

Quadratic Form based Multiple Contrast Tests for Comparison of Group Means

Paavo Sattler¹, Markus Pauly^{1,2} and Merle Munko³

¹ Department of Statistics, TU Dortmund University, Joseph-von-Fraunhofer-Straße 2-4, 44227 Dortmund, Germany

email: paavo.sattler@tu-dortmund.de

² Research Center Trustworthy Data Science and Security, UA Ruhr, Joseph-von-Fraunhofer-Straße 25, 44227 Dortmund, Germany

³ Department of Mathematics, Otto-von-Guericke University Magdeburg, Germany

Abstract

Comparing the mean vectors across different groups is a cornerstone in the realm of multivariate statistics, with quadratic forms commonly serving as test statistics. However, when the overall hypothesis is rejected, identifying specific vector components or determining the groups among which differences exist requires additional investigations. Conversely, employing multiple contrast tests (MCT) allows conclusions about which components or groups contribute to these differences. However, they come with a trade-off, as MCT lose some benefits inherent to quadratic forms. In this paper, we combine both approaches to get a quadratic form based multiple contrast test that leverages the advantages of both. To understand its theoretical properties, we investigate its asymptotic distribution in a semiparametric model. We thereby focus on two common quadratic forms — the Wald-type statistic and the Anova-type statistic — although our findings are applicable to any quadratic form.

Furthermore, we employ Monte-Carlo and resampling techniques to enhance the test's performance in small sample scenarios. Through an extensive simulation study, we assess the performance of our proposed tests against existing alternatives, highlighting their advantages.

Keywords: Bootstrap, MANOVA, Monte Carlo techniques, Multiple Testing, Resampling.

1 Motivation and introduction

In multivariate analysis, comparing the mean vectors of two or more groups is one of the most common null hypothesis, also known as one-way MANOVA. As traditional methods like Hotelling's T^2 or Wilk's Lambda tests [Johnson et al., 2002, e.g.] rely on (too) restrictive assumptions like normality and covariance equality, the last two decades have seen the development of alternative approaches to address these limitations. For example, Krishnamoorthy and Xia [2006] or Smaga [2017] proposed methods for the two sample case while Konietschke et al. [2015], Roy et al. [2015], Smaga [2015], Hu et al. [2017], Friedrich and Pauly [2018] or Sattler and Pauly [2018] studied one-way and more general hypotheses in complex MANOVA designs.

An appropriate way to investigate hypotheses regarding vector-valued parameters, like expectation vectors, are quadratic forms. This already holds for the common Hotelling's T^2 and Wilk's Lambda tests but also applies for most of the generalizations mentioned above as well as other multivariate hypotheses, e.g. in completely nonparametric settings Brunner et al. [2017, 2019], Dobler et al. [2020], Rubarth et al. [2022]. In these papers, only one or a few specific quadratic forms are examined. More recently, Sattler [2021] and Baumeister et al. [2024] also allow for a broader range of general quadratic forms.

One of the key strengths of quadratic forms as test statistics is that they yield a univariate value out of multivariate input, but this also introduces a significant limitation: When the hypothesis of equal mean vectors is rejected, this univariate test statistic does not indicate where the differences in expectation vectors lie. As a result, this often necessitates further pairwise comparisons with multiplicity adjustments, a process that can be notably inefficient. In contrast, so-called multiple contrast tests (MCT) offer a compelling approach by not only detecting differences in parameters but, at the same time, draw conclusions where the differences occur. MCT usually use a maximum of univariate t-type statistics as test statistics and equicoordinate quantiles from multivariate normal, t- or resampling distributions as critical values, see, e.g., Hasler and Hothorn [2008], Konietschke et al. [2012], Konietschke et al. [2013], Gunawardana and Konietschke [2019], Umlauf et al. [2019], Noguchi et al. [2020] for some examples. However, for multivariate settings, the component-wise dependence is only taken into account through the estimated correlation of these t-test statistics, and not within the test statistics themselves. Thus, it would be beneficial to integrate quadratic forms with MCT as this combination could yield a test procedure that leverages the advantages of both.

The aim of the present paper is to propose such quadratic form based MCT. Thereby, the procedures will allow general quadratic forms and specifically address the question where differences occur should the global hypothesis be rejected. For ease of presentation, we will initially focus on the one-way hypothesis of equal group expectation vectors with subsequent identification of responsible components. However, we later explain how the method can be extended to more general settings and hypotheses. We examine the large sample properties of all procedures by means of asymptotic theory. Here, it turns out that the commonly used equicoordinate quantiles for MCT with t-type statistics will not lead to appropriate critical values for MCT based on quadratic forms, making adoptions of other approaches necessary Bühlmann [1998], Munko et al. [2023a,b]. The small sample performance of the resulting tests will be investigated in Monte Carlo simulations.

The paper is organized as follows: In the subsequent section, the statistical model will be introduced, together with the underlying assumptions. Afterwards, focusing on the equality of expectation vectors, the basic concept of multiple contrast tests and the test statistics are proposed,

and the asymptotic distributions are derived (Section 3). Subsequently, this is generalized for general partitions of global hypotheses without concentrating on the responsible components for rejection. In Section 5, a resampling strategy is used to generate quantiles with the aim to improve our tests' small sample behaviour. This is demonstrated in Section 6, which contains simulation results on type-I-error control and power for different alternatives, followed by an illustrative application for a real EEG-data set (Section 7). In Section 8 we finally draw conclusions and mention potential future applications. All proofs are deferred to a technical supplement, where additional details and simulations can be found.

2 Model

We consider a general semiparametric model to achieve a broad applicability by independent d -dimensional random vectors

$$\mathbf{X}_{ik} = \boldsymbol{\mu}_i + \boldsymbol{\epsilon}_{ik}.$$

Here, the index $k = 1, \dots, n_i$ represents the subject on which d -variate observations are measured while the index $i = 1, \dots, a$ refers to the respective group. Moreover, $\boldsymbol{\mu}_i$ denotes the expectation vector in group i while $\boldsymbol{\epsilon}_{ik} = (\epsilon_{ikj})_{j=1}^d$ is the corresponding random error of subject k in that group. Thereby we consider $a \geq 2$ groups and assume that for fixed $i \in \{1, \dots, a\}$ and $j \in \{1, \dots, d\}$ the random variables ϵ_{ikj} are identical distributed with $\mathbb{E}(\epsilon_{ikj}) = 0$. It is important to note that by splitting up the indices, this setting implies factorial designs inside a group as in Konietzschke et al. [2015].

We allow unbalanced sample sizes n_1, \dots, n_a and denote with $N = \sum_{i=1}^a n_i$ the total sample size. For our asymptotic investigations, we employ two additional assumptions:

$$(A1) \quad \frac{n_i}{N} \rightarrow \kappa_i \in (0, 1], \quad i = 1, \dots, a$$

$$(A2) \quad \boldsymbol{\Sigma}_i = \text{Var}(\boldsymbol{\epsilon}_{ik}) \geq 0, \quad i = 1, \dots, a.$$

Here, (A1) is the common converging group size assumption that ensures that no group is negligible, while assumption (A2) guarantees the existence of second moments.

To test hypotheses about mean vectors, we use different kinds of pooled mean vectors, a group-wise mean vector $\bar{\mathbf{X}}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} \mathbf{X}_{ik}$ with corresponding $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1^\top, \dots, \bar{\mathbf{X}}_a^\top)^\top$ and a component-wise mean vector

$$\bar{\mathbf{X}}^{(j)} = \left(\frac{1}{n_1} \sum_{k=1}^{n_1} X_{1kj}, \dots, \frac{1}{n_a} \sum_{k=1}^{n_a} X_{akj} \right)^\top.$$

Assumptions (A1) and (A2) guarantee that the normalized version of the latter $\sqrt{N} \bar{\mathbf{X}}^{(j)}$ possess the asymptotic block-diagonal covariance matrix $\boldsymbol{\Sigma}^{(j)} := \text{diag}(\kappa_1^{-1}(\boldsymbol{\Sigma}_1)_{jj}, \dots, \kappa_a^{-1}(\boldsymbol{\Sigma}_a)_{jj}) = \lim_{N \rightarrow \infty} \text{Var}(\sqrt{N} \bar{\mathbf{X}}^{(j)}) \geq 0$ for each $j = 1, \dots, d$.

In the following, we use $\xrightarrow{\mathcal{P}}$ to denote convergence in probability and $\xrightarrow{\mathcal{D}}$ for convergence in distribution, both as $N \rightarrow \infty$.

3 Equality of mean vectors

We first consider the one-way MANOVA hypothesis

$$\mathcal{H}_0 : (\boldsymbol{\mu}_{1j})_{j=1}^d = \cdots = (\boldsymbol{\mu}_{aj})_{j=1}^d.$$

We refer to it as *global hypothesis* since it is the intersection of the local hypotheses $\mathcal{H}_0^{(j)} : \mu_{1j} = \cdots = \mu_{aj}$. With the Kronecker-product and the centering matrix $\mathbf{P}_a = \mathbf{I}_a - \mathbf{1}_a \mathbf{1}_a^\top / a$, this hypothesis can be formulated using the hypothesis matrix $\mathbf{C} = \mathbf{P}_a \otimes \mathbf{I}_d$ and $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \dots, \boldsymbol{\mu}_a^\top)^\top$ through $\mathbf{C}\boldsymbol{\mu} = \mathbf{0}_{ad}$. Thereby, \mathbf{I}_a denotes the $a \times a$ identity matrix, $\mathbf{1}_a$ is an a -dimensional vector of 1s and the superscript \top indicates a transposed matrix or vector.

The mean vector is a consistent estimator for the expectation vector and with (A1) and (A2) it follows from the multivariate central limit theorem that $\mathbf{T} := \sqrt{N} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \stackrel{\mathcal{D}}{\rightarrow} \mathbf{Z} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} := \bigoplus_{i=1}^a \kappa_i^{-1} \boldsymbol{\Sigma}_i$ while \bigoplus denotes the direct sum (i.e. it defines the block diagonal matrix consisting of the entries). Therefore, under the null hypothesis \mathcal{H}_0 it holds that $\sqrt{N} \mathbf{C} \bar{\mathbf{X}} \stackrel{\mathcal{D}}{\rightarrow} \mathbf{C} \mathbf{Z} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \mathbf{C} \boldsymbol{\Sigma} \mathbf{C}^\top)$ allowing to consider test statistics in quadratic forms. Common examples cover the Wald-Type-Statistic

$$WTS(\boldsymbol{\Sigma}) := N \cdot (\mathbf{C} \bar{\mathbf{X}})^\top (\mathbf{C} \boldsymbol{\Sigma} \mathbf{C}^\top)^+ (\mathbf{C} \bar{\mathbf{X}})$$

using a Moore-Penrose inverse (denoted by the superscript $+$) and the Anova-Type-Statistic

$$ATS(\boldsymbol{\Sigma}) := N \cdot (\mathbf{C} \bar{\mathbf{X}})^\top (\mathbf{C} \bar{\mathbf{X}}) / \text{tr}(\mathbf{C} \boldsymbol{\Sigma} \mathbf{C}^\top)$$

as well as their empirical counterparts, where the unknown covariance matrix is substituted by a consistent estimator $\hat{\boldsymbol{\Sigma}}$. Since in our setting, the empirical covariance matrices for the i -th group, $\hat{\boldsymbol{\Sigma}}_i$ are consistent estimators for $\boldsymbol{\Sigma}_i$, $\hat{\boldsymbol{\Sigma}} = N \cdot \bigoplus_{i=1}^a n_i^{-1} \hat{\boldsymbol{\Sigma}}_i$ is a suitable choice. Then, we obtain asymptotic correct level α tests for the global hypothesis by using, for example, Monte-Carlo-based quantiles for these statistics. However, in case of a rejection, it would often be necessary to conduct further investigations to find the responsible components. A solution could be given through multiple contrast tests, where we will outline the general idea for $\boldsymbol{\Sigma} > \mathbf{0}$ and use the hypothesis matrix $\mathbf{C} = [\mathbf{c}_1^\top, \dots, \mathbf{c}_L^\top]^\top$. Here, \mathbf{C} can be for example the *Tukey*-type-contrast matrix for parameter $a \cdot d$, see e.g. Munko et al. [2023b], as in the simulations in Section 6, where each row of \mathbf{D}_{ad} compares two components of $\boldsymbol{\mu}$ and builds one of ad local hypotheses.

