathbf{Y}}_{ij} - \frac{1}{2} \mathbf{1}$ and $\mathbf{Y}_{ij,\mu} = \mathbf{Y}_{ij} - \boldsymbol{\mu}_{ij}$. Indeed,

$$\begin{aligned} & \mathbb{E} \left(\frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} [(\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu})^2 - (\mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu})^2] \right)^2 \leq \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \mathbb{E} \left((\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu})^2 - (\mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu})^2 \right)^2 \\ &= \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \mathbb{E} \left(\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu} - \mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu} \right)^2 \left(\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu} + \mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu} \right)^2 \\ &= \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \mathbb{E} \left(\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu} - \mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu} \right)^2 \left(\sum_{k=1}^p \sum_{l=1}^p a_{kl} (\hat{Y}_{ij,\mu}^{(k)} \hat{Y}_{ij,\mu}^{(l)} + Y_{ij,\mu}^{(k)} Y_{ij,\mu}^{(l)}) \right)^2 \\ &\leq \frac{4p^4 a_0^2}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \mathbb{E} \left(\hat{\mathbf{Y}}'_{ij,\mu} A \hat{\mathbf{Y}}_{ij,\mu} - \mathbf{Y}'_{ij,\mu} A \mathbf{Y}_{ij,\mu} \right)^2 \\ &\leq \frac{4p^4 a_0^2}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} p^4 a_0^2 \sum_{k=1}^p \sum_{l=1}^p \mathbb{E} (\hat{Y}_{ij,\mu}^{(k)} \hat{Y}_{ij,\mu}^{(l)} - Y_{ij,\mu}^{(k)} Y_{ij,\mu}^{(l)})^2 \\ &= \frac{4p^6 a_0^4}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \sum_{k=1}^p \sum_{l=1}^p \mathbb{E} \left(\hat{Y}_{ij,\mu}^{(k)} (\hat{Y}_{ij,\mu}^{(l)} - Y_{ij,\mu}^{(l)}) + Y_{ij,\mu}^{(l)} (\hat{Y}_{ij,\mu}^{(k)} - Y_{ij,\mu}^{(k)}) \right)^2 \\ &\leq \frac{8p^6 a_0^4}{N} \left[\sum_{i=1}^a \sum_{j=1}^{n_i} \sum_{k=1}^p \sum_{l=1}^p \mathbb{E} (\hat{Y}_{ij,\mu}^{(k)})^2 (\hat{Y}_{ij,\mu}^{(l)} - Y_{ij,\mu}^{(l)})^2 + \sum_{i=1}^a \sum_{j=1}^{n_i} \sum_{k=1}^p \sum_{l=1}^p \mathbb{E} (Y_{ij,\mu}^{(l)})^2 (\hat{Y}_{ij,\mu}^{(k)} - Y_{ij,\mu}^{(k)})^2 \right] \\ &\rightarrow 0. \end{aligned}$$

□

Proof of Corollary 1

$$\sqrt{a}(T_D - 1) = \sqrt{a} \left(\frac{(1/N^2) \text{tr}(H_1 - G_1)}{(1/N^2) \text{tr}(G_1)} \right) \doteq \sqrt{a} \frac{1}{N^2} \text{tr}(H_1 - G_1) \frac{1}{\text{tr} \Sigma_1}.$$

Choosing $A = (1/\text{tr} \Sigma_1) I_p$ in Theorem 7, we get the first result. The consistency of $\frac{1}{N^2} G_1$ for Σ_1 is proved in Theorem 1. The rest follows by choosing $A = I_p$ in Theorem 8. □

Proof of Corollary 2

$$\begin{aligned} \sqrt{a} \left(\left(\frac{N-a}{a-1} \right) T_{LH} - r_1 \right) &= \sqrt{a} (\text{tr}(H_1 G_1^-) - \text{tr}(G_1 G_1^-)) \\ &= \sqrt{a} \cdot \text{tr} \left(\frac{1}{N^2} (H_1 - G_1) N^2 G_1^- \right) \end{aligned}$$

$$\doteq \sqrt{a} \frac{1}{N^2} \text{tr}(H_1 - G_1) \Sigma_1^-$$

The last line follows from the fact that $(1/N^2)G_1 \xrightarrow{p} \Sigma_1$ (Theorem 1) and the continuity of the Moore-Penrose inverse. Now, setting $A = \Sigma_1^-$ and applying Theorem 7, we get the first result.

