Chapter 7
Rank-Based Inference for Multivariate Data in Factorial Designs

Arne C. Bathke and Solomon W. Harrar

Abstract We introduce fully nonparametric, rank-based test statistics for inference on multivariate data in factorial designs, and derive their asymptotic sampling distribution. The focus here is on the asymptotic setting where the number of levels of one factor tends to infinity, while the number of levels of the other factor, as well as the replication size per factor level combination, are fixed. The resulting test statistics can be calculated directly, they don't involve any iterative computational procedures. To our knowledge, they provide the first viable approach to a fully nonparametric analysis of, for example, multivariate ordinal responses, or a mix of ordinal with other response variables, in a factorial design setting.

7.1 Introduction

Multivariate data in factorial designs with few replications arise in agricultural, behavioral and biomedical studies, just to mention a few. However, due to the lack of appropriate inference procedures, such data are often analyzed using simplistic univariate methods or questionable model assumptions (e.g., multivariate normality). In this article, we develop fully nonparametric methods for the analysis of such data. These nonparametric methods allow for the analysis of data with ordinal responses, and they contain desirable invariance properties, as each variable (endpoint) can be monotonically transformed without changing the outcome of the analysis. Furthermore, the proposed methods represent the first nonparametric

---

A.C. Bathke (✉)
Fachbereich Mathematik, Universität Salzburg, 5020 Salzburg, Austria

Department of Statistics, University of Kentucky, Lexington, KY 40536, USA
e-mail: Arne.Bathke@sbg.ac.at

S.W. Harrar
Department of Statistics, University of Kentucky, Lexington, KY 40536, USA
e-mail: Solomon.Harrar@uky.edu

approach that is asymptotically valid when the number (*a*) of different samples or treatment groups (more generally, the number of levels of one of the factors) is large. In order to illustrate application of the procedure, we use the following data example.

7.1.1   Agricultural Field Trial

In an agricultural experiment that stands here exemplary for many similarly conducted field trials, several varieties of crabapples are examined with regard to their disease resistance (Chatfield et al. 2000). The response variable is a rating of tree health, on an ordinal scale from 0 to 5. Trees are evaluated at different times during the growing season, generating a multivariate observation vector per tree. When the experiment is repeated in a different year or at a different location, a second treatment factor is introduced whose main effect and interaction with the plant variety need to be considered, in addition to the variety effect. In Chatfield et al. (2000), the number of crabapple varieties was $a = 63$, justifying the use of methods derived for the asymptotic situation of $a \to \infty$. The number n_{ij} of replicates per variety were between 3 and 5. If we assume that the same study is performed at two different agricultural experiment stations or in two different years, we would be in the situations with $b = 2$.

7.1.2   Model

We describe the model using a two-factor layout corresponding to the data example. Generalization to higher-way layouts can be done using the techniques described here. On each experimental unit, a *p*-dimensional response vector is observed. These vectors are described by

$$\mathbf{X}_{ijr} = (X_{ijr}^{(1)}, X_{ijr}^{(2)}, \ldots, X_{ijr}^{(p)})'.$$

Here, the first two indices, $i = 1, \ldots, a$, and $j = 1, \ldots, b$, denote the levels of the two explanatory factors considered (in the example, year or location and variety, respectively). The index $r = 1, \ldots, n_{ij}$ stands for the replication or experimental unit within a factor level combination, and the super-index $d = 1, \ldots, p$ denotes the respective variable, among the total of *p* response variables considered (*p*-dimensional response). A possible multivariate additive linear model for $\mathbf{X}_{ijr}$ could be:

$$\mathbf{X}_{ijr} = \boldsymbol{\mu} + \boldsymbol{\xi}_i + \boldsymbol{\lambda}_j + \boldsymbol{\gamma}_{ij} + \boldsymbol{\varepsilon}_{ijr},$$

where $\boldsymbol{\xi}$, $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$ are the effects due to experimental condition, variety, and interaction between experimental condition and variety, and $\boldsymbol{\varepsilon}$ is the random variation assumed to be independently distributed with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{ij}$.

Some drawbacks of the linear model approach are that the results depend on the type of transformation used and can be heavily influenced by outliers. In this manuscript, we are proposing a completely nonparametric alternative to the linear model approach. This nonparametric model can be written as

$$\mathbf{X}_{ijr} \sim F_{ij}, \tag{7.1}$$

where F_{ij} is the multivariate p-dimensional distribution of the response vector for factor level combination (i, j). This model imposes no restriction on distributions or correlations of error terms or random effects. The dependence in the data, induced by observing several outcome variables on the same subject, is entirely absorbed by modeling them as multivariate observation vectors, allowing for arbitrary, unspecified dependence structures among the response variables. The vectors $\mathbf{X}_{ijr}$ are independent for different indices i, j, or r, but the components of the vectors are possibly dependent.

In this manuscript, we are proposing a completely nonparametric model for the analysis of multivariate data from factorial experiments, applicable in a variety of situations. Inferential methods for the two-factor heteroscedastic model have relatively been well developed in the univariate case in the parametric as well as nonparametric contexts (for the latter, see for example, the monograph Brunner et al. (2002), and the references therein). There is some recent work for the semiparametric multivariate counterpart (Harrar and Bathke 2012; Konietschke et al. 2015; Van Aelst and Willems 2011), and several procedures have been proposed under the assumption of multivariate normality (Belloni and Didier 2008; Girón and del Castillo 2010; Kawasaki and Seo 2012; Krishnamoorthy and Lu 2010; Krishnamoorthy and Yu 2004, 2012; Nel and Van der Merwe 1986; Zhang 2011, 2012; Zhang and Liu 2013). However, not much has been done under the nonparametric paradigm, in particular under the asymptotic framework of a large number of factor levels. This asymptotic setup is becoming increasingly popular due to high throughput diagnostics and other bioinformatics tools which generate massive amounts of data. More motivations for this type of asymptotics in agriculture, health sciences, and other disciplines are found in Boos and Brownie (1995), Akritas and Arnold (2000), Bathke (2002), Bathke (2004) and Harrar and Gupta (2007) in univariate settings, and Gupta et al. (2006), Gupta et al. (2008), Bathke and Harrar (2008) and Harrar and Bathke (2008) in the multivariate setting. Whereas the work of Gupta et al. (2006, 2008) is restricted to the equal covariance matrix case, Bathke and Harrar (2008) and Harrar and Bathke (2008) consider the single factor nonparametric situation.

In the following sections, hypotheses and corresponding test statistics are introduced, and their asymptotic properties are derived. A section is devoted to the cumbersome task of consistent estimation of the variance-covariance matrix, and one section shows empirical evidence regarding the performance of the proposed tests, based on a simulation study.

Regarding the notation, a block diagonal matrix with blocks $\mathbf{A}$ and $\mathbf{B}$ will be written as $\mathbf{A} \oplus \mathbf{B}$, and the Kronecker product of matrices will be denoted as $\mathbf{A} \otimes \mathbf{B}$. See, for example, Schott (2005, Sect. 8.2) or Harville (2008, Sect. 16.1) for the definition and basic properties of the Kronecker product.

7.2 Hypotheses and Test Statistics

In general notation, the two experimental factors are denoted as factor A and factor B, respectively. Based on the nonparametric model (7.1), we will test hypotheses pertaining to these factors. The hypotheses are formulated in terms of the distribution functions F_{ij}. To this end, define the vector $\mathbf{F} = (F_{11}, \ldots, F_{1b}, F_{21}, \ldots, F_{ab})$ of cumulative distribution functions. Here, we assume the normalized versions of the distribution functions, allowing naturally for ties (Kruskal 1952; Lévy 1925; Ruymgaart 1980) and thus not restricting the methodology to absolutely continuous distributions.

Particular hypotheses of interest will be of the form $\mathscr{H}^{\psi} : D_{\psi} \mathbf{F} = \mathbf{0}$, postulating absence of the effect ψ. More specifically,

$$
D_{\psi} = \begin{cases}
\mathbf{P}_a \otimes \frac{1}{b} \mathbf{J}_b, & \text{for } \psi = A, \\
\mathbf{P}_a \otimes \mathbf{I}_b, & \text{for } \psi = A|B, \\
\frac{1}{a} \mathbf{J}_a \otimes \mathbf{P}_b, & \text{for } \psi = B, \\
\mathbf{I}_a \otimes \mathbf{P}_b, & \text{for } \psi = B|A, \\
\mathbf{P}_a \otimes \mathbf{P}_b, & \text{for } \psi = AB,
\end{cases}
$$

where $\mathbf{I}_d$ is the $d \times d$ identity matrix, $\mathbf{J}_d$ is the $d \times d$ unity matrix (matrix of ones), and $\mathbf{P}_d = \mathbf{I}_d - d^{-1} \mathbf{J}_d$.