With $\tilde{\mathbf{T}}_\ell = \mathbf{c}_\ell \mathbf{T} / \sqrt{\mathbf{c}_\ell \boldsymbol{\Sigma} \mathbf{c}_\ell^\top}$, a standardized version of $\mathbf{c}_\ell \mathbf{T}$, the ℓ -th local hypothesis is rejected if $|\tilde{\mathbf{T}}_\ell| > q_\alpha$, where q_α is so called equicoordinate quantile depending on the covariance matrix $\mathbf{C} \boldsymbol{\Sigma} \mathbf{C}^\top$. The global hypothesis is rejected if and only if at least one local hypothesis is rejected. Here, standardization of \mathbf{T} is necessary to justify the usage of the same quantile for all components, which allows formulating the global test through $\mathbf{1}\{\max(|\tilde{\mathbf{T}}_1|, \dots, |\tilde{\mathbf{T}}_{ad}|) > q_\alpha\}$. The Tukey-type matrix with its local hypotheses would here allow deducting between which groups differences can be verified and in which components.

For situations where the test statistic can not be standardized or the standardization is not sufficient to have the same distribution in each component, there also exists a generalization of an MCT, see Munko et al. [2023a] and Munko et al. [2023b], which are based on an approach from Bretz et al. [2001].

Based on this concept, we want to construct a test for the global hypothesis using general quadratic forms (see e.g. Sattler et al. [2022]). To this aim, based on a matrix-value function $\widetilde{M} \in \mathbb{R}^{a \times a}$ with two arguments we define for our j -th local hypothesis from Section 2

$$\widetilde{Q}_{Nj} = N \cdot \frac{[\mathbf{P}_a \overline{\mathbf{X}}^{(j)}]^\top \widetilde{M}(\mathbf{P}_a, \widetilde{\boldsymbol{\Sigma}}^{(j)}) [\mathbf{P}_a \overline{\mathbf{X}}^{(j)}]}{\sqrt{2 \operatorname{tr}([\mathbf{P}_a^\top \widetilde{M}(\mathbf{P}_a, \widetilde{\boldsymbol{\Sigma}}^{(j)}) \mathbf{P}_a \widetilde{\boldsymbol{\Sigma}}^{(j)}]^2)}} \cdot \mathbb{1} \left(\operatorname{tr} \left([\mathbf{P}_a^\top \widetilde{M}(\mathbf{P}_a, \widetilde{\boldsymbol{\Sigma}}^{(j)}) \mathbf{P}_a \widetilde{\boldsymbol{\Sigma}}^{(j)}]^2 \right) > 0 \right)$$

with $0/0 := 1$.

Necessary conditions are $\widetilde{M}^\top = \widetilde{M}$, $\widetilde{M} \geq 0$ and $\widetilde{M}(\mathbf{P}_a, \widetilde{\boldsymbol{\Sigma}}^{(j)}) \xrightarrow{\mathcal{P}} \widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)})$ for an estimator $\widetilde{\boldsymbol{\Sigma}}^{(j)} \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}^{(j)}$, where a possible choice would be $\widetilde{\boldsymbol{\Sigma}}^{(j)} = N \cdot \operatorname{diag}(n_1^{-1}(\widehat{\boldsymbol{\Sigma}}_1)_{jj}, \dots, n_a^{-1}(\widehat{\boldsymbol{\Sigma}}_a)_{jj})$. Common functions are $\widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) := (\mathbf{P}_a \boldsymbol{\Sigma}^{(j)} \mathbf{P}_a^\top)^+$ for the WTS, or $\widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) := \mathbf{I}_a / \operatorname{tr}(\mathbf{P}_a \boldsymbol{\Sigma}^{(j)} \mathbf{P}_a^\top)$ for the ATS but also many other quadratic forms like the modified Anova-Type-statistic (MATS) from Friedrich and Pauly [2017] or different ATS versions [Brunner et al., 2019].

Lemma 3.1:

Let $\widetilde{v}_{jj} = 2 \operatorname{tr}([\mathbf{P}_a^\top \widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) \mathbf{P}_a \boldsymbol{\Sigma}^{(j)}]^2) > 0$. Then, under the j th local hypothesis, the following convergence in distribution holds

$$\widetilde{Q}_{Nj} \xrightarrow{\mathcal{D}} \widetilde{Q}_j \stackrel{\mathcal{D}}{=} \frac{1}{\sqrt{\widetilde{v}_{jj}}} \sum_{i=1}^a \lambda_i \Upsilon_i$$

with $\Upsilon_1, \dots, \Upsilon_a \stackrel{i.i.d.}{\sim} \chi_1^2$ and $\lambda_1, \dots, \lambda_a \in \operatorname{eigen} \left(\boldsymbol{\Sigma}^{(j)1/2} \mathbf{P}_a^\top \widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) \mathbf{P}_a \boldsymbol{\Sigma}^{(j)1/2} \right)$.

In addition, we have for $\mathbf{Z} = (Z_{11}, Z_{12}, \dots, Z_{1d}, Z_{21}, \dots, Z_{ad})^\top \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \boldsymbol{\Sigma})$ under the global null hypothesis \mathcal{H}_0 convergence in distribution

$$\widetilde{\mathbf{Q}}_N = (\widetilde{Q}_{N1}, \dots, \widetilde{Q}_{Nd})^\top \xrightarrow{\mathcal{D}} \widetilde{\mathbf{Q}} = (\widetilde{Q}_1, \dots, \widetilde{Q}_d)^\top := \left(\widetilde{v}_{jj}^{-1/2} \cdot [\mathbf{P}_a \mathbf{Z}^{(j)}]^\top \widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) \mathbf{P}_a \mathbf{Z}^{(j)} \right)_{j \in \{1, \dots, d\}}$$

with $\mathbf{Z}^{(j)} := (Z_{1j}, \dots, Z_{aj})^\top$. Moreover, if $\mathcal{T} \subset \{1, \dots, d\}$ denote the indices of true local hypotheses, we have $(\widetilde{Q}_{Nj})_{j \in \mathcal{T}} \xrightarrow{\mathcal{D}} (\widetilde{Q}_j)_{j \in \mathcal{T}}$.

From Mathai and Provost [1992] it follows $\operatorname{Var}(\widetilde{Q}_j) = 1$ if $\widetilde{v}_{jj} > 0$ which we in the following assume for $j = 1, \dots, d$. This is for example given if $\mathbf{P}_a^\top \widetilde{M}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) \mathbf{P}_a \neq \mathbf{0}_{a \times a}$ and $\operatorname{Var}(X_{ij1}) > 0$ for each $i = 1, \dots, a$ resulting in $\boldsymbol{\Sigma}^{(j)} > 0$.

However, since different components from the same group are allowed to depend on each other, the covariance matrix of this vector, called \mathbf{R} , is not necessarily diagonal, while you find the concrete form in the supplement.

Remark 3.2:

In the case of a WTS, the distribution of the components is quite simple through $\widetilde{Q}_j \sim \chi_{a-1}^2$. But in

Dickhaus and Royen [2015], it was shown with simple and reasonable examples, that quite different distributions can lead to this kind of marginal distribution even with the same dependence structure \mathbf{R} . The appropriate distribution of the vector $\tilde{\mathbf{Q}}$, called multivariate chi-square distribution of generalized Wishart-type, is also given therein, together with the fact that it is not possible to calculate the corresponding quantiles in general.

Nevertheless, it is reasonable to use the standardized vector $\tilde{\mathbf{Q}}$, and it is important to mention that here the marginal distributions are known, but usually the distribution of $\tilde{\mathbf{Q}}$ is not.

Remark 3.3:

In contrast to the multivariate normal distribution of $\tilde{\mathbf{T}}$ although $\tilde{\mathbf{Q}}$ is standardized, and therefore all components have the same variance, this does not mean that all components follow the same distribution. As shown, the distribution, in general, depends on a matrix's eigenvalues, while the variance only depends on the trace of the square of this matrix. For this reason, it is essential to use the approach from Munko et al. [2023a] to get adequate quantiles. Using equicoordinate quantiles also leads to asymptotic correct tests, but this can affect the ability to detect derivations from the null hypothesis, and therefore its power.

Using empirical quantiles based on a Monte-Carlo technique is a convenient approach to get the required quantiles of the limit distribution. A detailed procedure description can be found in the appendix, but we sketch it shortly. First $\mathbf{Z}_1^{MC} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \hat{\Sigma})$ is generated, and since $\tilde{\mathbf{Q}}$ can be seen as a function of such random vectors (see appendix for more details) we calculate a $\tilde{\mathbf{Q}}^{MC,1}$ based on $\mathbf{Z}_1^{MC}/\sqrt{N}$. If we repeat this frequently (e.g. $B=10,000$ times) we receive $\tilde{\mathbf{Q}}^{MC} = (\tilde{Q}^{MC,1}, \dots, \tilde{Q}^{MC,B})$ which allows us to calculate the needed quantiles.

Lemma 3.4:

For $\alpha \in (0, 1)$, let $\tilde{q}_{1,B}^{MC}, \dots, \tilde{q}_{d,B}^{MC}$ be the corresponding quantiles of $\tilde{\mathbf{Q}}^{MC}$ according to Munko et al. [2023a]. Then

$$\varphi_{\tilde{\mathbf{Q}}}^{MC}(N, B) := \max_{j=1, \dots, d} \mathbb{1} \left\{ \left(\frac{\tilde{Q}_{Nj}}{\tilde{q}_{j,B}^{MC}} \right) > 1 \right\}$$

is an asymptotic correct level α test for the global hypothesis $\mathcal{H}_0 : \boldsymbol{\mu}_1 = \dots = \boldsymbol{\mu}_a$, i.e., for all sequences $B_N \rightarrow \infty$ as $N \rightarrow \infty$, we have

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\varphi_{\tilde{\mathbf{Q}}}^{MC}(N, B_N) \right] = \alpha$$

under the global null hypothesis. Moreover, the test controls the family-wise error rate asymptotically in the strong sense, which means

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\max_{j \in \mathcal{T}} \mathbb{1} \left\{ \left(\frac{\tilde{Q}_{Nj}}{\tilde{q}_{j,B_N}^{MC}} \right) > 1 \right\} \right] \leq \alpha$$

for all index subsets $\mathcal{T} \subset \{1, \dots, d\}$ of true local hypotheses.

We define $0/0:=1$, and from the fact which $\tilde{Q}_{N_j}/\tilde{q}_{j,B}^{MC}$ are larger than 1, it follows which local hypotheses are rejected and thus, in which components the groups differ. Also, resampling techniques can be used to calculate appropriate quantiles and, therefore, receive corresponding tests. This will be handled in Section 5.

Remark 3.5:

Especially in the context of resampling approaches, large dimensions or large number of groups the computing effort of an approach becomes more relevant. Here, the quadratic form based multiple contrast test (QFMCT) is often preferable over classical MCT or quadratic forms, while it of course, depends on the chosen quadratic form. For the simplest quadratic form, the ATS without standardization through the trace and using $\tilde{\mathbf{M}}(\mathbf{P}_a, \boldsymbol{\Sigma}^{(j)}) = \mathbf{I}_a$ it is easy to calculate and compare the computational complexity. In the case of the classical multiple contrast tests with a Tukey-type matrix, it is $\mathcal{O}(a^3 d^3)$ and for this ATS, even $\mathcal{O}(a^3 d^3 + a^2 d^2)$. For the quadratic forms based multiple contrast test, of course, d quadratic forms are calculated, but each of them with a substantially smaller vector. This leads to a computational complexity of $\mathcal{O}(da^3 + da^2 + d^2)$, which is considerably smaller. With the results from Sattler and Rosenbaum [2025], this can further be reduced.