Regarding the consistency, note that,

$$r_1 = \text{tr} \left(\frac{1}{N^2} G_1 \left(\frac{1}{N^2} G_1 \right)^- \right) \xrightarrow{p} \text{tr}(\Sigma_1 \Sigma_1^-) = \rho_1.$$

Then, the result follows by choosing $A = \Sigma_1^-$ in Theorem 8 and observing,

$$\frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} \left[\hat{\mathbf{Y}}'_{ij} \left(\frac{1}{N^2} G_1 \right)^- \hat{\mathbf{Y}}_{ij} \right]^2 - \frac{1}{N} \sum_{i=1}^a \sum_{j=1}^{n_i} [\mathbf{Y}'_{ij} \Sigma_1^- \mathbf{Y}_{ij}]^2 = o_p(1)$$

because $(1/N^2)G_1^- = \Sigma_1^- + o_p(1)$ as $a \rightarrow \infty$. □

Proof of Corollary 3

$$\begin{aligned} \sqrt{a} \left(\frac{N-1}{N-a} \right) \left(\left(\frac{N-1}{a-1} \right) T_{BNP} - r_2 \right) &= \sqrt{a} \left(\frac{N-1}{N-a} \right) \left(\left(\frac{N-1}{a-1} \right) T_{BNP} \right. \\ &\quad \left. - \text{tr} \frac{1}{N-1} ((a-1)H_1 + (N-a)G_1) \left(\frac{1}{N-1} ((a-1)H_1 + (N-a)G_1) \right)^- \right). \end{aligned}$$

From Theorem 2 and continuity of the Moore-Penrose inverse we get,

$$N^2 \left(\frac{1}{N-1} ((a-1)H_1 + (N-a)G_1) \right)^- \xrightarrow{p} \Sigma_1^-. \quad (14)$$

Simplifying and using (14) yield

$$\sqrt{a} \left(\frac{N-1}{N-a} \right) \left(\left(\frac{N-1}{a-1} \right) T_{BNP} - r_2 \right) \doteq \sqrt{a} \cdot \text{tr} \frac{1}{N^2} (H_1 - G_1) \Sigma_1^-.$$

□

Proof of Theorem 9 By Theorem 5, it suffices to derive the asymptotic distribution of $\sqrt{a} \cdot \text{tr}(\tilde{H}_1 - \tilde{G}_2)A$.

$$\sqrt{a}(\tilde{H}_1 - \tilde{G}_2) = \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N) C_1 (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N)' - \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N) C_2 (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N)'$$

where C_1 and C_2 are as defined in the proof of Theorem 5. Denote

$$Q = \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N) C_2 (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N)'$$

Now,

$$E(Q) = \sqrt{a} \sum_{i=1}^a \sum_{k=1}^{n_i} \sum_{j=1}^a \sum_{l=1}^{n_j} c_{2:ij:kl} E((\mathbf{Y}_{ik} - \boldsymbol{\mu})(\mathbf{Y}_{jl} - \boldsymbol{\mu})')$$

where $c_{2:ij:kl}$ is the (k, l) th entry of the (i, j) th block of C_2 . Notice that $c_{2:ii:kl} = 0$ for $i = 1, \dots, a$. Also that $E((\mathbf{Y}_{ik} - \boldsymbol{\mu})(\mathbf{Y}_{jl} - \boldsymbol{\mu})') = 0$ for $i \neq j$. Therefore $E(Q) = 0$. Applying Lemma A.1 and noting that $c_{2:ii:kl} = 0$ for $i = 1, \dots, a$, we see that

$$\begin{aligned} \text{var}(Q) &= a \sum_{i=1}^a \sum_{j=1}^a \sum_{k=1}^a \sum_{l=1}^a c_{2:ij:kl}^2 (I_{p^2} + K_{p,p})(\Sigma_i \otimes \Sigma_j) \\ &= \frac{a}{N^2(a-1)^2} (I_{p^2} + K_{p,p}) \sum_{i \neq j} n_i n_j (\Sigma_i \otimes \Sigma_j) \\ &= \frac{a}{(a-1)^2} (I_{p^2} + K_{p,p}) \left(\sum_{i \neq j} \left(\frac{n_i}{N} \Sigma_i \right) \otimes \left(\frac{n_j}{N} \Sigma_j \right) \right) \\ &= O\left(\frac{1}{a}\right). \end{aligned}$$