These nonparametric hypotheses imply their corresponding parametric counterparts (see, e.g., Brunner et al. 2002; Harrar and Bathke 2008). As an illustration in the univariate context, the interaction effect in a parametric linear model can be expressed as $(\mathbf{P}_a \otimes \mathbf{P}_b)\boldsymbol{\mu}$, where $\boldsymbol{\mu}$ is the lexicographically arranged vector of cell means, $\boldsymbol{\mu} = (\mu_{11}, \ldots, \mu_{1b}, \mu_{21}, \ldots, \mu_{ab})'$. The implication between nonparametric and parametric hypotheses is immediately clear when expressing $(\mathbf{P}_a \otimes \mathbf{P}_b)\boldsymbol{\mu}$ in terms of the distribution functions as $(\mathbf{P}_a \otimes \mathbf{P}_b) \int x d\mathbf{F}(x)$. The same relation holds between the multivariate nonparametric and parametric analogs. The converse of this relation is not true: the parametric hypotheses do not imply their nonparametric counterparts.

In order to define nonparametric (rank-based) test statistics, let $\mathbf{R}_{ij} = (\mathbf{R}_{ij1}, \mathbf{R}_{ij2}, \ldots, \mathbf{R}_{ijn_{ij}})$ where $\mathbf{R}_{ijk} = (R_{ijk}^{(1)}, \ldots, R_{ijk}^{(p)})'$ and $R_{ijk}^{(l)}$ is the (mid-)rank of $X_{ijk}^{(l)}$ among all $N = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}$ random variables $X_{111}^{(l)}, \ldots, X_{abn_{ab}}^{(l)}$. Use of mid-ranks follows naturally from the normalized version of the cumulative distribution function (see above). Arranging these mid-ranks $R_{ijk}^{(l)}$ into a $p \times N$ matrix, put $\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_a)$ where $\mathbf{R}_i = (\mathbf{R}_{i1}, \mathbf{R}_{i2}, \ldots, \mathbf{R}_{ib})$. Then, denote the $p \times p$ hypothesis and error sum of squares and cross product matrices based on the ranks

as $\mathbf{H}^{(A)}(\mathbf{R})$, $\mathbf{H}^{(A|B)}(\mathbf{R})$, $\mathbf{H}^{(AB)}(\mathbf{R})$ and $\mathbf{G}(\mathbf{R})$. The corresponding matrices for testing main and simple effects of factor B can be written analogously to those of factor A. However, due to the large a asymptotics considered in this manuscript, we will not consider tests for the main effect of factor B in detail here.

$$\mathbf{H}^{(A)} = \frac{1}{a-1} \sum_{i=1}^{a} \sum_{j=1}^{b} (\tilde{\mathbf{R}}_{i..} - \tilde{\mathbf{R}}_{...})(\tilde{\mathbf{R}}_{i..} - \tilde{\mathbf{R}}_{...})'$$

$$= \frac{1}{a-1} \mathbf{R} \left[\left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}_{n_{ij}} \right) (\mathbf{P}_a \otimes \frac{1}{b} \mathbf{J}_b) \left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}'_{n_{ij}} \right) \right] \mathbf{R}',$$

$$\mathbf{H}^{(A|B)} = \frac{1}{(a-1)b} \sum_{i=1}^{a} \sum_{j=1}^{b} (\tilde{\mathbf{R}}_{ij.} - \tilde{\mathbf{R}}_{.j.})(\tilde{\mathbf{R}}_{ij.} - \tilde{\mathbf{R}}_{.j.})'$$

$$= \frac{1}{(a-1)b} \mathbf{R} \left[\left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}_{n_{ij}} \right) (\mathbf{P}_a \otimes \mathbf{I}_b) \left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}'_{n_{ij}} \right) \right] \mathbf{R}',$$

$$\mathbf{H}^{(B)} = \frac{1}{b-1} \sum_{i=1}^{a} \sum_{j=1}^{b} (\tilde{\mathbf{R}}_{.j.} - \tilde{\mathbf{R}}_{...})(\tilde{\mathbf{R}}_{.j.} - \tilde{\mathbf{R}}_{...})'$$

$$= \frac{1}{b-1} \mathbf{R} \left[\left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}_{n_{ij}} \right) (\frac{1}{a} \mathbf{J}_a \otimes \mathbf{P}_b) \left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}'_{n_{ij}} \right) \right] \mathbf{R}',$$

$$\mathbf{H}^{(B|A)} = \frac{1}{a(b-1)} \sum_{i=1}^{a} \sum_{j=1}^{b} (\tilde{\mathbf{R}}_{ij.} - \tilde{\mathbf{R}}_{i..})(\tilde{\mathbf{R}}_{ij.} - \tilde{\mathbf{R}}_{i..})'$$

$$= \frac{1}{a(b-1)} \mathbf{R} \left[\left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}_{n_{ij}} \right) (\mathbf{I}_a \otimes \mathbf{P}_b) \left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}'_{n_{ij}} \right) \right] \mathbf{R}',$$

$$\mathbf{H}^{(AB)} = \frac{1}{(a-1)(b-1)} \sum_{i=1}^{a} \sum_{j=1}^{b} (\tilde{\mathbf{R}}_{ij.} - \tilde{\mathbf{R}}_{i..} - \tilde{\mathbf{R}}_{.j.} + \tilde{\mathbf{R}}_{...})(\tilde{\mathbf{R}}_{ij.} - \tilde{\mathbf{R}}_{i..} - \tilde{\mathbf{R}}_{.j.} + \tilde{\mathbf{R}}_{...})'$$

$$= \frac{1}{(a-1)(b-1)} \mathbf{R} \left[\left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}_{n_{ij}} \right) (\mathbf{P}_a \otimes \mathbf{P}_b) \left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{1}'_{n_{ij}} \right) \right] \mathbf{R}', \quad \text{and}$$

$$\mathbf{G} = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)} \sum_{k=1}^{n_{ij}} (\mathbf{R}_{ijk} - \bar{\mathbf{R}}_{ij.})(\mathbf{R}_{ijk} - \bar{\mathbf{R}}_{ij.})' = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}} \mathbf{S}_{ij}$$

$$= \frac{1}{ab} \mathbf{R} \left(\bigoplus_{i=1}^{a} \bigoplus_{j=1}^{b} \frac{1}{n_{ij}(1-n_{ij})} \mathbf{P}_{n_{ij}} \right) \mathbf{R}',$$

where $\bar{\mathbf{R}}_{ij.} = \frac{1}{n_{ij}} \sum\limits_{k=1}^{n_{ij}} \mathbf{R}_{ijk}$, $\tilde{\mathbf{R}}_{i..} = \frac{1}{b} \sum\limits_{j=1}^{b} \bar{\mathbf{R}}_{ij.}$, $\tilde{\mathbf{R}}_{.j.} = \frac{1}{a} \sum\limits_{i=1}^{a} \bar{\mathbf{R}}_{ij.}$, $\tilde{\mathbf{R}}_{...} = \frac{1}{ab} \sum\limits_{i=1}^{a} \sum\limits_{j=1}^{b} \bar{\mathbf{R}}_{ij.}$, and

$\mathbf{S}_{ij} = \frac{1}{(n_{ij}-1)} \sum\limits_{k=1}^{n_{ij}} (\mathbf{R}_{ijk} - \bar{\mathbf{R}}_{ij.})(\mathbf{R}_{ijk} - \bar{\mathbf{R}}_{ij.})'$.

These sum of squares and cross product matrices constitute essentially a nonparametric multivariate unweighted means analysis. The matrix notation above shows the pattern after which they can also be defined in higher-way layouts. To keep this manuscript concise, this extension to higher-way layouts is not carried out in detail here. Under the hypothesis $\mathscr{H}^{\psi}$, the expectation of $\mathbf{H}^{(\psi)}$ is equal to the expectation of $\mathbf{G}$, thus allowing for the following way to construct multivariate test statistics.

Let ψ be one of the effects under consideration: AB, $A|B$, A, B, or $B|A$. We propose the following multivariate test statistics for testing $\mathscr{H}^{\psi}$.

(a) Dempster's ANOVA Type criterion: $T_{\mathrm{D}}^{(\psi)} = \mathrm{tr}(\mathbf{H}^{(\psi)})/\mathrm{tr}(\mathbf{G})$.
(b) Wilks' Λ criterion: $T_{\mathrm{LR}}^{(\psi)} = \log|\mathbf{I} + \mathbf{H}^{(\psi)}\mathbf{G}^-|$.
(c) The Lawley-Hotelling criterion: $T_{\mathrm{LH}}^{(\psi)} = \mathrm{tr}(\mathbf{H}^{(\psi)}\mathbf{G}^-)$.
(d) The Bartlett-Nanda-Pillai criterion: $T_{\mathrm{BNP}}^{(\psi)} = \mathrm{tr}\left(\mathbf{H}^{(\psi)}\mathbf{G}^-(\mathbf{I} + \mathbf{H}^{(\psi)}\mathbf{G}^-)\right)$.

These test statistics are similar to the four test statistics considered in Harrar and Bathke (2012) in the context of a two-factor semiparametric MANOVA under heteroscedasticity. Their use in this manuscript is distinct in two important ways. In the present article, the sums of squares and cross products $\mathbf{H}^{(\psi)}$ and $\mathbf{G}$ are computed from the ranks which can not be assumed to be independent across subjects. Due to the discreteness of the rankings, it may not be reasonable to assume non-singularity of the matrices $\mathbf{G}$ and $\mathbf{H}^{(\psi)} + \mathbf{G}$. Thus we use here Moore-Penrose generalized inverses in defining the test statistics. The Moore-Penrose generalized inverse has the useful continuity property (Schott 2005, Sect. 5.7; for a proof see, e.g., Penrose 1955).