For the standardized ATS or the WTS, the quadratic forms based multiple contrast test is even more preferable, since for the required covariance matrix estimation and matrix multiplications, the complexity is growing exponentially with the dimension. This makes this approach attractive from a computational perspective.

In this section, we focus on the question in which component the difference occurs, which is one common interrogation, especially since the number of groups is usually much smaller than the respective dimension. This would be similar possible for the hypothesis $\mathcal{H}_0 : \boldsymbol{\mu} = \boldsymbol{\zeta}, \boldsymbol{\zeta} \in \mathbb{R}^{ad}$. Moreover, it is possible to adapt our approach to find the groups that differ instead of the components. Thereto, for each of the $\binom{a}{2}$ pairwise comparisons between two groups, a quadratic form would be defined based on the group mean vectors and $\mathbf{P}_2 \otimes \mathbf{I}_d$. Now, we generalize the considered hypotheses essentially, such that all mentioned hypotheses are special cases.

4 General Hypotheses

In this section, we consider the general linear hypothesis

$$\mathcal{H}_0 : \mathbf{C}\boldsymbol{\mu} = \boldsymbol{\beta}$$

for a hypothesis matrix $\mathbf{C} \in \mathbb{R}^{r \times ad}$ and corresponding vector $\boldsymbol{\beta} \in \mathbb{R}^r$ for $r \in \mathbb{N}$. The local hypotheses can be obtained by partitioning the hypothesis matrix $\mathbf{C} = \left(\mathbf{C}_1^\top, \dots, \mathbf{C}_L^\top \right)^\top$ into $L \in \mathbb{N}$ block matrices $\mathbf{C}_1 \in \mathbb{R}^{r_1 \times ad}, \dots, \mathbf{C}_L \in \mathbb{R}^{r_L \times ad}$ and the vector $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_L^\top)^\top$ into L vectors $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_L$ with corresponding number of rows r_ℓ as

$$\mathcal{H}_0^{(\ell)} : \mathbf{C}_\ell \boldsymbol{\mu} = \boldsymbol{\beta}_\ell, \quad \ell \in \{1, \dots, L\}.$$

Of course, this contains equality of expectation vectors with a focus on components, as was treated in the previous section, but also with other partitions allowing to determine the respon-

sible groups for a rejection. Now, we want to formulate the analogous tests for this general hypotheses. This leads to quadratic forms for the ℓ -th local hypothesis through

$$Q_{N\ell} = N \cdot \frac{[\mathbf{C}_\ell \bar{\mathbf{X}} - \boldsymbol{\beta}_\ell]^\top \mathbf{M}(\mathbf{C}_\ell, \tilde{\boldsymbol{\Sigma}}) [\mathbf{C}_\ell \bar{\mathbf{X}} - \boldsymbol{\beta}_\ell]}{\sqrt{2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \tilde{\boldsymbol{\Sigma}}) \mathbf{C}_\ell \tilde{\boldsymbol{\Sigma}}]^2)}} \cdot \mathbb{1} \left(\operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \tilde{\boldsymbol{\Sigma}}) \mathbf{C}_\ell \tilde{\boldsymbol{\Sigma}}]^2) > 0 \right).$$

In contrast to $\tilde{\mathbf{M}}$, the dimension of the second argument of \mathbf{M} changes, while all other requirements on the function are still needed.

Lemma 4.1:

Let $v_{\ell\ell} = 2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \boldsymbol{\Sigma}) \mathbf{C}_\ell \boldsymbol{\Sigma}]^2) > 0$. Then under the ℓ -th local hypothesis, we get

$$Q_{N\ell} \stackrel{\mathcal{D}}{\rightarrow} Q_\ell \stackrel{\mathcal{D}}{=} \frac{1}{\sqrt{v_{\ell\ell}}} \sum_{i=1}^{r_\ell} \lambda_i \Upsilon_i$$

with $\Upsilon_1, \dots, \Upsilon_{r_\ell} \stackrel{i.i.d.}{\sim} \chi_1^2$ and $\lambda_1, \dots, \lambda_{r_\ell} \in \operatorname{eigen}(\boldsymbol{\Sigma}^{1/2} \mathbf{C}_\ell^\top \mathbf{M}(\mathbf{P}_a, \boldsymbol{\Sigma}) \mathbf{C}_\ell \boldsymbol{\Sigma}^{1/2})$. In addition, with $\mathbf{Z} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \boldsymbol{\Sigma})$ under the global null hypothesis \mathcal{H}_0 also

$$\mathbf{Q}_N = (Q_{N1}, \dots, Q_{NL})^\top \stackrel{\mathcal{D}}{\rightarrow} \mathbf{Q} = (Q_1, \dots, Q_L)^\top := \left(v_{\ell\ell}^{-1/2} \cdot [\mathbf{C}_\ell \mathbf{Z}]^\top \mathbf{M}(\mathbf{C}_\ell, \boldsymbol{\Sigma}) \mathbf{C}_\ell \mathbf{Z} \right)_{\ell \in \{1, \dots, L\}}.$$

Moreover, if $\mathcal{T} \subset \{1, \dots, L\}$ denote the indices of true local hypotheses, we have $(Q_{N\ell})_{\ell \in \mathcal{T}} \stackrel{\mathcal{D}}{\rightarrow} (Q_\ell)_{\ell \in \mathcal{T}}$.

Furthermore, to fulfill the condition of this Lemma, an additional assumption

(A3) for each $\ell \in \{1, \dots, L\}$, there is a nonzero eigenvalue of $\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \boldsymbol{\Sigma}) \mathbf{C}_\ell \boldsymbol{\Sigma}$

is required. A sufficient condition for (A3) is that $\boldsymbol{\Sigma}$ is positive definite and $\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \boldsymbol{\Sigma}) \mathbf{C}_\ell \neq \mathbf{0}_{ad \times ad}$ for all ℓ . However, this condition is not necessary, as we have seen in Section 3, and the restrictiveness of this condition strongly depends on the chosen \mathbf{M} .

The previous Lemma allows, in a similar way as earlier, to construct Monte-Carlo realizations $\mathbf{Q}^{MC} = (\mathbf{Q}^{MC,1}, \dots, \mathbf{Q}^{MC,B})$ of \mathbf{Q} and based on this an asymptotic correct test.

Lemma 4.2:

For $\alpha \in (0, 1)$, let $q_{1,B}^{MC}, \dots, q_{L,B}^{MC}$ be the corresponding quantiles of \mathbf{Q}^{MC} according to Munko et al. [2023a]. Then

$$\varphi_{\mathbf{Q}}^{MC}(N, B) := \max_{\ell=1, \dots, L} \mathbb{1} \left\{ \left(\frac{Q_{N\ell}}{q_{\ell,B}^{MC}} \right) > 1 \right\}$$

is an asymptotic correct level α test for the global hypothesis $\mathcal{H}_0 : \mathbf{C}\boldsymbol{\mu} = \boldsymbol{\beta}$. Moreover, the test controls the family-wise error rate asymptotically in the strong sense.

5 Resampling Procedures

In the context of multiple contrast tests and as well as quadratic forms, bootstrap techniques lead in many situations to preferable tests, regarding type-I error and power. This holds especially for small sample sizes see, for example, Sattler [2021], Segbehoe et al. [2022] and Baumeister et al. [2024]. For this reason, we introduce two bootstrap approaches, focusing on the parametric bootstrap and handling the wild bootstrap only in the supplement to increase readability.

First, based on the empirical covariance matrix $\widehat{\Sigma}_i$ random vectors $\mathbf{X}_{i1}^*, \dots, \mathbf{X}_{in_i}^* \stackrel{i.i.d.}{\sim} \mathcal{N}_d(\mathbf{0}_d, \widehat{\Sigma}_i)$ which are independent of the realizations are generated for each group $i = 1, \dots, a$.

Note that we still do not postulate a parametric model on our data, although we use the parametric bootstrap, since drawing the bootstrap observations from a normal distribution is motivated by the asymptotic behaviour of \mathbf{T} , cf. Section 4.

Furthermore, let $\overline{\mathbf{X}}^*$ denote the parametric bootstrap counterpart of the mean vector and $\widehat{\Sigma}^* = N \cdot \bigoplus_{i=1}^a n_i^{-1} \widehat{\Sigma}_i^*$, where $\widehat{\Sigma}_i^*$ denotes the empirical covariance matrix of the i th parametric bootstrap sample. Then, the parametric bootstrap quadratic forms are given by

$$Q_{N\ell}^* = N \cdot \frac{[\mathbf{C}_\ell \overline{\mathbf{X}}^*]^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^*) [\mathbf{C}_\ell \overline{\mathbf{X}}^*]}{\sqrt{2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^*) \mathbf{C}_\ell \widehat{\Sigma}^*]^2)}} \cdot \mathbb{1} \left(\operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^*) \mathbf{C}_\ell \widehat{\Sigma}^*]^2) > 0 \right)$$

for $\ell \in \{1, \dots, L\}$.

Example 5.1:

In the situation of Section 3, the parametric bootstrap quadratic forms are

$$\widetilde{Q}_{Nj}^* = N \cdot \frac{[\mathbf{P}_a \overline{\mathbf{X}}^{(j)*}]^\top \widetilde{\mathbf{M}}(\mathbf{P}_a, \widehat{\Sigma}^{(j)*}) [\mathbf{P}_a \overline{\mathbf{X}}^{(j)*}]}{\sqrt{2 \operatorname{tr}([\mathbf{P}_a^\top \widetilde{\mathbf{M}}(\mathbf{P}_a, \widehat{\Sigma}^{(j)*}) \mathbf{P}_a \widehat{\Sigma}^{(j)*}]^2)}} \cdot \mathbb{1} \left(\operatorname{tr}([\mathbf{P}_a^\top \widetilde{\mathbf{M}}(\mathbf{P}_a, \widehat{\Sigma}^{(j)*}) \mathbf{P}_a \widehat{\Sigma}^{(j)*}]^2) > 0 \right)$$

with sample covariances $\widehat{\Sigma}^{(j)*}$ and component-wise mean vectors

$$\overline{\mathbf{X}}^{(j)*} = \left(\frac{1}{n_1} \sum_{k=1}^{n_1} X_{1kj}^*, \dots, \frac{1}{n_a} \sum_{k=1}^{n_a} X_{akj}^* \right).$$

Theorem 5.1:

If Assumptions (A1) and (A2) are fulfilled, it holds: Given the data, the conditional distribution of

- (a) $\sqrt{N} \overline{\mathbf{X}}^*$ converges weakly to $\mathcal{N}_{ad}(\mathbf{0}_{ad}, \Sigma)$ in probability.
- (b) $\widetilde{\mathbf{Q}}_N^* = (\widetilde{Q}_{N1}^*, \dots, \widetilde{Q}_{Nd}^*)^\top$ converges weakly to $\widetilde{\mathbf{Q}}$ in probability.
- (c) $\mathbf{Q}_N^* = (Q_{N1}^*, \dots, Q_{NL}^*)^\top$ converges weakly to \mathbf{Q} in probability, if also (A3) is fulfilled.

Moreover, we have $\widehat{\Sigma}^ \xrightarrow{\mathcal{P}} \Sigma$ and the unknown covariance matrix can be estimated through the parametric bootstrap estimator.*

A large number of repetitions (B) leads to a corresponding number of realizations $\mathbf{Q}_N^{*,1}, \dots, \mathbf{Q}_N^{*,B}$ of \mathbf{Q}_N^* and the possibility to calculate bootstrap quantiles.