The last line follows because the entries of \mathbf{Y}_{ij} are uniformly bounded which implies that

$$\left(\sum_{i \neq j} \left(\frac{n_i}{N} \Sigma_i \right) \otimes \left(\frac{n_j}{N} \Sigma_j \right) \right) = O(1).$$

Therefore $Q \xrightarrow{p} 0$. Let us next find the distribution of $\sqrt{a} \cdot \text{tr}(\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N) C_1 (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N)' = \sqrt{a} \cdot \text{tr} \mathbf{Y}_\mu C_1 \mathbf{Y}'_\mu$.

$$\sqrt{a} \cdot \text{tr}(\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N) C_1 (\mathbf{Y} - \boldsymbol{\mu} \mathbf{1}'_N)' A = \frac{\sqrt{a}}{a-1} \sum_{i=1}^a Z_i \doteq \frac{1}{\sqrt{a}} \sum_{i=1}^a Z_i$$

where

$$Z_i = \text{tr}(\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i - 1} (J_{n_i} - I_{n_i}) (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})' A.$$

Since the diagonal entries of $J_{n_i} - I_{n_i}$ are zeros and Y_{ij} 's are independent, it follows that $E(Z_i) = 0$.

Moreover,

$$\frac{1}{a} \sum_{i=1}^a \text{var}(Z_i) = \text{vec}(A')' \left(\frac{1}{a} \sum_{i=1}^a \text{var}(\mathbf{W}_i) \right) \text{vec}(A')$$

where

$$\mathbf{W}_i = (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i}) \left(1 - \frac{n_i}{N}\right) \frac{1}{n_i - 1} (J_{n_i} - I_{n_i}) (\mathbf{Y}_i - \boldsymbol{\mu} \mathbf{1}'_{n_i})'.$$

Applying Lemma A.1, we get

$$\text{var}(\mathbf{W}_i) = \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N}\right)^2 (I_{p^2} + K_{p,p})(\Sigma_i \otimes \Sigma_i).$$

Thus,

$$\begin{aligned}
\lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \text{var}(Z_i) &= \lim_{a \rightarrow \infty} \frac{1}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N}\right)^2 \text{vec}(A)'(I_{p^2} + K_{p,p})(\Sigma_i \otimes \Sigma_i) \text{vec}(A') \\
&= \lim_{a \rightarrow \infty} \frac{2}{a} \sum_{i=1}^a \frac{n_i}{n_i - 1} \left(1 - \frac{n_i}{N}\right)^2 \text{tr}(\Sigma_i A)^2 \\
&= \tau_2(A).
\end{aligned}$$

Since the Z_i are bounded random variables, the theorem is proved. \square

Proof of Corollary 4

$$\sqrt{a}(T_D^{(2)} - 1) = \sqrt{a} \left(\frac{(1/N^2)\text{tr}(H_1 - G_2)}{(1/N^2)\text{tr}(G_2)} \right) \doteq \sqrt{a} \frac{1}{N^2} \text{tr}(H_1 - G_2) \frac{1}{\text{tr}\Sigma}.$$

Choosing $A = (1/\text{tr}\Sigma)I_p$ in Theorem 9, we get the desired first result. Simplification and consistency under the multivariate hypothesis are straightforward. \square