7.3 Asymptotic Results

For the asymptotic derivations in this section, we will assume that $a \to \infty$, b bounded, and $\forall i, j : n_{ij}$ bounded. The asymptotics are somewhat involved as the quadratic forms $\mathbf{H}^{(\psi)}$ and $\mathbf{G}$ are based on a matrix of ranks $\mathbf{R}$ which has both row-wise and column-wise dependence.

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. For the concept of ART, see also Brunner et al. (2002, p. 77).

Definition 7.1. Let $\mathbf{X}_{ijk} = (X_{ijk}^{(1)}, \ldots, X_{ijk}^{(p)})'$, $i = 1, \ldots, a$, $j = 1, \ldots, b$, and $k = 1, \ldots, n_{ij}$, be independent random vectors with possibly dependent components $X_{ijk}^{(l)}$ whose marginal distribution is $F_{ij}^{(l)}$, $l = 1, \ldots, p$. Let $N = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}$. Further let

$$H^{(l)}(x) = \frac{1}{N} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} F_{ij}^{(l)}(x)$$

denote the *average cdf* for variable (l),

$$\hat{H}^{(l)}(x) = \frac{1}{N} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n_{ij}} c(x - X_{ijk}^{(l)}),$$

where $c(t) = 0, 1/2, 1$ if $t < 0, t = 0, t > 0$, respectively, denotes the *average empirical cdf*, and $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_a)$ where $\mathbf{Y}_i = (\mathbf{Y}_{i1}, \ldots, \mathbf{Y}_{ib})$, $\mathbf{Y}_{ij} = (\mathbf{Y}_{ij1}, \ldots, \mathbf{Y}_{ijn_{ij}})$ and $\mathbf{Y}_{ijk} = (Y_{ijk}^{(1)}, \ldots, Y_{ijk}^{(p)})'$ where $Y_{ijk}^{(l)} = H^{(l)}(X_{ijk}^{(l)})$ is known as the *asymptotic rank transform (ART)* of $X_{ijk}^{(l)}$. The matrix of *rank transforms (RT)*, $\hat{\mathbf{Y}}$, is defined analogously, with elements $\hat{Y}_{ijk}^{(l)} = \hat{H}^{(l)}(X_{ijk}^{(l)})$.

The expression "rank transform" pays tribute to the fact that $\hat{Y}_{ijk}^{(l)}$ is related to the (mid-)rank $R_{ijk}^{(l)}$ by $\hat{Y}_{ijk}^{(l)} = N^{-1}(R_{ijk}^{(l)} - \frac{1}{2})$. 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.

Denote $\mathrm{Var}(\mathbf{Y}_{ij1}) = \boldsymbol{\Sigma}_{ij}$ and assume that the following limit exists:

$$\lim_{a \to \infty} \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \boldsymbol{\Sigma}_{ij} = \boldsymbol{\Sigma}.$$

For later use, we also introduce the notation $\mathbf{M} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_a)$, $\boldsymbol{\mu}_i = (\boldsymbol{\mu}_{i1}, \ldots, \boldsymbol{\mu}_{ib})$, $\boldsymbol{\mu}_{ij} = (\boldsymbol{\mu}_{ij1}, \ldots, \boldsymbol{\mu}_{ijn_{ij}})$, where $\boldsymbol{\mu}_{ijk} = (\mu_{ijk}^{(1)}, \ldots, \mu_{ijk}^{(p)})'$ is the vector of expectations of the ART vector $\mathbf{Y}_{ijk}$, that is $\mu_{ijk}^{(l)} = E(Y_{ijk}^{(l)})$, and $\mathbf{Y}_\mu = \mathbf{Y} - \mathbf{M}$, $\hat{\mathbf{Y}}_\mu = \hat{\mathbf{Y}} - \mathbf{M}$.

For $\psi \in \{A, B, A|B, B|A, AB\}$, we denote the ART analogs of the matrices $\mathbf{H}^{(\psi)}$ and $\mathbf{G}$ defined in Sect. 7.2 by $\tilde{\mathbf{H}}^{(\psi)}$ and $\tilde{\mathbf{G}}$, respectively. In order to prove asymptotic normality results for the rank-based test statistics considered in this paper, we need to first establish the asymptotic equivalence of certain quadratic forms defined in terms of $(\mathbf{H}^{(\psi)}, \mathbf{G})$ (based on "rank transforms") and the corresponding quadratic forms defined in terms of $(\tilde{\mathbf{H}}^{(\psi)}, \tilde{\mathbf{G}})$ (based on "asymptotic rank transforms").

We begin this task by showing the asymptotic equivalence between certain matrix differences in "rank transforms' and the corresponding ones in "asymptotic

rank transforms". Recall that the ranks $R_{ijk}^{(l)}$ take values in $[1, N]$, while the "rank transforms" $\hat{Y}_{ijk}^{(l)} = N^{-1}(R_{ijk}^{(l)} - \frac{1}{2})$ and the "asymptotic rank transforms" $Y_{ijk}^{(l)}$ take values within the unit interval, making it necessary to divide the rank matrices by N^2 in order to be able to establish asymptotic equivalence.

Proposition 7.1. *Assume b, p and n are bounded. Then*

(i) Under the hypothesis $\mathscr{H}^\psi$ for $\psi \in \{A, A|B, AB\}$

$$\sqrt{a} \left\{ \frac{1}{N^2}(\mathbf{H}^{(\psi)} - \mathbf{G}) - (\tilde{\mathbf{H}}^{(\psi)} - \tilde{\mathbf{G}}) \right\} = o_p(1) \quad as \ a \to \infty \ .$$

(ii) Under the hypothesis $\mathscr{H}^\psi$ for $\psi \in \{B, B|A\}$

$$\frac{1}{N^2}\mathbf{H}^{(\psi)} - \tilde{\mathbf{H}}^{(\psi)} = o_p(1) \quad as \ a \to \infty \ .$$

and
(iii) $N^{-2}\mathbf{G} - \tilde{\mathbf{G}} = o_p(1)$ as $a \to \infty$.

Proof. The proof can be established using the same techniques as in the proof of Theorem 4 in Harrar and Bathke (2008).

The following proposition asserts that the difference $N^{-2}\mathbf{G} - \boldsymbol{\Sigma}$ is asymptotically $(a \to \infty)$ stochastically negligible.

Proposition 7.2. *Assume that the n_{ij} are bounded. Then $N^{-2}\mathbf{G} - \boldsymbol{\Sigma} \xrightarrow{p} 0$ as $a \to \infty$.*

Proof. Since $N^{-2}\mathbf{G} - \tilde{\mathbf{G}} = o_p(1)$ by (iii) of Proposition 7.1, it suffices to show that $\tilde{\mathbf{G}} - \boldsymbol{\Sigma} \xrightarrow{p} 0$. This follows from Theorem 1 of Harrar and Bathke (2012) if we show that $\sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}^{-2}(n_{ij} - 1)^{-1} \boldsymbol{\Sigma}_{ij} \otimes \boldsymbol{\Sigma}_{ij} = o(a^2)$ and $\sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}^{-3} K_4(\mathbf{Y}_{ij1}) = o(a^2)$ as $a \to \infty$. These two follow from the fact that the components of $\mathbf{Y}_{ij1}$ are uniformly bounded random variables.

Next, we obtain the asymptotic null distributions of the four test statistics for testing the main, simple, and interaction effects. Since the results for testing $\mathscr{H}^{(AB)}$, $\mathscr{H}^{(A)}$, $\mathscr{H}^{(A|B)}$ and $\mathscr{H}^{(B|A)}$ are similar in form and their derivations proceed along the same lines, we consider them together.

We know from Proposition 7.2 that $N^{-2}\mathbf{G} - \boldsymbol{\Sigma} = o_p(1)$ as $a \to \infty$, and it is established in Theorem 7.1 below that $\sqrt{a}(\mathbf{H}^{(\psi)} - \mathbf{G})\boldsymbol{\Omega} = O_p(1)$ as $a \to \infty$, for any matrix of constants $\boldsymbol{\Omega}$. All four test statistics, scaled and centered suitably, can be expressed as

$$\sqrt{a} \left(\ell T_{\mathscr{G}}^{(\psi)} - h \right) = \sqrt{a} \ \mathrm{tr}(\mathbf{H}^{(\psi)} - \mathbf{G})\boldsymbol{\Omega} + o_p(1), \tag{7.2}$$

where $\ell = 1, 2, 1, 4$, $h = 1, 2p \log 2, p, 2p$ and $\boldsymbol{\Omega} = (1/\mathrm{tr}\boldsymbol{\Sigma})\mathbf{I}_p, \boldsymbol{\Sigma}^-, \boldsymbol{\Sigma}^-, \boldsymbol{\Sigma}^-$ for $\mathscr{G} = \mathrm{D, LR, LH, BNP}$, respectively (see Harrar and Bathke 2012, for more

details). Therefore, the null distributions of the four test statistics can be derived in a unified manner by obtaining the null distribution of $\sqrt{a}\,\text{tr}(\mathbf{H}^{(\psi)} - \mathbf{G})\boldsymbol{\Omega}$ for any fixed matrix $\boldsymbol{\Omega}$. The null distribution of this latter quantity is given in Theorem 7.1.