Lemma 5.2:

For $\alpha \in (0, 1)$, let $q_{1,B}^*, \dots, q_{L,B}^*$ be the corresponding quantiles of $\mathbf{Q}_N^{*,1}, \dots, \mathbf{Q}_N^{*,B}$ according to Munko et al. [2023a]. Then

$$\varphi_{\mathbf{Q}}^*(N, B) := \mathbb{1} \left\{ \max_{\ell=1, \dots, L} \left(\frac{Q_{N\ell}}{q_{\ell,B}^*} \right) > 1 \right\}$$

is an asymptotic correct level α test for the global hypothesis $\mathcal{H}_0 : \mathbf{C}\boldsymbol{\mu} = \boldsymbol{\beta}$ that controls the family-wise error rate asymptotically in the strong sense.

Remark 5.3:

Since the vector \mathbf{Q}_N is standardized, it is unnecessary and inefficient to calculate the standardized version of the ATS, instead of the version using $\mathbf{M}(\mathbf{C}_\ell, \boldsymbol{\Sigma}) = \mathbf{I}_{r_\ell}$. Also, other versions of the ATS, which only differ by a kind of standardization, therefore lead to the same vector \mathbf{Q}_N .

Example 5.1 (continued):

To check the hypothesis $\mathcal{H}_0 : \boldsymbol{\mu}_1 = \dots = \boldsymbol{\mu}_a$ based on $\tilde{\mathbf{Q}}_N$ from Section 3, the appropriate test would be given through

$$\varphi_{\tilde{\mathbf{Q}}}^*(N, B) := \mathbb{1} \left\{ \max_{j=1, \dots, d} \left(\frac{\tilde{Q}_{Nj}}{\tilde{q}_{j,B}^*} \right) > 1 \right\},$$

where $\tilde{q}_{1,B}^*, \dots, \tilde{q}_{d,B}^*$ are the quantiles calculated using realisations $\tilde{\mathbf{Q}}_N^{*,1}, \dots, \tilde{\mathbf{Q}}_N^{*,B}$ of $\tilde{\mathbf{Q}}_N^*$.

6 Simulations

Since it is easy to interpret and at the same time allows to construct alternatives in a simple and uniform way, we focus here on the hypothesis of equal expectation vectors $\mathcal{H}_0 : \boldsymbol{\mu}_1 = \dots = \boldsymbol{\mu}_a$. From Sattler and Zimmermann [2024], it follows in the case of only two groups, that for the ATS and WTS using a hypothesis matrix with only one row would lead to the same values of the test statistic. Following this, the value of the classical MCT and the QFMCT coincides. Also, this does not influence the previous results, it makes it reasonable to focus here on $a > 2$, which, for computational reasons, let us choose $a = 3$. Moreover, we consider unbalanced group sizes $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$ with $N \in \{25, 50, 100\}$ since especially the performance for small sample sizes is of great interest. We considered 5-dimensional observation vectors generated independently according to the model $\mathbf{X}_{ik} = \boldsymbol{\mu}_i + \boldsymbol{\Sigma}^{1/2} \mathbf{Z}_{ik}$, $i = 1, \dots, a$, $k = 1, \dots, n_i$ and error terms $\mathbf{Z}_{ik} = (Z_{ik1}, \dots, Z_{ik5})^\top$ based on independently generated Z_{ikj} , $i = 1, \dots, a$, $k = 1, \dots, n_i$, $j = 1, \dots, 5$, with the following distributions:

- a standard normal distribution, i.e. $Z_{ikj} \sim \mathcal{N}(0, 1)$.
- a student-t distribution with 9 degrees of freedom, i.e. $\sqrt{7/9} \cdot Z_{ikj} \sim t_9$.

We choose $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$ and $(\Sigma_3)_{j_1 j_2} = 0.65^{|j_1 - j_2|}$ to have not only heterogeneous Σ_i but also $\Sigma^{(j)}$ and under the respective null hypothesis we know for the ATS, $Q_j \sim \sum_{\ell=1}^d \lambda_\ell B_\ell$, with $B_\ell \stackrel{i.i.d.}{\sim} \chi_1^2$ and $\lambda_1, \dots, \lambda_d \in \text{eigen}(\mathbf{P}_a \Sigma^{(j)} \mathbf{P}_a^\top)$. Therefore, despite their standardization, we have different distributions of the quadratic forms, as mentioned in Remark 3.3. For this reason, usage of the same quantile for all components of \tilde{Q} makes less sense, and the approach of Munko et al. [2023a] is required. Besides the QFMCT with parametric bootstrap and Monte-Carlo approach, we simulated the classical ATS with parametric bootstrap (ψ_{ATS}^*). Moreover, we considered the classical MCT using the Tukey-type matrix, with equicoordinate quantiles and based on the parametric bootstrap (φ and φ^*).

Since we are interested in particular in the global power of the tests, we consider a one-point-alternative, given through $\mu_1 = \mu_2 = \mathbf{0}_5$ and $\mu_3 = \delta \cdot (1, 0, 0, 0, 0)^\top$. We print the type-I rate in bold, if it is within the 95% binomial interval [0.0458, 0.0543], and in the supplement also the results for a shift-alternative can be found.

Here, we expect the one-point alternative to be challenging since the deviation from the null hypothesis is only in one component and, therefore, difficult to detect. We use 1,000 bootstrap steps for our parametric bootstrap, 10,000 simulation steps for the Monte-Carlo approach and 10,000 runs for all tests to get reliable results.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	15.29	14.79	17.86	24.58	32.73	45.19	58.62	70.16	81.71
φ^*	25	5.22	5.22	6.04	9.30	15.25	23.62	34.13	45.46	59.29
$\varphi_{\tilde{Q}}^{MC}(ATS)$	25	9.70	9.96	12.00	17.71	25.78	38.09	52.15	66.15	79.64
$\varphi_{\tilde{Q}}^*(ATS)$	25	4.95	5.12	6.61	9.83	16.12	25.09	37.65	51.12	66.53
ψ_{ATS}^*	25	3.95	4.22	5.01	6.03	8.91	12.60	18.18	25.59	36.74
φ	50	8.99	10.75	16.72	27.92	47.47	68.79	84.30	94.06	98.57
φ^*	50	4.75	5.99	10.48	18.77	36.35	57.82	76.53	89.32	96.72
$\varphi_{\tilde{Q}}^{MC}(ATS)$	50	7.09	8.25	14.01	24.78	45.34	68.33	84.28	94.61	98.84
$\varphi_{\tilde{Q}}^*(ATS)$	50	4.93	5.98	10.70	19.74	38.79	62.09	79.81	92.26	97.95
ψ_{ATS}^*	50	4.58	5.47	7.23	10.97	18.31	31.99	48.77	68.94	85.43
φ	100	6.82	10.07	22.86	50.79	78.96	95.23	99.40	99.92	100.00
φ^*	100	5.07	7.65	19.11	45.18	74.80	93.59	99.09	99.89	100.00
$\varphi_{\tilde{Q}}^{MC}(ATS)$	100	6.14	9.28	21.90	50.66	79.76	95.67	99.61	99.94	100.00
$\varphi_{\tilde{Q}}^*(ATS)$	100	5.22	8.13	19.90	47.41	77.55	95.10	99.43	99.94	100.00
ψ_{ATS}^*	100	4.91	6.01	10.58	22.93	45.67	74.25	92.92	99.06	99.97

Table 1: Power of different test statistics under an one-point-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on the standard normal distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell - k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

In Table 1 it can be seen that the type-I error rate for the bootstrap version of the MC-test and QFMCT is comparably good while the bootstrap ATS has a worse small sample approximation, but overall still good values. The power under the one-point-alternative is continually higher

for the QFMCT, which, especially for the small sample size, leads to differences of about 0.06. In Table 2, the difference in power between bootstrap multiple contrast tests is smaller, while again, the approach based on the ATS is favourable. Moreover, for the skew normal distribution it has a clear better type-I error rate than the classical bootstrap.

For both distributions, it is noticeable that the classical ATS with the same bootstrap approach has not only a worse type-I error rate under the null hypothesis but especially a substantially smaller power than the MC-tests. For all sample sizes, it is partwise half the value of the QFMCT with parametric bootstrap. Therefore, it is clearly worse at detecting deviation from the null hypothesis.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	14.69	14.29	18.29	24.31	33.35	44.86	59.28	71.07	81.07
φ^*	25	4.76	3.94	6.41	9.36	14.91	23.36	35.41	47.65	60.53
$\varphi_{\tilde{Q}}^{MC}(ATS)$	25	9.89	9.12	12.29	17.44	25.53	36.52	52.35	65.95	77.82
$\varphi_{\tilde{Q}}^*(ATS)$	25	4.93	4.51	6.39	9.43	15.69	24.66	37.94	51.70	64.85
ψ_{ATS}^*	25	4.27	4.02	5.12	6.26	8.15	12.18	18.79	25.39	36.62
φ	50	8.54	10.41	16.41	29.80	47.79	68.69	84.92	94.03	98.14
φ^*	50	4.59	5.84	10.07	20.55	36.86	58.22	76.71	89.99	96.00
$\varphi_{\tilde{Q}}^{MC}(ATS)$	50	6.78	8.21	13.39	25.99	45.09	66.57	84.42	94.04	98.25
$\varphi_{\tilde{Q}}^*(ATS)$	50	4.62	5.98	10.08	20.58	38.32	60.39	79.69	91.66	97.28
ψ_{ATS}^*	50	4.71	5.10	6.85	10.92	18.45	30.56	49.70	68.70	84.96
φ	100	6.89	9.49	22.89	50.62	79.53	95.05	99.27	99.91	100.00
φ^*	100	5.07	7.25	18.73	45.11	75.46	93.42	98.90	99.85	99.99
$\varphi_{\tilde{Q}}^{MC}(ATS)$	100	6.30	8.57	21.72	49.82	79.90	95.50	99.38	99.97	100.00
$\varphi_{\tilde{Q}}^*(ATS)$	100	5.14	7.32	19.27	46.90	77.69	94.58	99.24	99.93	100.00
ψ_{ATS}^*	100	5.16	6.12	10.58	22.26	46.21	74.33	92.99	99.05	99.92

Table 2: Power of different test statistics under an one-point-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a t_9 distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

Finally, the non-bootstrap approaches in this simulation need larger sample sizes to have good results. Here again, the QFMCT is clearly favourable. Since the algorithm to calculate the equicoordinate quantiles is quite demanding, it is even faster, although it needs 10,000 simulations runs.

7 Data Example

To illustrate the application of the method, the EEG data set from the R-package *manova.rm* by Friedrich et al. [2019] is considered closer. This study from Staffen et al. [2014] was conducted at the University Clinic of Salzburg (Department of Neurology), where from 160 patients, electroencephalography (EEG) data were measured. All participants are diagnosed with different diagnoses of impairments, namely Alzheimer’s disease (AD), mild cognitive impairment (MCI), and subjective cognitive complaints (SCC). In Table 3, the number of patients divided by sex and

diagnosis can be found.

Table 3: Number of observations for the different factor level combinations of sex and diagnosis.

	AD	MCI	SCC
male	12	27	20
female	24	30	47

MANOVA-based comparisons were already made in Bathke et al. [2018]), and we now want to complete this with our new approach. The original study also further differentiated between subjective cognitive complaints with minimal cognitive dysfunction (SCC+) and without (SCC-). However, because of sample sizes, this was not done in the data set provided through the package.

For each participant, three different electrode positions (frontal, temporal, and central) were used together with two kinds of measurements (z-score for brain rate and Hjorth complexity), resulting in observation vector's dimension $d = 6$. In Sattler et al. [2022], homogeneity of covariance matrices between different diagnoses as well as different sexes were investigated for this data set. Therefore, we have six groups with heterogeneous covariance matrices and unbalanced sample sizes, making many test procedures incapable.