Proof of Corollary 5

$$\begin{aligned}
\sqrt{a} \left(T_{LH}^{(2)} - r_1^{(2)} \right) &= \sqrt{a} (\text{tr}(H_1 G_2^-) - \text{tr}(G_2 G_2^-)) \\
&= \sqrt{a} \cdot \text{tr} \left(\frac{1}{N^2} (H_1 - G_2) N^2 G_2^- \right) \\
&\doteq \sqrt{a} \frac{1}{N^2} \text{tr}(H_1 - G_2) \Sigma^-
\end{aligned}$$

Now, setting $A = \Sigma^-$ and applying Theorem 9, we get the desired result. \square

Proof of Theorem 10 Here also it suffices to obtain the asymptotic distribution of $\sqrt{a} \cdot \text{tr}(H_2 - G_3)A$.

$$\sqrt{a}(\tilde{H}_2 - \tilde{G}_3) = \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)C_1(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)' + \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)C_2(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)'$$

where C_1 and C_2 are as defined in the proof of Theorem 6. Letting

$$Q = \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)C_2(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)'$$

we can show as in Theorems 4 and 5 that $Q \xrightarrow{p} 0$. Hence,

$$\begin{aligned}
\sqrt{a} \cdot \text{tr}(\tilde{H}_2 - \tilde{G}_3)A &\doteq \sqrt{a}(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)C_1(\mathbf{Y} - \boldsymbol{\mu}\mathbf{1}'_N)'A \\
&= \frac{1}{\sqrt{a}} \sum_{i=1}^a \text{tr} \left((\mathbf{Y}_i - \boldsymbol{\mu}\mathbf{1}'_{n_i}) \frac{1}{n_i(n_i - 1)} (J_{n_i} - I_{n_i})(\mathbf{Y}_i - \boldsymbol{\mu}\mathbf{1}'_{n_i})'A \right) \\
&= \frac{1}{\sqrt{a}} \sum_{i=1}^a Z_i
\end{aligned}$$

where

$$Z_i = \text{tr} \left((\mathbf{Y}_i - \boldsymbol{\mu}\mathbf{1}'_{n_i}) \frac{1}{n_i(n_i - 1)} (J_{n_i} - I_{n_i})(\mathbf{Y}_i - \boldsymbol{\mu}\mathbf{1}'_{n_i})'A \right).$$

Moreover, $E(Z_i) = 0$ and $\lim_{a \rightarrow \infty} (1/a) \sum_{i=1}^a \text{var}(Z_i) = \tau_3(A)$. \square

Proof of Corollary 6

$$\sqrt{a}(T_D^{(3)} - 1) = \sqrt{a} \left(\frac{(1/N^2)\text{tr}(H_2 - G_3)}{(1/N^2)\text{tr}(G_3)} \right) \doteq \sqrt{a} \frac{1}{N^2} \text{tr}(H_2 - G_3) \frac{1}{\text{tr}\check{\Sigma}}.$$

Choosing $A = (1/\text{tr}\check{\Sigma})I_p$ in Theorem 10, we get the desired result. \square

Proof of Corollary 7

$$\begin{aligned} \sqrt{a} \left(T_{LH}^{(3)} - r_1^{(3)} \right) &= \sqrt{a} \left(\text{tr}(H_2 G_3^-) - \text{tr}(G_3 G_3^-) \right) \\ &= \sqrt{a} \cdot \text{tr} \left(\frac{1}{N^2} (H_2 - G_3) N^2 G_3^- \right) \\ &\doteq \sqrt{a} \frac{1}{N^2} \text{tr}(H_2 - G_3) \check{\Sigma}^- \end{aligned}$$

Now, setting $A = \check{\Sigma}^-$ and applying Theorem 10, we get the desired result. \square