Theorem 7.1. *Let* $\psi = AB, A, A|B,$ *or* $B|A.$ *Under the hypothesis* $\mathcal{H}_0^{(\psi)},$
$\sqrt{a}\,\text{tr}(\mathbf{H}^{(\psi)} - \mathbf{G})\boldsymbol{\Omega} \xrightarrow{\mathcal{L}} N\left(0, \tau_\psi^2(\boldsymbol{\Omega})\right)$ *as* $a \to \infty$ *and* n_{ij} *and* b *bounded, where*

$$
\tau_\psi^2(\boldsymbol{\Omega}) = \begin{cases}
\frac{2}{b}\left\{v_1(\boldsymbol{\Omega}) + \frac{v_2(\boldsymbol{\Omega})}{(b-1)^2}\right\} & \text{when} \quad \psi = AB, \\[2mm]
\frac{2}{b}\left\{v_1(\boldsymbol{\Omega}) + v_2(\boldsymbol{\Omega})\right\} & \text{when} \quad \psi = A, \\[2mm]
\frac{2}{b}v_1(\boldsymbol{\Omega}) & \text{when} \quad \psi = A|B, \\[2mm]
\frac{2}{b^2}\left\{v_1(\boldsymbol{\Omega}) + \frac{v_2(\boldsymbol{\Omega})}{(b-1)^2}\right\} & \text{when} \quad \psi = B|A.
\end{cases}
$$

Here,

$$
v_1(\boldsymbol{\Omega}) = \lim_{a \to \infty} \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{\text{tr}(\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij})^2}{n_{ij}(n_{ij} - 1)},
$$

and

$$
v_2 = \lim_{a \to \infty} \frac{1}{ab} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \frac{\text{tr}(\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij}\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij'})}{n_{ij}n_{ij'}},
$$

assuming the limits exist.

Proof. Considering Proposition 7.1, it is enough to show that $N^{-2}\sqrt{a}\,\text{tr}(\mathbf{H}^{(\psi)} - \mathbf{G})\boldsymbol{\Omega} \xrightarrow{\mathcal{L}} N\left(0, \tau_\psi^2(\boldsymbol{\Omega})\right)$ as $a \to \infty$ and n_{ij} and b bounded. This follows from Theorem 2 of Harrar and Bathke (2012) if for some $\delta > 0$, $E|(\mathbf{Y}_{ij1} - \frac{1}{2}\mathbf{1})'\boldsymbol{\Sigma}_{ij}^{-1}(\mathbf{Y}_{ij1} - \frac{1}{2}\mathbf{1})|^{2+\delta} < \infty$ and

$$
\lim_{a \to \infty} \frac{1}{a} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}^{1+\delta/2}(n_{ij} - 1)^{1+\delta/2}}\text{tr}(\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij})^{2+\delta} < \infty \text{ and}
$$

$$
\lim_{a \to \infty} \frac{1}{a} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \frac{1}{n_{ij}^{1+\delta/2}n_{ij'}^{1+\delta/2}}\text{tr}(\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij}\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij'})^{1+\delta/2} < \infty.
$$

Recalling again that the components of $\mathbf{Y}_{ij1}$ are bounded random variables completes the proof.

Under the assumptions and notations of Theorem 7.1, the asymptotic distribution of Dempster's ANOVA type criterion can be obtained by setting $\boldsymbol{\Omega} = (1/\mathrm{tr}\,\boldsymbol{\Sigma})\mathbf{I}_p$. For the other three criteria, we set $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$ to get the asymptotic null distributions.

Needless to say, the asymptotic null distributions of T_{LR}, T_{LH} and T_{BNP}, scaled and centered as in (7.2), are the same up to the order $O(a^{-1/2})$. A comparison of the asymptotic variances in Theorem 7.1 reveals that the test statistic for the interaction effect has smaller variance than that of the main effect. Also we see from the asymptotic variances in Theorem 7.1 that the test statistic for the simple effect of A has smaller variance compared to that of either the interaction or main effects.

7.4 Consistent Variance and Covariance Matrix Estimation

Multivariate data in factorial designs present a major technical difficulty considering the derivation of valid nonparametric test statistics: Unlike in the multivariate one-way design discussed in Harrar and Bathke (2008), the covariance matrices do not simplify under the null hypotheses that are considered here. Therefore, it is more complicated to devise consistent variance estimators.

The following theorem provides an asymptotic result formulated in terms of the unobservable "asymptotic rank transforms". The expression is analogous to the variance estimator defined in Theorem 2.3 of Harrar and Bathke (2012) in the semiparametric context. However, due to the fact that the "asymptotic rank transforms" are per definition bounded between 0 and 1, it is not necessary to require a moment condition as in Harrar and Bathke (2012).

Theorem 7.2. *Let the model and assumptions be as in Theorem 7.1. Define*

$$\widetilde{\boldsymbol{\psi}}_{ij}(\boldsymbol{\Omega}) = \frac{1}{4c_{ij}} \sum_{(k_1,k_2,k_3,k_4)\in\mathscr{K}} \boldsymbol{\Omega}\,(\mathbf{Y}_{ijk_1} - \mathbf{Y}_{ijk_2})(\mathbf{Y}_{ijk_1} - \mathbf{Y}_{ijk_2})'$$
$$\times\, \boldsymbol{\Omega}\,(\mathbf{Y}_{ijk_3} - \mathbf{Y}_{ijk_4})(\mathbf{Y}_{ijk_3} - \mathbf{Y}_{ijk_4})',$$

where $\mathscr{K}$ is the set of all quadruples $\kappa = (k_1, k_2, k_3, k_4)$ where no element in κ is equal to any other element in κ, and $c_{ij} = n_{ij}(n_{ij}-1)(n_{ij}-2)(n_{ij}-3)$. Also, define

$$\tilde{\mathbf{S}}_{ij} = \frac{1}{(n_{ij}-1)} \sum_{k=1}^{n_{ij}} (\mathbf{Y}_{ijk} - \bar{\mathbf{Y}}_{ij.})(\mathbf{Y}_{ijk} - \bar{\mathbf{Y}}_{ij.})'.$$

Then,

$$\frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}(n_{ij}-1)}\mathrm{tr}(\widetilde{\boldsymbol{\psi}}_{ij}(\boldsymbol{\Omega})) - \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}(n_{ij}-1)}\mathrm{tr}(\boldsymbol{\Omega}\,\boldsymbol{\Sigma}_{ij})^2 = o_p(1),$$

and

$$\frac{1}{ab} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \frac{1}{n_{ij} n_{ij'}} \operatorname{tr}(\boldsymbol{\Omega} \tilde{\mathbf{S}}_{ij} \boldsymbol{\Omega} \tilde{\mathbf{S}}_{ij'}) - \frac{1}{ab} \sum_{i=1}^{a} \sum_{j \neq j'}^{b} \frac{1}{n_{ij} n_{ij'}} \operatorname{tr}(\boldsymbol{\Omega} \boldsymbol{\Sigma}_{ij} \boldsymbol{\Omega} \boldsymbol{\Sigma}_{ij'}) = o_p(1),$$

as $a \to \infty$.

The proof follows similar to that of Theorem 2.3 in Harrar and Bathke (2012), or rather from the theory of U-statistics (see, e.g., Serfling 1980).

Since the "variance estimator" presented in the previous theorem is not observable and therefore can not be used in practice, in the next two theorems we are introducing observable rank-based estimators and establish their asymptotic equivalence to corresponding expressions formulated in terms of the "asymptotic rank transforms".

Theorem 7.3. *Let $\widetilde{\boldsymbol{\Psi}}_{ij}(\boldsymbol{\Omega})$ be defined as in Theorem 7.2. Define $\widehat{\boldsymbol{\Psi}}_{ij}(\boldsymbol{\Omega})$ analogously, but using rank transforms instead of asymptotic rank transforms (see Definition 7.1). Then,*

$$D = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)} \operatorname{tr}(\widehat{\boldsymbol{\Psi}}_{ij}(\boldsymbol{\Omega})) - \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)} \operatorname{tr}(\widetilde{\boldsymbol{\Psi}}_{ij}(\boldsymbol{\Omega})) = o_p(1),$$

as $a \to \infty$.