In Bathke et al. [2018]), equality of expectation vectors was rejected, and therefore an influence of the diagnosis was proven. However, it was never identified where the differences occur. Since all previous analyses of this data set could neither verify an influence of the location nor measurement, identifying the responsible component is, here, not the main issue. Therefore, we will focus on identifying the responsible groups and determine three local hypotheses

$$\mathcal{H}_0^{(1)} : \mu_{AD} = \mu_{MCI} \quad \mathcal{H}_0^{(2)} : \mu_{AD} = \mu_{SCC} \quad \mathcal{H}_0^{(3)} : \mu_{MCI} = \mu_{SCC}$$

which together build the global hypothesis $\mathcal{H}_0 : \mu_{AC} = \mu_{MCI} = \mu_{SCC}$. Since we have vectors with dimension 6, each of these local hypotheses consists of 6 sub-hypotheses, one for each component of the vector. So for the classical MCT we have 18 single hypotheses, in partitions of 6 for each local hypothesis. The test is conducted with 5.000 bootstrap runs and for both genders separately.

	p-values of φ^* for 6 sub-hypotheses						p-value of $\varphi_Q^*(ATS)$
$\mathcal{H}_0^{(1)}$ -Male	74.40	75.06	69.92	68.54	88.30	65.64	31.04
$\mathcal{H}_0^{(2)}$ -Male	1.88	3.08	1.50	8.04	30.62	24.30	0.98
$\mathcal{H}_0^{(3)}$ -Male	1.32	1.32	<0.01	1.14	27.78	0.30	0.04
$\mathcal{H}_0^{(1)}$ -Female	97.96	100.00	99.38	78.38	99.96	88.46	75.54
$\mathcal{H}_0^{(2)}$ -Female	1.46	9.42	3.34	3.12	39.28	7.88	0.38
$\mathcal{H}_0^{(3)}$ -Female	2.72	2.92	1.82	9.62	57.52	7.02	0.38

Table 4: Adjusted p-values in percent for QFMTC with ATS and classical MCT, both based on parametric bootstrap and investigating local hypotheses $\mathcal{H}_0^{(1)} : \mu_{AD} = \mu_{MCI}$, $\mathcal{H}_0^{(2)} : \mu_{AD} = \mu_{SCC}$ and $\mathcal{H}_0^{(3)} : \mu_{MCI} = \mu_{SCC}$.

In Table 4, the results can be found, where it contains for each local hypothesis 6 adjusted p-

values when considering φ^* and only one for $\varphi_Q^*(ATS)$. It is not surprising that for both genders, the global hypotheses can be rejected for the global level $\alpha = 5\%$ since this was also a result of Bathke et al. [2018]). Furthermore, all multiple contrast tests determine the differences between AD and MCI as well as MCI and SCC. The classic MCT can also determine the components responsible for this and reject $\mathcal{H}_0^{(3)}$ for men also to the smaller global level $\alpha = 1\%$. In contrast, all 4 rejections of QFMCT hold for this level, again showing QFMCT's advantage by identifying even relatively small deviations from the null hypothesis.

8 Conclusion

In the presented work, we present an approach combining two useful and common methods, the multiple contrast test and quadratic forms, to get the best from both. This leads to a quadratic form based multiple contrast test, defined for quite general quadratic forms. With an appropriate bootstrap approach, this results in a consistent test to compare the group means. Although the approach is introduced for equality of expectations with a focus on the components, it is generalized for a large class of hypotheses later on. The conducted simulation shows that although the type-I error is comparable to the bootstrap version of the classical multiple contrast test, it is clearly better regarding the power. Since it usually also needs fewer computations, it is therefore overall preferable for general linear hypotheses regarding expectation vectors in multiple groups, in particular for large numbers of groups.

We focused here on the group expectation vectors as parameters of interest since they are important and often used. However, the technique of a quadratic form based multiple contrast test can be used for a variety of parameter vectors like quantiles, relative effects, vectorized covariance matrices and many more. So, in future work, we will investigate them and existing limit theorems to develop corresponding QFMCTs.

Acknowledgements

Merle Munko gratefully acknowledges the support of the Deutsche Forschungsgemeinschaft (grant no. GRK 2297 MathCoRe).

References

- Bathke, A. C., Friedrich, S., Pauly, M., Konietzschke, F., Staffen, W., Strobl, N., and Höller, Y. (2018). Testing mean differences among groups: Multivariate and repeated measures analysis with minimal assumptions. *Multivariate Behavioral Research*, 53:348–359.
- Baumeister, M., Ditzhaus, M., and Pauly, M. (2024). Quantile-based manova: A new tool for inferring multivariate data in factorial designs. *Journal of Multivariate Analysis*, 199:105246.
- Bretz, F., Genz, A., and Hothorn, L. A. (2001). On the numerical availability of multiple comparison procedures. *Biometrical Journal*, 43(5):645–656.
- Brunner, E., Bathke, A. C., and Konietzschke, F. (2019). *Rank and Pseudo-Rank Procedures for Independent Observations in Factorial Designs*. Springer.

- Brunner, E., Konietschke, F., Pauly, M., and Puri, M. L. (2017). Rank-based procedures in factorial designs: hypotheses about non-parametric treatment effects. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(5):1463–1485.
- Bühlmann, P. (1998). Sieve bootstrap for smoothing in nonstationary time series. *The Annals of Statistics*, 26(1):48–83.
- Dickhaus, T. and Royen, T. (2015). A survey on multivariate chi-square distributions and their applications in testing multiple hypotheses. *Statistics*, 49.
- Dobler, D., Friedrich, S., and Pauly, M. (2020). Nonparametric manova in meaningful effects. *Annals of the Institute of Statistical Mathematics*, 72(4):997–1022.
- Friedrich, S., Konietschke, F., and Pauly, M. (2019). *MANOVA.RM: Analysis of Multivariate Data and Repeated Measures Designs*. R package version 0.3.2.
- Friedrich, S. and Pauly, M. (2017). Mats: Inference for potentially singular and heteroscedastic manova. *Journal of Multivariate Analysis*, 165:166–179.
- Friedrich, S. and Pauly, M. (2018). Mats: Inference for potentially singular and heteroscedastic manova. *Journal of Multivariate Analysis*, 165:166–179.
- Gunawardana, A. and Konietschke, F. (2019). Nonparametric multiple contrast tests for general multivariate factorial designs. *Journal of Multivariate Analysis*, 173:165–180.
- Hasler, M. and Hothorn, L. A. (2008). Multiple contrast tests in the presence of heteroscedasticity. *Biometrical Journal*, 50(5):793–800.
- Hu, J., Bai, Z., Wang, C., and Wang, W. (2017). On testing the equality of high dimensional mean vectors with unequal covariance matrices. *Annals of the Institute of Statistical Mathematics*, 69:1572–9052.
- Johnson, R. A., Wichern, D. W., et al. (2002). *Applied multivariate statistical analysis*.
- Konietschke, F., Bathke, A. C., Harrar, S. W., and Pauly, M. (2015). Parametric and nonparametric bootstrap methods for general manova. *Journal of Multivariate Analysis*, 140:291–301.
- Konietschke, F., Bösigler, S., Brunner, E., and Hothorn, L. A. (2013). Are multiple contrast tests superior to the anova? *The International Journal of Biostatistics*, 9(1):63–73.
- Konietschke, F., Hothorn, L. A., and Brunner, E. (2012). Rank-based multiple test procedures and simultaneous confidence intervals. *Electronic Journal of Statistics*, 6:738–759.
- Krishnamoorthy, K. and Xia, Y. (2006). On selecting tests for equality of two normal mean vectors. *Multivariate Behavioral Research*, 41(4):533–548. PMID: 26794917.
- Mammen, E. (1993). Bootstrap and Wild Bootstrap for High Dimensional Linear Models. *The Annals of Statistics*, 21(1):255 – 285.
- Mathai, A. M. and Provost, S. B. (1992). Quadratic forms in random variables, vol. 126 of statistics: Textbooks and monographs.
- Munko, M., Ditzhaus, M., Dobler, D., and Genuneit, J. (2023a). Rmst-based multiple contrast tests in general factorial designs.

- Munko, M., Ditzhaus, M., Pauly, M., Smaga, Ł., and Zhang, J.-T. (2023b). General multiple tests for functional data.
- Munko, M., Dobler, D., and Ditzhaus, M. (2024). Multiple tests for restricted mean time lost with competing risks data.
- Noguchi, K., Abel, R. S., Marmolejo-Ramos, F., and Konietzschke, F. (2020). Nonparametric multiple comparisons. *Behavior Research Methods*, 52:489–502.
- Roy, A., Leiva, R., Žežula, I., and Klein, D. (2015). Testing the equality of mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup. *Journal of Multivariate Analysis*, 137:50–60.
- Rubarth, K., Sattler, P., Zimmermann, H. G., and Konietzschke, F. (2022). Estimation and testing of wilcoxon-mann-whitney effects in factorial clustered data designs. *Symmetry*, 14(2).
- Sattler, P. (2021). A comprehensive treatment of quadratic-form-based inference in repeated measures designs under diverse asymptotics. *Electronic Journal of Statistics*, 15(1):3611 – 3634.
- Sattler, P., Bathke, A. C., and Pauly, M. (2022). Testing hypotheses about covariance matrices in general manova designs. *Journal of Statistical Planning and Inference*, 219:134–146.
- Sattler, P. and Pauly, M. (2018). Inference for high-dimensional split-plot-designs: A unified approach for small to large numbers of factor levels. *Electron. J. Statist.*, 12(2):2743–2805.
- Sattler, P. and Rosenbaum, M. (2025). Choice of the hypothesis matrix for using the anova-type-statistic. *arXiv2409.12592*.
- Sattler, P. and Zimmermann, G. (2024). Choice of the hypothesis matrix for using the wald-type-statistic. *Statistics & Probability Letters*, page 110038.
- Segbehoe, L. S., Schaarschmidt, F., and Djira, G. D. (2022). Simultaneous confidence intervals for contrasts of quantiles. *Biometrical Journal*, 64(1):7–19.
- Smaga, Ł. (2015). Wald-type statistics using $\{2\}$ -inverses for hypothesis testing in general factorial designs. *Statistics & Probability Letters*, 107:215–220.
- Smaga, Ł. (2017). A two-sample test based on cluster subspaces for equality of mean vectors in high dimension. *Discussiones Mathematicae: Probability & Statistics*, 37.
- Staffen, W., Strobl, N., Zauner, H., Höller, Y., Dobesberger, J., and Trinka, E. (2014). Combining spect and eeg analysis for assessment of disorders with amnesic symptoms to enhance accuracy in early diagnostics. *Poster A19 Presented at the 11th Annual Meeting of the Austrian Society of Neurology, 26–29 March, Salzburg, Austria*.
- Umlauft, M., Placzek, M., Konietzschke, F., and Pauly, M. (2019). Wild bootstrapping rank-based procedures: Multiple testing in nonparametric factorial repeated measures designs. *Journal of Multivariate Analysis*, 171:176–192.

A Proofs

We will here only conduct the proofs for the results for Section 4 and Section 5, since the results from Section 3 can be seen as a special case thereof. This follows since there exist a matrix $\mathbf{A}_j \in \mathbb{R}^{a \times ad}$ fulfilling $\overline{\mathbf{X}}^{(j)} = \mathbf{A}_j \overline{\mathbf{X}}$. Hence, by setting $\mathbf{C}_j = \mathbf{P}_a \mathbf{A}_j$ and $\beta_j = \mathbf{0}_a$, all statements from Section 3 are direct consequences from Section 4.