REFERENCES

- [1] U. Munzel, E. Brunner, Nonparametric Methods in Multivariate Factorial Designs, *Journal of Statistical Planning and Inference* 88, 1 (2000) 117–132.
- [2] U. Munzel, E. Brunner, Nonparametric Tests in the Unbalanced Multivariate One-Way Design, *Biometrical Journal* 42, 7 (2000) 837–854.
- [3] E. Brunner, U. Munzel, M.L. Puri, Rank-Score Tests in Factorial Designs with Repeated Measures, *Journal of Multivariate Analysis* 70, 2 (1999) 286–317.
- [4] J.A. Chatfield, E.A. Draper, K.D. Cochran, D.A. Herms, Evaluation of Crabapples for Apple Scab at the Secret Arboretum in Wooster, Ohio, The Ohio State University, Ohio Agricultural Research and Development Center Special Circular 177 (2000) Ornamental Plants: Annual Reports and Research Reviews.
- [5] M.A. Omer, D.A. Johnson, R.C. Rowe, Recovery of *Vorticillium dahliae* from North American Certified Seed Potatoes and Characterization of Strains by Vegetative Compatibility and Aggressiveness, *American Journal of Potato Research* 77 (2000) 325–331.
- [6] A.C. Bathke, S.W. Harrar, Nonparametric Methods in Multivariate Factorial Designs for Large Number of Factor Levels, Submitted (2006).
- [7] S.W. Harrar, A.C. Bathke, A Nonparametric Version of the Bartlett-Nanda-Pillai Multivariate Test. Asymptotics, Approximations, and Applications, Submitted (2006).
- [8] G.L. Thompson, Asymptotic Distribution of Rank Statistics under Dependencies with Multivariate Application, *Journal of Multivariate Analysis* 33 (1990) 183–211.
- [9] G.L. Thompson, A Unified Approach to Rank Tests for Multivariate and Repeated Measures Designs, *Journal of the American Statistical Association* 86 (1991) 410–419.
- [10] M.L. Puri, P.K. Sen, On A Class of Multivariate Multisample Rank-Order Tests, *Sankhya A* 28 (1966) 353–376.

- [11] P.K. Sen, Asymptotic Distribution of a Class of Multivariate Rank Order Statistics, *Calcutta Statistical Association Bulletin* 19, 73 (1970) 23–31.
- [12] P. Lévy, *Calcul des Probabilités*, Gauthiers-Villars, Paris, 1925.
- [13] W.H. Kruskal, A Nonparametric Test for the Several Sample Problem, *Annals of Mathematical Statistics* 23 (1952) 525–540.
- [14] F.H. Ruymgaart, A Unified Approach to the Asymptotic Distribution Theory of Certain Midrank Statistics, In: *Statistique non Paramétrique Asymptotique*, 1–18, J.P. Raoult (Ed.), *Lecture Notes on Mathematics*, No. 821, Springer, Berlin, 1980.
- [15] U. Munzel, Linear Rank Score Statistics When Ties Are Present, *Statistics and Probability Letters* 41 (1999) 389–395.
- [16] J.R. Schott, *Matrix Analysis for Statistics*, Wiley, New York, 1997.
- [17] H. Scheffé, *The Analysis of Variance*, Wiley Classics Library, New York, 1999.
- [18] M.G. Akritas, N. Papadatos, Heteroscedastic One-Way ANOVA and Lack-of-Fit Tests, *Journal of the American Statistical Association* 99, 466 (2004) 368–382.
- [19] H. Wang, H., M.G. Akritas, Rank Tests for ANOVA with Large Number of Factor Levels, *Journal of Nonparametric Statistics* 16 (2004) 563–589.
- [20] T.P. Hettmansperger, H. Oja, Affine Invariant Multivariate Multisample Sign Tests, *Journal of the Royal Statistical Society, Series B*, 56, (1994) 235–249.
- [21] T.P. Hettmansperger, J. Möttönen, H. Oja, Affine Invariant Multivariate Rank Tests for Several Samples, *Statistica Sinica*, 8, (1998) 785–800.
- [22] H. Oja, R.H. Randles, Multivariate Nonparametric Tests, *Statistical Science*, 19, (2004) 598–605.
- [23] H. Oja, Affine Invariant Multivariate Sign and Rank Tests and Corresponding Estimates: a Review, *Scandinavian Journal of Statistics*, 26 (1999), 319–343.
- [24] M.G. Akritas, The Rank Transform Method in Some Two-Factor Designs, *Journal of the American Statistical Association* 85 (1990) 73–78.
- [25] R. Gentleman, R. Ihaka, R: A Language for Data Analysis and Graphics, *Journal of Computational and Graphical Statistics* 5 (1996) 299–314.