Proof. Without loss of generality, assume that $\boldsymbol{\Omega} = \mathbf{I}$. Define

$$\tilde{D} = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)} \frac{1}{4c_{ij}}$$

$$\sum_{(k_1,k_2,k_3,k_4) \in \mathscr{K}}^{n_{ij}} \left[(\hat{\mathbf{Y}}_{ijk_1} - \hat{\mathbf{Y}}_{ijk_2})(\hat{\mathbf{Y}}_{ijk_1} - \hat{\mathbf{Y}}_{ijk_2})' \otimes (\hat{\mathbf{Y}}_{ijk_3} - \hat{\mathbf{Y}}_{ijk_4})(\hat{\mathbf{Y}}_{ijk_3} - \hat{\mathbf{Y}}_{ijk_4})' \right.$$

$$\left. - (\mathbf{Y}_{ijk_1} - \mathbf{Y}_{ijk_2})(\mathbf{Y}_{ijk_1} - \mathbf{Y}_{ijk_2})' \otimes (\mathbf{Y}_{ijk_3} - \mathbf{Y}_{ijk_4})(\mathbf{Y}_{ijk_3} - \mathbf{Y}_{ijk_4})' \right],$$

where the c_{ij} are as defined in Theorem 7.2, and consider an arbitrary element of this $p^2 \times p^2$ matrix. Each element $\tilde{D}_{q_1,q_2,q_3,q_4}$ is uniquely determined by a combination of four indices q_1, q_2, q_3, q_4, where $q_r = 1, \ldots, p$, $r = 1, \ldots, 4$. Then, we have

$$\tilde{D}_{q_1,q_2,q_3,q_4} = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)} \frac{1}{4c_{ij}} \sum_{(k_1,k_2,k_3,k_4) \in \mathscr{K}}^{n_{ij}}$$

$$\left[(\hat{Y}_{ijk_1}^{(q_1)} - \hat{Y}_{ijk_2}^{(q_1)})(\hat{Y}_{ijk_1}^{(q_2)} - \hat{Y}_{ijk_2}^{(q_2)})(\hat{Y}_{ijk_3}^{(q_3)} - \hat{Y}_{ijk_4}^{(q_3)})(\hat{Y}_{ijk_3}^{(q_4)} - \hat{Y}_{ijk_4}^{(q_4)}) \right.$$

$$\left. - (Y_{ijk_1}^{(q_1)} - Y_{ijk_2}^{(q_1)})(Y_{ijk_1}^{(q_2)} - Y_{ijk_2}^{(q_2)})(Y_{ijk_3}^{(q_3)} - Y_{ijk_4}^{(q_3)})(Y_{ijk_3}^{(q_4)} - Y_{ijk_4}^{(q_4)}) \right]$$

$$= \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)4c_{ij}} \sum_{(k_1,k_2,k_3,k_4)\in\mathscr{K}}^{n_{ij}}$$

$$\left[(\hat{H}^{(q_1)}(X_{ijk_1}^{(q_1)}) - \hat{H}^{(q_1)}(X_{ijk_2}^{(q_1)}))(\hat{H}^{(q_2)}(X_{ijk_1}^{(q_2)}) - \hat{H}^{(q_2)}(X_{ijk_2}^{(q_2)})) \right.$$

$$\times (\hat{H}^{(q_3)}(X_{ijk_3}^{(q_3)}) - \hat{H}^{(q_3)}(X_{ijk_4}^{(q_3)}))(\hat{H}^{(q_4)}(X_{ijk_3}^{(q_4)}) - \hat{H}^{(q_4)}(X_{ijk_4}^{(q_4)}))$$

$$- (H^{(q_1)}(X_{ijk_1}^{(q_1)}) - H^{(q_1)}(X_{ijk_2}^{(q_1)}))(H^{(q_2)}(X_{ijk_1}^{(q_2)}) - H^{(q_2)}(X_{ijk_2}^{(q_2)}))$$

$$\left. \times (H^{(q_3)}(X_{ijk_3}^{(q_3)}) - H^{(q_3)}(X_{ijk_4}^{(q_3)}))(H^{(q_4)}(X_{ijk_3}^{(q_4)}) - H^{(q_4)}(X_{ijk_3}^{(q_4)})) \right]$$

$$= \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{1}{n_{ij}(n_{ij}-1)4c_{ij}} \sum_{(k_1,k_2,k_3,k_4)\in\mathscr{K}}^{n_{ij}} \frac{1}{N^4} \sum_{s_1=1}^{N} \sum_{s_2=1}^{N} \sum_{s_3=1}^{N} \sum_{s_4=1}^{N}$$

$$\zeta(X_{ijk_1}^{(q_1)}, X_{ijk_2}^{(q_1)}, X_{ijk_1}^{(q_2)}, X_{ijk_2}^{(q_2)}, X_{ijk_3}^{(q_3)}, X_{ijk_4}^{(q_3)}, X_{ijk_3}^{(q_4)}, X_{ijk_4}^{(q_4)}, X_{s_1}, X_{s_2}, X_{s_3}, X_{s_4}),$$

where $\zeta(X_{ijk_1}^{(q_1)}, X_{ijk_2}^{(q_1)}, X_{ijk_1}^{(q_2)}, X_{ijk_2}^{(q_2)}, X_{ijk_3}^{(q_3)}, X_{ijk_4}^{(q_3)}, X_{ijk_3}^{(q_4)}, X_{ijk_4}^{(q_4)}, X_{s_1}, X_{s_2}, X_{s_3}, X_{s_4})$

$$= \left([c(X_{ijk_1}^{(q_1)} - X_{s_1}) - c(X_{ijk_2}^{(q_1)} - X_{s_1})][c(X_{ijk_1}^{(q_2)} - X_{s_2}) - c(X_{ijk_2}^{(q_2)} - X_{s_2})] \right.$$

$$\times [c(X_{ijk_3}^{(q_3)} - X_{s_3}) - c(X_{ijk_4}^{(q_3)} - X_{s_3})][c(X_{ijk_3}^{(q_4)} - X_{s_4}) - c(X_{ijk_4}^{(q_4)} - X_{s_4})]$$

$$- [F_{s_1}(X_{ijk_1}^{(q_1)}) - F_{s_1}(X_{ijk_2}^{(q_1)})][F_{s_2}(X_{ijk_1}^{(q_2)}) - F_{s_2}(X_{ijk_2}^{(q_2)})]$$

$$\left. \times [F_{s_3}(X_{ijk_3}^{(q_3)}) - F_{s_3}(X_{ijk_4}^{(q_3)})][F_{s_4}(X_{ijk_3}^{(q_4)}) - F_{s_4}(X_{ijk_4}^{(q_4)})] \right),$$

$c(\cdot)$ denotes the normalized counting function $c(x) = (I\{x > 0\} + I\{x \geq 0\})$,

and F_t denotes the cdf of X_t.

Note that $E(\zeta) = 0$ if all indices s_1, s_2, s_3, s_4 are different from each other, and the corresponding random variables independent of the other eight random variables. This holds because the first part of ζ, integrated over $(X_{s_1}, X_{s_2}, X_{s_3}, X_{s_4})$, equals the second part. Therefore, $E(\tilde{D}) \to 0$ since the number of (s_1, s_2, s_3, s_4) index combinations resulting in nonzero expectation is of order N^3, but the sum is divided by N^4. Consider now

$$\tilde{D}^2_{q_1,q_2,q_3,q_4} = \frac{1}{a^2 b^2} \sum_{i_1=1}^{a} \sum_{i_2=1}^{a} \sum_{j_1=1}^{b} \sum_{j_2=1}^{b} \frac{1}{n_{i_1 j_1} n_{i_2 j_2} (n_{i_1 j_1}-1)(n_{i_2 j_2}-1)16 c_{i_1 j_1} c_{i_2 j_2}}$$

$$\sum_{(k_1,k_2,k_3,k_4)\in\mathscr{K}}^{n_{i_1 j_1}} \sum_{(l_1,l_2,l_3,l_4)\in\mathscr{K}}^{n_{i_2 j_2}} \frac{1}{N^8} \sum_{s_1=1}^{N} \sum_{s_2=1}^{N} \sum_{s_3=1}^{N} \sum_{s_4=1}^{N} \sum_{t_1=1}^{N} \sum_{t_2=1}^{N} \sum_{t_3=1}^{N} \sum_{t_4=1}^{N}$$

$$\zeta(X_{i_1 j_1 k_1}^{(q_1)}, X_{i_1 j_1 k_2}^{(q_1)}, X_{i_1 j_1 k_1}^{(q_2)}, X_{i_1 j_1 k_2}^{(q_2)}, X_{i_1 j_1 k_3}^{(q_3)}, X_{i_1 j_1 k_4}^{(q_3)}, X_{i_1 j_1 k_3}^{(q_4)}, X_{i_1 j_1 k_4}^{(q_4)}, X_{s_1}, X_{s_2}, X_{s_3}, X_{s_4})$$

$$\times \zeta(X_{i_2 j_2 l_1}^{(q_1)}, X_{i_2 j_2 l_2}^{(q_1)}, X_{i_2 j_2 l_1}^{(q_2)}, X_{i_2 j_2 l_2}^{(q_2)}, X_{i_2 j_2 l_3}^{(q_3)}, X_{i_2 j_2 l_4}^{(q_3)}, X_{i_2 j_2 l_3}^{(q_4)}, X_{i_2 j_2 l_4}^{(q_4)}, X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4}).$$

Again, when all involved random variables with indices $s_1, s_2, s_3, s_4, t_1, t_2, t_3, t_4$ are independent of each other, and of the remaining random variables, the expectation of each ζ-function is zero, and therefore also the expectation of the product. Similar to above, this can be seen by first integrating over the random variables with indices (s_1, s_2, s_3, s_4), conditional on those with indices (k_1, k_2, k_3, k_4). The number of cases with nonzero expectation is again of smaller order, in this case N^7, while division is by N^8. It follows that $E(\tilde{D}^2_{q_1,q_2,q_3,q_4}) \to 0$ and therefore $\tilde{D}_{q_1,q_2,q_3,q_4} = o_p(1)$ for each element of $\tilde{D}$, which proves $\tilde{D} = o_p(1)$.