Proof of Lemma 4.1: With the central limit theorem it holds $\sqrt{N}(\overline{\mathbf{X}} - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} \mathbf{Z} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \boldsymbol{\Sigma})$. Let $\mathcal{T} \subset \{1, \dots, L\}$ denote the index subset of true local null hypotheses. From the consistency of $\tilde{\boldsymbol{\Sigma}}$ it follows $\hat{v}_{\ell\ell} = 2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \tilde{\boldsymbol{\Sigma}}) \mathbf{C}_\ell \tilde{\boldsymbol{\Sigma}}]^2) \xrightarrow{\mathcal{P}} v_{\ell\ell}$. With continuous mapping and Slutsky's theorem, therefore

$$\sqrt{N} \cdot \left(\hat{v}_{\ell\ell}^{-1/2} \cdot \mathbf{C}_\ell \overline{\mathbf{X}} - \beta_\ell, \mathbf{M}(\mathbf{C}_\ell, \hat{\boldsymbol{\Sigma}})(\mathbf{C}_\ell \overline{\mathbf{X}} - \beta_\ell) \right)_{\ell \in \mathcal{T}} \xrightarrow{\mathcal{D}} \left(v_{\ell\ell}^{-1/2} \cdot \mathbf{C}_\ell \mathbf{Z}, \mathbf{M}(\mathbf{C}_\ell, \boldsymbol{\Sigma}) \mathbf{C}_\ell \mathbf{Z} \right)_{\ell \in \mathcal{T}} =: \boldsymbol{\Upsilon}.$$

Now consider the continuous function

$$\mathbf{f} : \mathbb{R}^{2 \sum_{\ell \in \mathcal{T}} r_\ell} \rightarrow \mathbb{R}^{|\mathcal{T}|}, (\mathbf{x}_1, \mathbf{y}_1, \dots, \mathbf{x}_d, \mathbf{y}_d) \mapsto (\mathbf{y}_1^\top \mathbf{x}_1, \dots, \mathbf{y}_d^\top \mathbf{x}_d)^\top.$$

Together with the continuous mapping theorem, it holds

$$\begin{aligned} (Q_{N\ell})_{\ell \in \mathcal{T}} &= N \cdot \left(\hat{v}_{\ell\ell}^{-1/2} \cdot (\mathbf{C}_\ell \overline{\mathbf{X}} - \beta_\ell)^\top \mathbf{M}(\mathbf{C}_\ell, \hat{\boldsymbol{\Sigma}})(\mathbf{C}_\ell \overline{\mathbf{X}} - \beta_\ell) \right)_{\ell \in \mathcal{T}} \\ &= \mathbf{f} \left(\sqrt{N} \cdot \left(\hat{v}_{\ell\ell}^{-1/2} \cdot (\mathbf{C}_\ell \overline{\mathbf{X}} - \beta_\ell), \mathbf{M}(\mathbf{C}_\ell, \hat{\boldsymbol{\Sigma}})(\mathbf{C}_\ell \overline{\mathbf{X}} - \beta_\ell) \right)_{\ell \in \mathcal{T}} \right) \xrightarrow{\mathcal{D}} \mathbf{f}(\boldsymbol{\Upsilon}) = (Q_\ell)_{\ell \in \mathcal{T}}. \end{aligned}$$

The distribution of the components then follows by the representation theorem for quadratic forms. □

With $\sqrt{N}(\overline{\mathbf{X}} - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} \mathbf{Z} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \boldsymbol{\Sigma})$ it holds

$$\begin{aligned} (Q_{N\ell_1}, Q_{N\ell_2}) &\xrightarrow{\mathcal{D}} ([\mathbf{C}_{\ell_1} \mathbf{Z}]^\top \mathbf{M}(\mathbf{C}_{\ell_1}, \boldsymbol{\Sigma})[\mathbf{C}_{\ell_1} \mathbf{Z}], [\mathbf{C}_{\ell_2} \mathbf{Z}]^\top \mathbf{M}(\mathbf{C}_{\ell_2}, \boldsymbol{\Sigma})[\mathbf{C}_{\ell_2} \mathbf{Z}]) \cdot (v_{\ell_1 \ell_1} v_{\ell_2 \ell_2})^{-1/2} \\ &= \left(\mathbf{Z}^\top \mathbf{C}_{\ell_1}^\top \mathbf{M}(\mathbf{C}_{\ell_1}, \boldsymbol{\Sigma}) \mathbf{C}_{\ell_1} \mathbf{Z}, \mathbf{Z}^\top \mathbf{C}_{\ell_2}^\top \mathbf{M}(\mathbf{C}_{\ell_2}, \boldsymbol{\Sigma}) \mathbf{C}_{\ell_2} \mathbf{Z} \right) \cdot (v_{\ell_1 \ell_1} v_{\ell_2 \ell_2})^{-1/2} \end{aligned}$$

if the local null hypotheses $\mathcal{H}_0^{(\ell_1)}$ and $\mathcal{H}_0^{(\ell_2)}$ are true, which fulfills the situation from Theorem 3.2d.4 in Mathai and Provost [1992]. Therefore we know

$$\mathbf{R}_{\ell_1, \ell_2} := \operatorname{Cov}(Q_{\ell_1}, Q_{\ell_2}) = (v_{\ell_1 \ell_1} v_{\ell_2 \ell_2})^{-1/2} \cdot 2 \operatorname{tr} \left(\mathbf{C}_{\ell_2}^\top \mathbf{M}(\mathbf{C}_{\ell_2}, \boldsymbol{\Sigma}) \mathbf{C}_{\ell_2} \boldsymbol{\Sigma} \mathbf{C}_{\ell_1}^\top \mathbf{M}(\mathbf{C}_{\ell_1}, \boldsymbol{\Sigma}) \mathbf{C}_{\ell_1} \boldsymbol{\Sigma} \right).$$

Monte-Carlo-approach For the Monte-Carlo-approach, we proceed as follows

1. Based on your data \mathbf{X} , calculate the empirical covariance matrix $\hat{\boldsymbol{\Sigma}}$.
2. Generate $\mathbf{Z}_1^{MC} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \hat{\boldsymbol{\Sigma}})$.

$$3. \text{ Calculate the Monte-Carlo quadratic form } \mathbf{Q}^{MC,1} := \begin{pmatrix} \widehat{v}_{11}^{-1/2} \cdot (\mathbf{C}_1 \mathbf{Z}_1^{MC})^\top \mathbf{M}(\mathbf{C}_1, \widehat{\boldsymbol{\Sigma}}) \mathbf{C}_1 \mathbf{Z}_1^{MC} \\ \vdots \\ \widehat{v}_{LL}^{-1/2} \cdot (\mathbf{C}_L \mathbf{Z}_L^{MC})^\top \mathbf{M}(\mathbf{C}_L, \widehat{\boldsymbol{\Sigma}}) \mathbf{C}_L \mathbf{Z}_L^{MC} \end{pmatrix}$$

$$\text{with } \widehat{v}_{\ell\ell} := 2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\boldsymbol{\Sigma}}) \mathbf{C}_\ell \widehat{\boldsymbol{\Sigma}}]).$$

4. Repeat steps 2. and 3. a large number of times, e.g. $B=10.000$ times, to obtain $\mathbf{Q}^{MC,1}, \dots, \mathbf{Q}^{MC,B}$ and determine quantiles $q_{1,B}^{MC}, \dots, q_{L,B}^{MC}$.

Lemma 4.2 ensures that this method is asymptotically valid.

Proof of Theorem 5.1: First we define $\overline{\mathbf{X}}_i^* := \frac{1}{n_i} \sum_{k=1}^{n_i} \mathbf{X}_{ik}^*$ for $i=1, \dots, a$ and based on this $\overline{\mathbf{X}}^* = (\overline{\mathbf{X}}_1^*, \dots, \overline{\mathbf{X}}_a^*)^\top$ as bootstrap pendant to $\overline{\mathbf{X}}_i$ resp. $\overline{\mathbf{X}}$. Now we use the conditional Lindeberg-Feller-theorem to show that given the data $\overline{\mathbf{X}}_i^* \stackrel{\mathcal{D}}{\rightarrow} \mathbf{Y}_i \sim \mathcal{N}_a(\mathbf{0}_a, \boldsymbol{\Sigma}_i)$. From the generation of we know that given the data, $\mathbf{X}_{i1}^*, \dots, \mathbf{X}_{in_i}^*$ are independent with

$$\begin{aligned} 1.) \quad & \sum_{k=1}^{n_i} \mathbb{E} \left(\frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^* \mid \mathbf{X} \right) = \sum_{k=1}^{n_i} \frac{\sqrt{N}}{n_i} \mathbb{E}(\mathbf{X}_{ik}^* \mid \mathbf{X}) = \mathbf{0}_d \\ 2.) \quad & \sum_{k=1}^{n_i} \operatorname{Cov} \left(\frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^* \mid \mathbf{X} \right) = \frac{N}{n_i^2} \sum_{k=1}^{n_i} \widehat{\boldsymbol{\Sigma}}_i \xrightarrow{\mathcal{P}} \frac{1}{\kappa_i} \boldsymbol{\Sigma}_i \\ 3.) \quad & \lim_{N \rightarrow \infty} \sum_{k=1}^{n_i} \mathbb{E} \left(\left\| \frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^* \right\|^2 \cdot \mathbb{1}_{\left\| \frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^* \right\| > \delta} \mid \mathbf{X} \right) \\ & = \lim_{N \rightarrow \infty} \frac{N}{n_i^2} \sum_{k=1}^{n_i} \mathbb{E} \left(\left\| \mathbf{X}_{ik}^* \right\|^2 \cdot \mathbb{1}_{\left\| \mathbf{X}_{ik}^* \right\| > \delta \cdot n_i / \sqrt{N}} \mid \mathbf{X} \right) \\ & \leq \frac{1}{\kappa_i} \lim_{N \rightarrow \infty} \sum_{k=1}^{n_i} \sqrt{\mathbb{E}(\left\| \mathbf{X}_{ik}^* \right\|^2 \mid \mathbf{X})} \cdot \sqrt{\mathbb{E} \left(\mathbb{1}_{\left\| \mathbf{X}_{ik}^* \right\| > \delta \cdot n_i / \sqrt{N}} \mid \mathbf{X} \right)} \stackrel{\mathcal{P}}{=} 0, \end{aligned}$$

where Cauchy-Bunjakowski-Schwarz inequality is used in the last line. Together with the fact that, given the data, the observations follow a multivariate normal distribution, the first expectation value is bounded. From (A1) we know $n_i/N \rightarrow \kappa_i$ and therefore $\delta \cdot n_i / \sqrt{N} \rightarrow \infty$ so it follows with the conditional Markov inequality

$$P(\left\| \mathbf{X}_{i1}^* \right\| > \delta \cdot n_i / \sqrt{N} \mid \mathbf{X}) \leq \frac{\sqrt{N}}{n_i} \cdot \mathbb{E}(\left\| \mathbf{X}_{i1}^* \right\| \mid \mathbf{X}) \leq \frac{\sqrt{N}}{n_i} \cdot \|(\widehat{\boldsymbol{\Sigma}}_i)^{1/2}\|_{\max} \cdot \mathbb{E}(\left\| \mathbf{Y} \right\|)$$

with $\mathbf{Y} \sim \mathcal{N}_d(\mathbf{0}_d, \mathbf{I}_d)$. Using $\mathbb{E}(\left\| \mathbf{Y} \right\|) < \infty$ and $\|(\widehat{\boldsymbol{\Sigma}}_i)^{1/2}\|_{\max} \xrightarrow{\mathcal{P}} \|(\boldsymbol{\Sigma}_i)^{1/2}\|_{\max} < \infty$ this together with Slutsky's theorem completes the Lindeberg-Feller conditions. Now, from the independence of groups, it directly follows, that, given the data, the conditional distribution of $\sqrt{N} \overline{\mathbf{X}}^*$ converges weakly to $\mathbf{Z} \sim \mathcal{N}_{ad}(\mathbf{0}_{ad}, \boldsymbol{\Sigma})$.