Theorem 7.4. *Let $\tilde{\mathbf{S}}_{ij}$ be defined as in Theorem 7.2, and define $\widehat{\mathbf{S}}_{ij}$ analogously, but using rank transforms instead of asymptotic rank transforms. Then,*

$$K = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\frac{1}{n_{ij}n_{ij'}}\mathrm{tr}(\boldsymbol{\Omega}\widehat{\mathbf{S}}_{ij}\boldsymbol{\Omega}\widehat{\mathbf{S}}_{ij'}) - \frac{1}{ab}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\frac{1}{n_{ij}n_{ij'}}\mathrm{tr}(\boldsymbol{\Omega}\tilde{\mathbf{S}}_{ij}\boldsymbol{\Omega}\tilde{\mathbf{S}}_{ij'}) = o_p(1),$$

as $a \to \infty$.

Proof. As in the proof of Theorem 7.3, assume without loss of generality that $\boldsymbol{\Omega} = \mathbf{I}$, and define

$$\tilde{\mathbf{K}} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\frac{1}{n_{ij}(1-n_{ij})n_{ij'}(1-n_{ij'})}\sum_{k=1}^{n_{ij}}\sum_{k'=1}^{n_{ij'}}$$

$$\left[(\hat{\mathbf{Y}}_{ijk} - \hat{\bar{\mathbf{Y}}}_{ij\cdot})(\hat{\mathbf{Y}}_{ijk} - \hat{\bar{\mathbf{Y}}}_{ij\cdot})' \otimes (\hat{\mathbf{Y}}_{ij'k'} - \hat{\bar{\mathbf{Y}}}_{ij'\cdot})(\hat{\mathbf{Y}}_{ij'k'} - \hat{\bar{\mathbf{Y}}}_{ij'\cdot})'\right.$$
$$\left. - (\mathbf{Y}_{ijk} - \bar{\mathbf{Y}}_{ij\cdot})(\mathbf{Y}_{ijk} - \bar{\mathbf{Y}}_{ij\cdot})' \otimes (\mathbf{Y}_{ij'k'} - \bar{\mathbf{Y}}_{ij'\cdot})(\mathbf{Y}_{ij'k'} - \bar{\mathbf{Y}}_{ij'\cdot})'\right],$$

and consider again an arbitrary element $\tilde{K}_{q_1,q_2,q_3,q_4}$ of this $p^2 \times p^2$ matrix that is determined by a combination of four indices q_1, q_2, q_3, q_4, where $q_r = 1, \ldots, p$, $r = 1, \ldots, 4$.

$$\tilde{K}_{q_1,q_2,q_3,q_4} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\frac{1}{n_{ij}(1-n_{ij})n_{ij'}(1-n_{ij'})}\sum_{k=1}^{n_{ij}}\sum_{k'=1}^{n_{ij'}}$$

$$\left[(\hat{Y}^{(q_1)}_{ijk} - \hat{\bar{Y}}^{(q_1)}_{ij\cdot})\hat{Y}^{(q_2)}_{ijk}(\hat{Y}^{(q_3)}_{ij'k'} - \hat{\bar{Y}}^{(q_3)}_{ij'\cdot})\hat{Y}^{(q_4)}_{ij'k'} - (Y^{(q_1)}_{ijk} - \bar{Y}^{(q_1)}_{ij\cdot})Y^{(q_2)}_{ijk}(Y^{(q_3)}_{ij'k'} - \bar{Y}^{(q_3)}_{ij'\cdot})Y^{(q_4)}_{ij'k'}\right]$$

$$= \frac{1}{ab}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\frac{1}{n_{ij}(1-n_{ij})n_{ij'}(1-n_{ij'})}\sum_{k=1}^{n_{ij}}\sum_{k'=1}^{n_{ij'}}$$

$$\left\{\left[\hat{Y}^{(q_1)}_{ijk}\hat{Y}^{(q_2)}_{ijk}\hat{Y}^{(q_3)}_{ij'k'}\hat{Y}^{(q_4)}_{ij'k'} - Y^{(q_1)}_{ijk}Y^{(q_2)}_{ijk}Y^{(q_3)}_{ij'k'}Y^{(q_4)}_{ij'k'}\right]\right.$$

$$- \left[\hat{\bar{Y}}^{(q_1)}_{ij\cdot}\hat{Y}^{(q_2)}_{ijk}(\hat{Y}^{(q_3)}_{ij'k'} - \hat{\bar{Y}}^{(q_3)}_{ij'\cdot})\hat{Y}^{(q_4)}_{ij'k'} - \bar{Y}^{(q_1)}_{ij\cdot}Y^{(q_2)}_{ijk}(Y^{(q_3)}_{ij'k'} - \bar{Y}^{(q_3)}_{ij'\cdot})Y^{(q_4)}_{ij'k'}\right]$$

$$\left. - \left[\hat{Y}^{(q_1)}_{ijk}\hat{Y}^{(q_2)}_{ijk}\hat{\bar{Y}}^{(q_3)}_{ij'\cdot}\hat{Y}^{(q_4)}_{ij'k'} - Y^{(q_1)}_{ijk}Y^{(q_2)}_{ijk}\bar{Y}^{(q_3)}_{ij'\cdot}Y^{(q_4)}_{ij'k'}\right]\right\}.$$

The terms in each of the three square brackets can be considered separately, using basically the same techniques for each. We show details of the proof for the first term.

$$\hat{Y}_{ijk}^{(q_1)}\hat{Y}_{ijk}^{(q_2)}\hat{Y}_{ij'k'}^{(q_3)}\hat{Y}_{ij'k'}^{(q_4)} - Y_{ijk}^{(q_1)}Y_{ijk}^{(q_2)}Y_{ij'k'}^{(q_3)}Y_{ij'k'}^{(q_4)}$$

$$= \hat{H}^{(q_1)}(X_{ijk}^{(q_1)})\hat{H}^{(q_2)}(X_{ijk}^{(q_2)})\hat{H}^{(q_3)}(X_{ij'k'}^{(q_3)})\hat{H}^{(q_4)}(X_{ij'k'}^{(q_4)})$$

$$- H^{(q_1)}(X_{ijk}^{(q_1)})H^{(q_2)}(X_{ijk}^{(q_2)})H^{(q_3)}(X_{ij'k'}^{(q_3)})H^{(q_4)}(X_{ij'k'}^{(q_4)})$$

$$= \frac{1}{N^4}\sum_{s_1=1}^{N}\sum_{s_2=1}^{N}\sum_{s_3=1}^{N}\sum_{s_4=1}^{N}\Big[c(X_{ijk}^{(q_1)} - X_{s_1})c(X_{ijk}^{(q_2)} - X_{s_2})c(X_{ij'k'}^{(q_3)} - X_{s_3})c(X_{ij'k'}^{(q_4)} - X_{s_4})$$

$$- F_{s_1}(X_{ijk}^{(q_1)})F_{s_2}(X_{ijk}^{(q_2)})F_{s_3}(X_{ij'k'}^{(q_3)})F_{s_4}(X_{ij'k'}^{(q_4)})\Big].$$

Clearly, the expected value of this expression is 0 when all indices s_1, s_2, s_3, s_4 are different from each other, and the corresponding random variables independent of the other four random variables. This can be seen by integrating over $(X_{s_1}, X_{s_2}, X_{s_3}, X_{s_4})$ first. The number of (s_1, s_2, s_3, s_4) index combinations resulting in nonzero expectation is of order N^3, while the sum is divided by N^4. Using similar techniques for the remaining components of $\tilde{K}_{q_1,q_2,q_3,q_4}$, it follows that $E(\tilde{K}_{q_1,q_2,q_3,q_4}) \to 0$. Consider next

$$\tilde{K}_{q_1,q_2,q_3,q_4}^2 = \frac{1}{a^2 b^2}\sum_{i_1=1}^{a}\sum_{i_2=1}^{a}\sum_{j_1 \neq j_1'}^{b}\sum_{j_2 \neq j_2'}^{b}$$

$$\frac{1}{n_{i_1 j_1}n_{i_2 j_2}(1 - n_{i_1 j_1})(1 - n_{i_2 j_2})n_{i_1 j_1'}n_{i_2 j_2'}(1 - n_{i_1 j_1'})(1 - n_{i_2 j_2'})}$$

$$\sum_{k_1=1}^{n_{i_1 j_1}}\sum_{k_2=1}^{n_{i_2 j_2}}\sum_{k'=1}^{n_{i_1 j_1'}}\sum_{k_2'=1}^{n_{i_2 j_2'}}\Big[(\hat{Y}_{i_1 j_1 k_1}^{(q_1)} - \hat{\bar{Y}}_{i_1 j_1 \cdot}^{(q_1)})\hat{Y}_{i_1 j_1 k_1}^{(q_2)}(\hat{Y}_{i_1 j_1' k_1'}^{(q_3)} - \hat{\bar{Y}}_{i_1 j_1' \cdot}^{(q_3)})\hat{Y}_{i_1 j_1' k_1'}^{(q_4)}$$

$$- (Y_{i_1 j_1 k_1}^{(q_1)} - \bar{Y}_{i_1 j_1 \cdot}^{(q_1)})Y_{i_1 j_1 k_1}^{(q_2)}(Y_{i_1 j_1' k_1'}^{(q_3)} - \bar{Y}_{i_1 j_1' \cdot}^{(q_3)})Y_{i_1 j_1' k_1'}^{(q_4)}\Big]$$

$$\Big[(\hat{Y}_{i_2 j_2 k_2}^{(q_1)} - \hat{\bar{Y}}_{i_2 j_2 \cdot}^{(q_1)})\hat{Y}_{i_2 j_2 k_2}^{(q_2)}(\hat{Y}_{i_2 j_2' k_2'}^{(q_3)} - \hat{\bar{Y}}_{i_2 j_2' \cdot}^{(q_3)})\hat{Y}_{i_2 j_2' k_2'}^{(q_4)}$$

$$- (Y_{i_2 j_2 k_2}^{(q_1)} - \bar{Y}_{i_2 j_2 \cdot}^{(q_1)})Y_{i_2 j_2 k_2}^{(q_2)}(Y_{i_2 j_2' k_2'}^{(q_3)} - \bar{Y}_{i_2 j_2' \cdot}^{(q_3)})Y_{i_2 j_2' k_2'}^{(q_4)}\Big].$$

The product of the square brackets can be decomposed into the following and similar terms, using the same decomposition as in the first part of this proof.

$$\frac{1}{N^8} \sum_{s_1=1}^{N} \sum_{s_2=1}^{N} \sum_{s_3=1}^{N} \sum_{s_4=1}^{N} \sum_{t_1=1}^{N} \sum_{t_2=1}^{N} \sum_{t_3=1}^{N} \sum_{t_4=1}^{N}$$

$$\left[c(X_{i_1j_1k_1}^{(q_1)} - X_{s_1}) c(X_{i_1j_1k_1}^{(q_2)} - X_{s_2}) c(X_{i_1j_1'k_1'}^{(q_3)} - X_{s_3}) c(X_{i_1j_1'k_1'}^{(q_4)} - X_{s_4}) \right.$$

$$\left. - F_{s_1}(X_{i_1j_1k_1}^{(q_1)}) F_{s_2}(X_{i_1j_1k_1}^{(q_2)}) F_{s_3}(X_{i_1j_1'k_1'}^{(q_3)}) F_{s_4}(X_{i_1j_1'k_1'}^{(q_4)}) \right]$$

$$\left[c(X_{i_2j_2k_2}^{(q_1)} - X_{t_1}) c(X_{i_2j_2k_2}^{(q_2)} - X_{t_2}) c(X_{i_2j_2'k_2'}^{(q_3)} - X_{t_3}) c(X_{i_2j_2'k_2'}^{(q_4)} - X_{t_4}) \right.$$

$$\left. - F_{t_1}(X_{i_2j_2k_2}^{(q_1)}) F_{t_2}(X_{i_2j_2k_2}^{(q_2)}) F_{t_3}(X_{i_2j_2'k_2'}^{(q_3)}) F_{t_4}(X_{i_2j_2'k_2'}^{(q_4)}) \right].$$

As above, it can be seen that when all involved random variables with indices $s_1, s_2, s_3, s_4, t_1, t_2, t_3, t_4$ are independent of each other, and of the remaining random variables, the expectation of this expression is zero. The number of cases with nonzero expectation is of order N^7, while division is by N^8. A tedious calculation verifies that this is also the case for the remaining components of $\tilde{K}^2_{q_1,q_2,q_3,q_4}$. Thus, $E(\tilde{K}^2_{q_1,q_2,q_3,q_4}) \to 0$ and $\tilde{K}_{q_1,q_2,q_3,q_4} = o_p(1)$ for each element of $\tilde{K}$, proving $\tilde{K} = o_p(1)$.

The three previous theorems together establish the consistency of a rank-based estimator of the asymptotic variances. Aggregating the results so far, we can take advantage of the results from Harrar and Bathke (2012) and formulate Theorem 7.5.

Theorem 7.5. *Let* $\psi = AB, A, A|B,$ *or* $B|A$. *Under the hypothesis* $\mathscr{H}_0^{(\psi)}$, $\sqrt{a}\, \mathrm{tr}(\mathbf{H}^{(\psi)} - \mathbf{G})\hat{\boldsymbol{\Omega}}\,\hat{\tau}_\psi^{-1}(\hat{\boldsymbol{\Omega}}) \xrightarrow{\mathscr{L}} N(0,1)$ *as* $a \to \infty$ *and* n_{ij} *and* b *bounded, where* $\hat{\boldsymbol{\Omega}}$ *is the consistent estimator of* $\boldsymbol{\Omega}$ *obtained by replacing* $-\boldsymbol{\Sigma}$ *with* $N^{-2}\mathbf{G}$ *(see Proposition 7.2), and where*

$$\hat{\tau}_\psi^2(\hat{\boldsymbol{\Omega}}) = \begin{cases} \frac{2}{b}\left\{\hat{v}_1(\hat{\boldsymbol{\Omega}}) + \frac{\hat{v}_2(\hat{\boldsymbol{\Omega}})}{(b-1)^2}\right\} & when \quad \psi = AB, \\ \frac{2}{b}\left\{\hat{v}_1(\hat{\boldsymbol{\Omega}}) + \hat{v}_2(\hat{\boldsymbol{\Omega}})\right\} & when \quad \psi = A, \\ \frac{2}{b}\hat{v}_1(\hat{\boldsymbol{\Omega}}) & when \quad \psi = A|B, \\ \frac{2}{b^2}\left\{\hat{v}_1(\hat{\boldsymbol{\Omega}}) + \frac{\hat{v}_2(\hat{\boldsymbol{\Omega}})}{(b-1)^2}\right\} & when \quad \psi = B|A. \end{cases}$$

Here, $\hat{v}_1(\hat{\boldsymbol{\Omega}}) = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\mathrm{tr}(\widehat{\boldsymbol{\Psi}}_{ij}(\hat{\boldsymbol{\Omega}}))}{n_{ij}(n_{ij}-1)}$ *and* $\hat{v}_2(\hat{\boldsymbol{\Omega}}) = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j\neq j'}^{b}\frac{\mathrm{tr}(\hat{\boldsymbol{\Omega}}\hat{\mathbf{S}}_{ij}\hat{\boldsymbol{\Omega}}\hat{\mathbf{S}}_{ij'})}{n_{ij}n_{ij'}}$.

7.5 Simulation Study

In order to investigate the finite sample performance of the proposed inference methods under the exemplary setting of dimension $p = 3$, number of levels of factor A between $a = 6$ and $a = 50$, number of levels of factor B set to $b = 3$, sample sizes
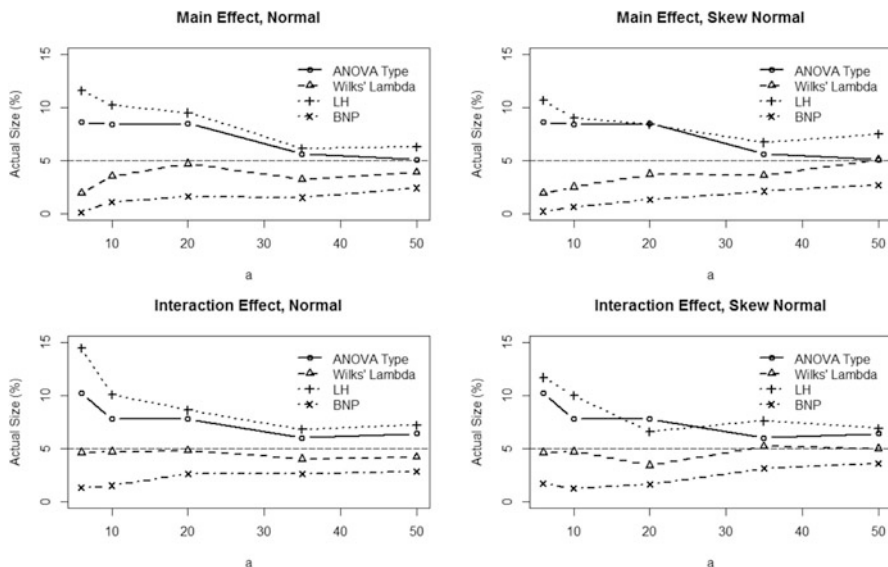
Fig. 7.1 Simulated α under null hypothesis for $p = 3$, $b = 3$, $a = 6$ to 50, $n_{ij} = 4, 5, 6$. Normal and skew normal underlying data distributions, nominal α 5 %. Main effect of A and interaction between A and B

per cell between $n_{ij} = 4$, 5, and 6. Underlying distributions chosen were normal and skew normal. The multivariate skew normal data were generated according to Proposition 6 in Azzalini and Dalla Valle (1996) where we used (in their notation) $\boldsymbol{\delta} = \frac{\sqrt{2}}{\sqrt{p(p+1)+2}}(1, \ldots, p)'$ and $\Omega = I_p + \frac{1}{2\pi}\boldsymbol{\delta}\boldsymbol{\delta}'$.

The results under null hypothesis are shown in Fig. 7.1. As expected from a fully nonparametric rank-based approach, the underlying distribution does not have a major effect on the performance. In all cases considered, Wilks' Λ type statistic performed best, in the sense of the simulated level being closest to the nominal level, while not exceeding it.