For the consistency of the covariance estimator, we use again the representation $\mathbf{X}_{ij}^* = \widehat{\boldsymbol{\Sigma}}_i^{1/2} \mathbf{Y}_{ij}$ for $\mathbf{Y}_{11}, \dots, \mathbf{Y}_{an_a} \sim \mathcal{N}_d(\mathbf{0}_d, \mathbf{I}_d)$ i.i.d. and independent of the data. Then, we have

$$\widehat{\boldsymbol{\Sigma}}_i^* = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\mathbf{X}_{ij}^* - \overline{\mathbf{X}}_i^*) (\mathbf{X}_{ij}^* - \overline{\mathbf{X}}_i^*)^\top = \widehat{\boldsymbol{\Sigma}}_i^{1/2} \left(\frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\mathbf{Y}_{ij} - \overline{\mathbf{Y}}_i) (\mathbf{Y}_{ij} - \overline{\mathbf{Y}}_i)^\top \right) \widehat{\boldsymbol{\Sigma}}_i^{1/2}.$$

The term in the middle is the empirical covariance matrix of $\mathbf{Y}_{i1}, \dots, \mathbf{Y}_{in_i}$ which converges in probability to \mathbf{I}_d . Hence, it follows $\widehat{\boldsymbol{\Sigma}}_i^* \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}_i$. With independence of groups it follows

$$\widehat{\Sigma}^* = N \bigoplus_{i=1}^a n_i^{-1} \widehat{\Sigma}_i^* \xrightarrow{\mathcal{P}} \Sigma.$$

Now as continuous function of consistent estimators, $2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^*) \mathbf{C}_\ell \widehat{\Sigma}^*]^2)$ is also a consistent estimator of $v_{\ell\ell}$. Therefore all necessary properties are shown, and with Slutsky's theorem, the result follows since \mathbf{Q}_N^* can be written by using the continuous function \mathbf{f} and the estimator $\widehat{\Sigma}^*$. \square

For the wild bootstrap approach, $i = 1, \dots, a$, let W_{ik} be i.i.d. with $\mathbb{E}(W_{ik}) = 0$ and $\operatorname{Var}(W_{ik}) = 1$, while usually distributions are standard normal or Mammen (Mammen [1993]). Then the wild bootstrap observation vectors be given through $\mathbf{X}_{ik}^\dagger = W_{ik}(\mathbf{X}_{ik} - \overline{\mathbf{X}}_i)$ and we define

$$\overline{\mathbf{X}}^\dagger = \left(\frac{1}{n_1} \sum_{k=1}^{n_1} \mathbf{X}_{1k}^{\dagger \top}, \dots, \frac{1}{n_a} \sum_{k=1}^{n_a} \mathbf{X}_{ak}^{\dagger \top} \right)^\top.$$

Analogues to the parametric bootstrap, based on our wild bootstrap we define

$$Q_{N\ell}^\dagger = N \cdot \frac{[\mathbf{C}_\ell \overline{\mathbf{X}}^\dagger]^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^\dagger) [\mathbf{C}_\ell \overline{\mathbf{X}}^\dagger]}{\sqrt{2 \operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^\dagger) \mathbf{C}_\ell \widehat{\Sigma}^\dagger]^2)}} \cdot \mathbb{1} \left(\operatorname{tr}([\mathbf{C}_\ell^\top \mathbf{M}(\mathbf{C}_\ell, \widehat{\Sigma}^\dagger) \mathbf{C}_\ell \widehat{\Sigma}^\dagger]^2) > 0 \right)$$

for $\ell \in \{1, \dots, L\}$, using the covariance estimators $\widehat{\Sigma}^\dagger$ and $\widetilde{\mathbf{Q}}_{Nj}^\dagger$ based on $\widehat{\Sigma}^{(j)\dagger}$.

Theorem A.1:

If Assumptions (A1) and (A2) are fulfilled, it holds: Given the data, the conditional distribution of

- (a) $\sqrt{N} \overline{\mathbf{X}}^\dagger$ converges weakly to $\mathcal{N}_a(\mathbf{0}_a, \Sigma)$ in probability.
- (b) $\widetilde{\mathbf{Q}}_N^\dagger = (\widetilde{\mathbf{Q}}_{N1}^\dagger, \dots, \widetilde{\mathbf{Q}}_{Nd}^\dagger)^\top$ converges weakly to $\widetilde{\mathbf{Q}}$ in probability.
- (c) $\mathbf{Q}_N^\dagger = (\mathbf{Q}_{N1}^\dagger, \dots, \mathbf{Q}_{NL}^\dagger)^\top$ converges weakly to \mathbf{Q} in probability, if also (A3) is fulfilled.

Moreover, we have $\widehat{\Sigma}^\dagger \xrightarrow{\mathcal{P}} \Sigma$ and the unknown covariance matrix can be estimated through this estimator.

Proof: For the wild bootstrap approach, we follow the same steps as for the parametric bootstrap.

- 1.)
$$\sum_{k=1}^{n_i} \mathbb{E} \left(\frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^\dagger | \mathbf{X} \right) = \sum_{k=1}^{n_i} \frac{\sqrt{N}}{n_i} \mathbb{E}(W_{ik})(\mathbf{X}_{ik} - \overline{\mathbf{X}}_i) = \mathbf{0}_d$$
- 2.)
$$\begin{aligned} & \sum_{k=1}^{n_i} \operatorname{Cov} \left(\frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^\dagger | \mathbf{X} \right) \\ &= \frac{N}{n_i^2} \sum_{k=1}^{n_i} \mathbb{E}((\mathbf{X}_{ik}^\dagger)(\mathbf{X}_{ik}^\dagger)^\top | \mathbf{X}) \\ &= \frac{N}{n_i^2} \sum_{k=1}^{n_i} \mathbb{E}(W_{ik}^2) \cdot (\mathbf{X}_{ik} - \overline{\mathbf{X}}_i)(\mathbf{X}_{ik} - \overline{\mathbf{X}}_i)^\top \\ &= \frac{N(n_i-1)}{n_i^2} \frac{1}{n_i-1} \sum_{k=1}^{n_i} 1 \cdot (\mathbf{X}_{ik} - \overline{\mathbf{X}}_i)(\mathbf{X}_{ik} - \overline{\mathbf{X}}_i)^\top \\ &= \frac{N(n_i-1)}{n_i^2} \widehat{\Sigma}_i \xrightarrow{\mathcal{P}} \frac{1}{\kappa_i} \Sigma_i \end{aligned}$$

$$\begin{aligned}
3.) \quad & \lim_{N \rightarrow \infty} \sum_{k=1}^{n_i} \mathbb{E} \left(\left\| \frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^* \right\|^2 \cdot \mathbb{1}_{\left\| \frac{\sqrt{N}}{n_i} \mathbf{X}_{ik}^* \right\| > \delta} \middle| \mathbf{X} \right) \\
&= \lim_{N \rightarrow \infty} \frac{N}{n_i^2} \sum_{k=1}^{n_i} \mathbb{E} \left(\left\| \mathbf{X}_{ik}^\dagger \right\|^2 \cdot \mathbb{1}_{\left\| \mathbf{X}_{ik}^\dagger \right\| > \delta \cdot n_i / \sqrt{N}} \middle| \mathbf{X} \right) \\
&\leq \frac{1}{\kappa_i} \lim_{N \rightarrow \infty} \sum_{k=1}^{n_i} \sqrt{\mathbb{E} \left(\left\| \mathbf{X}_{ik}^\dagger \right\|^2 \middle| \mathbf{X} \right)} \cdot \sqrt{\mathbb{E} \left(\mathbb{1}_{\left\| \mathbf{X}_{ik}^\dagger \right\| > \delta \cdot n_i / \sqrt{N}} \middle| \mathbf{X} \right)} \stackrel{\mathcal{P}}{=} 0.
\end{aligned}$$

Since, given the data, \mathbf{X}_{ik}^\dagger has fourth second moments, all remaining steps can be done analogues to the proof of Theorem 5.1. \square

Lemma A.1:

For $\alpha \in (0, 1)$, let $q_{1,B}^\dagger, \dots, q_{L,B}^\dagger$ be the corresponding quantiles of \mathbf{Q}_N^\dagger according to Munko et al. [2023a]. Then

$$\varphi_{\mathbf{Q}}^\dagger(N, B) := \mathbb{1} \left\{ \max_{\ell=1, \dots, L} \left(\frac{Q_{N\ell}}{q_{\ell,B}^\dagger} \right) > 1 \right\}$$

is an asymptotic correct level α test for the global hypothesis $\mathcal{H}_0 : \mathbf{C}\boldsymbol{\mu} = \boldsymbol{\beta}$ that controls the family-wise error rate asymptotically in the strong sense.

Proof of the asymptotic correctness and family-wise error rate control of the tests In this paragraph, we prove Lemma 4.2, Lemma 5.2, and Lemma A.1. Therefore, we aim to apply Lemma S8 in the supplement of Munko et al. [2024] with $\varepsilon_N := 1/B_N$ and $F : \mathbb{R}^L \rightarrow [0, 1]$ being the distribution function of \mathbf{Q} , which is continuous with strictly increasing marginal distribution functions on $[0, \infty)$ if (A3) holds. For Lemma 4.2, we choose F_n as empirical distribution function of the Monte-Carlo replicates $\mathbf{Q}^{MC,1}, \dots, \mathbf{Q}^{MC,B_N}$. For the bootstrap methods, we choose F_N as empirical distribution function of the Monte-Carlo replicates $\mathbf{Q}_N^{*,1}, \dots, \mathbf{Q}_N^{*,B_N}$ resp. $\mathbf{Q}_N^{\dagger,1}, \dots, \mathbf{Q}_N^{\dagger,B_N}$. Moreover, $F_{N,\ell}, \ell = 1, \dots, L$, are chosen as the marginal distribution functions of F_N . Then, the consistency of the empirical covariance matrix, Theorem 5.1 and Theorem A.1 ensure the condition of Lemma S7 in Munko et al. [2024], respectively, and, thus, (S10) therein follows. Furthermore, Remark 1 in Munko et al. [2024] yields (S11). Hence, Lemma S8 in Munko et al. [2024] is applicable and provides that the quantiles $q_{\ell,B_N}^{MC}, q_{\ell,B_N}^*, q_{\ell,B_N}^\dagger$ all converge in probability to q_ℓ fulfilling $F(q_1, \dots, q_L) = 1 - \alpha$. Thus, the asymptotic correctness follows. The asymptotic control of the family-wise error rate in the strong sense follows from

$$\begin{aligned}
\mathbb{E} \left[\max_{\ell \in \mathcal{T}} \mathbb{1} \left\{ \left(\frac{Q_{N\ell}}{q_{\ell,B_N}^{MC}} \right) > 1 \right\} \right] &\rightarrow \mathbb{E} \left[\max_{\ell \in \mathcal{T}} \mathbb{1} \left\{ \left(\frac{Q_\ell}{q_\ell} \right) > 1 \right\} \right] \\
&\leq \mathbb{E} \left[\max_{\ell \in \{1, \dots, L\}} \mathbb{1} \left\{ \left(\frac{Q_\ell}{q_\ell} \right) > 1 \right\} \right] \\
&= 1 - F(q_1, \dots, q_L) = \alpha
\end{aligned}$$

for all index subsets $\mathcal{T} \subset \{1, \dots, L\}$ of true hypotheses. Analogously, the asymptotic control of the family-wise error rate in the strong sense holds for the bootstrap methods.