Due to its best performance under null hypothesis, Wilks' Λ type test statistic was selected for a power simulation. Here, the statistic based on variablewise ranks, as proposed in the present article, was compared to the power of the analogous procedure using the original observations instead of the ranks (justified by Harrar and Bathke 2012). While there were no visible differences for underlying normal distributions, the power gain of the nonparametric rank-based method became quite pronounced when the underlying distribution was chosen as contaminated normal. Figure 7.2 shows simulation results for an exemplary situation with heteroscedastic contaminated multivariate normal distributions $0.9N_3(\mathbf{0}, \boldsymbol{\Sigma}_{ij}) + 0.1N_3(10 \cdot \mathbf{1}, \boldsymbol{\Sigma}_{ij})$. Here, $\boldsymbol{\Sigma}_{ij}$ were different compound symmetric variance-covariance matrices with off-diagonal elements $\rho_{ij} = \sqrt{ij}/(1 + ij)$ and diagonal elements $1 - \rho_{ij}$.

Alternatives were modeled by location shifts. Specifically, in the main effects power simulation, expected values were shifted up by δ units for levels 10–20 of
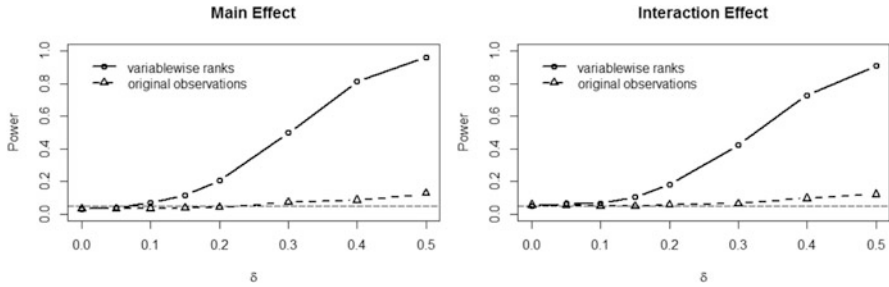
Fig. 7.2 Simulated power of Wilks' Λ type test statistic using ranks, and using raw data, $p = 3$, $a = 20$, $b = 3$, $n_{ij} = 4, 5, 6$. Contaminated normal underlying data distributions, nominal α 5 %. Main effect of A and interaction between A and B. Location shifts for main and interaction effects as described in the text

factor A, for all variables, while they were shifted down by δ units for the other levels 1–9. In the interaction effects power simulation, the upwards shift was for the factor level combinations (i, j) with $i \geq 10, j \geq 2$, whereas the downwards shift was for $i < 10, j < 2$. In both cases, $a = 20$, $b = 3$, $p = 3$.

The results show the rather striking advantages of a nonparametric rank-based approach over its semiparametric competitor using the original observations instead of ranks.

7.6 Discussions and Conclusions

In this somewhat theoretical manuscript, we have introduced fully nonparametric, rank-based test statistics for inference on multivariate data in factorial designs. To our knowledge, no comparable results in such general applicability (for example for fully ordinal data) have been established yet. Due to the rather cumbersome technicalities, the work has only been carried out here for a design with two factors, but it can be extended in a straightforward way to higher-way layouts. Also, we have focused here on large (a) asymptotics (number of factor levels of factor A tends to infinity) and only considered those test statistics in detail that yield asymptotic normality under this type of asymptotic setting. The asymptotic distribution of the test for main effect of factor B will be that of a weighted sum of χ^2 random variables.

It should be pointed out that the test statistics can be calculated directly, they don't involve any iterative computational procedures. The test statistics presented here can be taken as a basis for small sample approximations based on moment estimators or expansions. In future work, it would be interesting to compare their performance with resampling based methods such as those from Konietschke et al. (2015), or with other robust procedures based on semiparametric models.

References

Akritas, M., & Arnold S. (2000). Asymptotics for analysis of variance when the number of levels is large. *Journal of the American Statistical association, 95*(449), 212–226.

Azzalini, A., & Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika, 83*(4), 715–726.

Bathke, A. C. (2002). Anova for a large number of treatments. *Mathematical Methods of Statistics, 11*(1), 118–132.

Bathke, A. C. (2004). The anova F test can still be used in some balanced designs with unequal variances and nonnormal data. *Journal of Statistical Planning and Inference, 126*(2), 413–422.

Bathke, A. C., & Harrar, S.W. (2008). Nonparametric methods in multivariate factorial designs for large number of factor levels. *Journal of Statistical Planning and Inference, 138*(3), 588–610.

Belloni, A., & Didier, G. (2008). On the Behrens–Fisher problem: A globally convergent algorithm and a finite-sample study of the Wald, LR and LM tests. *The Annals of Statistics, 36*, 2377–2408.

Boos, D.D., & Brownie, C. (1995). Anova and rank tests when the number of treatments is large. *Statistics & Probability Letters, 23*(2), 183–191.

Brunner, E., Domhof, S., & Langer, F. (2002). *Nonparametric analysis of longitudinal data in factorial experiments*. New York: Wiley.

Chatfield, J. A., Draper, E. A., Cochran, K. D., & Herms, D. A. (2000). Evaluation of crabapples for apple scab at the Secrest Arboretum in Wooster, Ohio. *Special circular 177*, The Ohio State University, Ohio Agricultural Research and Development Center.

Girón, F. J., & del Castillo, C. (2010). The multivariate Behrens–Fisher distribution. *Journal of Multivariate Analysis, 101*(9), 2091–2102.

Gupta, A. K., Harrar, S. W., & Fujikoshi, Y. (2006). Asymptotics for testing hypothesis in some multivariate variance components model under non-normality. *Journal of Multivariate Analysis, 97*, 148–178.

Gupta, A. K., Harrar, S. W., & Fujikoshi, Y. (2008). Manova for large hypothesis degrees of freedom under non-normality. *Test, 17*(1), 120–137.

Harrar, S. W., & Bathke, A. C. (2008). Nonparametric methods for unbalanced multivariate data and many factor levels. *Journal of Multivariate Analysis, 99*(8), 1635–1664.

Harrar, S. W., & Bathke, A. C. (2012). A modified two-factor multivariate analysis of variance: asymptotics and small sample approximations (and erratum). *Annals of the Institute of Statistical Mathematics, 64*(1 & 5), 135–165 & 1087.

Harrar, S. W., & Gupta, A. K. (2007) Asymptotic expansion for the null distribution of the f-statistic in one-way anova under non-normality. *Annals of the Institute of Statistical Mathematics, 59*(3), 531–556.

Harville, D. A. (2008). *Matrix algebra from a statistician's perspective*. Berlin: Springer.

Kawasaki, T., & Seo, T. (2015). A two sample test for mean vectors with unequal covariance matrices. *Communications in Statistics-Simulation and Computation, 44*, 1850–1866.

Konietschke, F., Bathke, A. C., Harrar, S. W., & Pauly, M. (2015). Parametric and nonparametric bootstrap methods for general MANOVA. *Journal of Multivariate Analysis, 140*, 291–301.

Krishnamoorthy, K., & Lu, F. (2010). A parametric bootstrap solution to the MANOVA under heteroscedasticity. *Journal of Statistical Computation and Simulation, 80*(8), 873–887.

Krishnamoorthy, K., & Yu, J. (2004). Modified Nel and Van der Merwe test for the multivariate Behrens–Fisher problem. *Statistics & Probability Letters, 66*(2), 161–169.

Krishnamoorthy, K., & Yu, J. (2012). Multivariate Behrens–Fisher problem with missing data. *Journal of Multivariate Analysis, 105*(1), 141–150.

Kruskal, W. H. (1952). A nonparametric test for the several sample problem. *Annals of Mathematical Statistics, 23*, 525–540.

Lévy, P. (1925). *Calcul des Probabilités*. Paris: Gauthiers-Villars.

Nel, D.G., & Van der Merwe, C. A. (1986). A solution to the multivariate Behrens-Fisher problem. *Communications in Statistics-Theory and Methods, 15*(12), 3719–3735.

Penrose, R. (1955). A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society, 51*, 17–19.

Ruymgaart, F. H. (1980). A unified approach to the asymptotic distribution theory of certain midrank statistics. In *Statistique non parametrique asymptotique*. *Lecture notes on mathematics* (Vol. 821). Berlin: Springer.

Schott, J. R. (2005). *Matrix analysis for statistics*. Hoboken: Wiley.

Serfling, R. (1980). *Approximation theorems of mathematical statistics*. New York: Wiley.

Van Aelst, S., & Willems, G. (2011). Robust and efficient one-way MANOVA tests. *Journal of the American Statistical Association, 106*(494), 706–718.

Zhang, J. -T. (2011). Two-way MANOVA with unequal cell sizes and unequal cell covariance matrices. *Technometrics, 53*(4), 426–439.

Zhang, J. -T. (2012). An approximate Hotelling T^2-test for heteroscedastic one-way MANOVA. *Open Journal of Statistics, 2*(1), 1–11.

Zhang, J. -T., & Liu, X. (2013). A modified Bartlett test for heteroscedastic one-way MANOVA. *Metrika, 76*(1), 135–152.