B Further simulations

Since in the main paper, only the one-point-alternative was investigated, we now additionally consider a shift-alternative, given through $\mu_1 = \mu_2 = \mathbf{0}_5$ and $\mu_3 = \delta \cdot \mathbf{1}_5$. Furthermore, we consider one more distribution for Z , a t-distribution with 9 degrees of freedom, which is centred and standardized. While up to now, only the ATS was used for the QFMCT, here we additionally considered it based on the WTS, as comparison. For all tests, we further consider the wild bootstrap denoted by \dagger .

The results can be seen in Table 5-10. First of all, we want to focus on comparing both bootstrap approaches, parametric and wild. While for the classical MCT and the classical ATS the wild bootstrap seems to be slightly better, for the QFMCT the parametric bootstrap performs better. For example, over all distributions and sample size the type-I error rate of $\varphi_Q^*(ATS)$ is in the 95% confidence interval except for $N = 25$ and the skew normal distribution, which is more often the case for $\varphi_Q^\dagger(ATC)$, since it is a bit more liberal. Of course, this often leads to even higher power compared with the other test, which allows the choice of the procedure in accordance with the respective focus. Also, the QFMCT based on the WTS performs well for both resampling techniques, in a direct comparison the ATS is slightly better overall. Moreover, the ATS has a shorter calculation time and less restrictive conditions (for the ATS (A3) is equivalent to Σ having no zero rows). So, the usage of the WTS instead of the ATS for QFMCT is only reasonable if their other properties are required.

The results for the t_9 -distribution under the one-point-alternative are quite similar to the distributions of the main part, although the type-I-error rate is a bit worse for the very small sample size $N = 25$. Under the alternative, the better power of the QFMTC can be seen again.

As expected, the shift-alternative is easier to detect, leading to a higher power for all considered tests. Nevertheless, the differences and behaviour of the different approaches remain the same, although the power is higher for smaller δ . Together, the additional simulations support the conclusion from the main part that the QFMCT is preferable to test the hypothesis of equal group mean since it has the same type-I-error rate under the null hypothesis but usually has a better power.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	15.29	14.79	17.86	24.58	32.73	45.19	58.62	70.16	81.71
φ^*	25	5.22	5.22	6.04	9.30	15.25	23.62	34.13	45.46	59.29
φ^\dagger	25	5.00	4.97	7.01	10.32	15.12	22.34	33.49	46.83	59.03
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	25	9.70	9.96	12.00	17.71	25.78	38.09	52.15	66.15	79.64
$\varphi_{\mathcal{Q}}^*(ATS)$	25	4.95	5.12	6.61	9.83	16.12	25.09	37.65	51.12	66.53
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	25	5.64	5.57	7.70	11.32	17.04	26.35	39.26	54.33	67.92
$\varphi_{\mathcal{Q}}^*(WTS)$	25	4.87	5.23	6.51	9.69	15.78	24.32	35.40	47.02	61.18
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	25	4.96	5.19	6.93	10.50	15.49	23.64	35.01	48.07	60.13
ψ_{ATS}^*	25	3.95	4.22	5.01	6.03	8.91	12.60	18.18	25.59	36.74
ψ_{ATS}^\dagger	25	5.26	5.42	6.62	8.48	11.16	15.42	22.10	31.47	42.01
φ	50	8.99	10.75	16.72	27.92	47.47	68.79	84.30	94.06	98.57
φ^*	50	4.75	5.99	10.48	18.77	36.35	57.82	76.53	89.32	96.72
φ^\dagger	50	5.12	6.06	9.56	19.12	36.07	56.79	76.05	90.31	96.46
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	50	7.09	8.25	14.01	24.78	45.34	68.33	84.28	94.61	98.84
$\varphi_{\mathcal{Q}}^*(ATS)$	50	4.93	5.98	10.70	19.74	38.79	62.09	79.81	92.26	97.95
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	50	5.55	6.05	10.14	20.54	39.20	61.38	80.27	93.16	97.82
$\varphi_{\mathcal{Q}}^*(WTS)$	50	4.67	6.04	10.48	19.25	37.39	59.29	77.36	89.81	97.04
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	50	5.10	6.01	9.50	19.61	36.85	57.77	77.25	90.62	96.68
ψ_{ATS}^*	50	4.58	5.47	7.23	10.97	18.31	31.99	48.77	68.94	85.43
ψ_{ATS}^\dagger	50	5.08	5.69	7.90	11.84	20.74	33.72	52.07	72.06	87.18
φ	100	6.82	10.07	22.86	50.79	78.96	95.23	99.40	99.92	100.00
φ^*	100	5.07	7.65	19.11	45.18	74.80	93.59	99.09	99.89	100.00
φ^\dagger	100	4.87	7.68	18.47	44.45	74.52	92.98	99.05	99.90	100.00
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	100	6.14	9.28	21.90	50.66	79.76	95.67	99.61	99.94	100.00
$\varphi_{\mathcal{Q}}^*(ATS)$	100	5.22	8.13	19.90	47.41	77.55	95.10	99.43	99.94	100.00
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	100	5.13	8.04	19.13	47.19	77.95	94.65	99.38	99.94	100.00
$\varphi_{\mathcal{Q}}^*(WTS)$	100	4.97	7.78	19.38	46.36	76.31	94.06	99.28	99.91	100.00
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	100	5.05	7.50	18.65	45.43	75.75	93.60	99.13	99.93	100.00
ψ_{ATS}^*	100	4.91	6.01	10.58	22.93	45.67	74.25	92.92	99.06	99.97
ψ_{ATS}^\dagger	100	5.15	6.45	11.01	23.97	46.92	75.80	94.10	99.29	99.95

Table 5: Power of different test statistics under an one-point-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a standard normal distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	14.69	14.29	18.29	24.31	33.35	44.86	59.28	71.07	81.07
φ^*	25	4.76	3.94	6.41	9.36	14.91	23.36	35.41	47.65	60.53
φ^\dagger	25	4.50	5.47	6.92	10.23	15.45	25.41	35.75	49.25	62.79
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	25	9.89	9.12	12.29	17.44	25.53	36.52	52.35	65.95	77.82
$\varphi_{\mathcal{Q}}^*(ATS)$	25	4.93	4.51	6.39	9.43	15.69	24.66	37.94	51.70	64.85
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	25	5.08	5.72	7.33	10.71	17.38	27.87	40.41	55.22	69.18
$\varphi_{\mathcal{Q}}^*(WTS)$	25	4.78	4.14	6.70	9.66	15.56	24.37	37.14	49.66	61.71
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	25	4.49	5.42	7.00	10.39	16.49	26.39	37.16	50.32	63.76
ψ_{ATS}^*	25	4.27	4.02	5.12	6.26	8.15	12.18	18.79	25.39	36.62
ψ_{ATS}^\dagger	25	5.47	5.28	6.76	7.81	11.10	16.06	22.29	31.38	41.90
φ	50	8.54	10.41	16.41	29.80	47.79	68.69	84.92	94.03	98.14
φ^*	50	4.59	5.84	10.07	20.55	36.86	58.22	76.71	89.99	96.00
φ^\dagger	50	4.81	5.76	10.27	20.01	38.44	59.18	77.13	89.64	96.58
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	50	6.78	8.21	13.39	25.99	45.09	66.57	84.42	94.04	98.25
$\varphi_{\mathcal{Q}}^*(ATS)$	50	4.62	5.98	10.08	20.58	38.32	60.39	79.69	91.66	97.28
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	50	4.92	6.42	10.66	21.05	40.72	62.72	80.66	92.06	97.81
$\varphi_{\mathcal{Q}}^*(WTS)$	50	4.63	5.68	10.32	20.97	37.95	59.58	78.12	90.50	96.09
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	50	4.66	5.74	10.69	20.99	39.50	60.66	78.21	90.26	96.71
ψ_{ATS}^*	50	4.71	5.10	6.85	10.92	18.45	30.56	49.70	68.70	84.96
ψ_{ATS}^\dagger	50	5.13	5.74	7.81	12.22	20.49	34.07	51.88	71.87	87.13
φ	100	6.89	9.49	22.89	50.62	79.53	95.05	99.27	99.91	100.00
φ^*	100	5.07	7.25	18.73	45.11	75.46	93.42	98.90	99.85	99.99
φ^\dagger	100	5.23	7.45	18.32	45.81	75.28	93.08	99.09	99.93	99.99
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	100	6.30	8.57	21.72	49.82	79.90	95.50	99.38	99.97	100.00
$\varphi_{\mathcal{Q}}^*(ATS)$	100	5.14	7.32	19.27	46.90	77.69	94.58	99.24	99.93	100.00
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	100	5.28	7.77	19.40	47.54	77.68	94.59	99.37	99.98	100.00
$\varphi_{\mathcal{Q}}^*(WTS)$	100	4.79	7.22	19.12	46.47	76.54	94.05	98.93	99.87	99.98
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	100	5.02	7.54	18.86	46.70	76.59	93.75	99.16	99.95	99.99
ψ_{ATS}^*	100	5.16	6.12	10.58	22.26	46.21	74.33	92.99	99.05	99.92
ψ_{ATS}^\dagger	100	5.08	6.19	10.45	23.00	46.70	75.30	94.00	99.13	99.97

Table 6: Power of different test statistics under an one-point-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a t_9 distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	14.40	15.60	18.02	23.88	32.73	45.02	59.45	73.73	84.51
φ^*	25	4.54	4.96	6.19	9.38	14.16	22.49	34.64	47.20	63.06
φ^\dagger	25	4.90	5.26	6.66	9.78	14.67	22.71	33.44	48.73	61.96
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	25	9.12	9.61	11.60	16.63	24.10	35.47	50.66	65.74	79.17
$\varphi_{\mathcal{Q}}^*(ATS)$	25	4.42	4.66	5.89	8.83	14.19	22.68	35.20	50.42	65.48
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	25	5.40	5.21	7.42	9.95	15.79	24.38	37.89	53.17	67.53
$\varphi_{\mathcal{Q}}^*(WTS)$	25	4.82	5.39	6.34	9.61	14.36	22.43	35.13	48.21	63.77
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	25	5.17	5.37	6.98	9.67	14.79	22.51	33.72	48.62	62.60
ψ_{ATS}^*	25	4.25	4.40	4.43	6.18	8.38	11.97	17.87	25.44	36.25
ψ_{ATS}^\dagger	25	5.51	5.46	6.66	7.79	10.69	15.90	20.96	30.52	41.36
φ	50	9.31	10.52	16.09	28.51	48.44	70.50	87.36	96.05	99.31
φ^*	50	4.99	5.65	9.77	19.35	36.38	58.71	79.52	92.40	98.01
φ^\dagger	50	5.10	5.88	9.97	19.79	36.97	59.57	79.76	92.39	98.09
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	50	7.11	7.96	13.50	24.96	44.62	66.60	84.54	95.32	98.82
$\varphi_{\mathcal{Q}}^*(ATS)$	50	5.02	5.57	9.59	19.79	37.41	59.80	80.44	93.41	98.19
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	50	4.95	5.69	9.55	19.81	38.43	61.32	81.57	93.62	98.47
$\varphi_{\mathcal{Q}}^*(WTS)$	50	5.40	5.77	9.83	19.49	36.70	59.34	80.85	93.44	98.52
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	50	5.18	6.20	9.91	19.40	36.58	60.18	81.00	93.36	98.57
ψ_{ATS}^*	50	4.87	5.12	6.84	10.96	18.30	30.15	48.85	68.94	85.25
ψ_{ATS}^\dagger	50	5.26	5.88	7.81	12.06	20.10	33.20	52.37	71.51	87.23
φ	100	6.86	9.56	22.52	51.64	80.86	96.13	99.71	99.97	100.00
φ^*	100	4.96	7.47	18.42	45.62	76.26	94.74	99.57	99.96	100.00
φ^\dagger	100	4.82	7.41	18.39	46.44	77.21	95.15	99.48	99.98	100.00
$\varphi_{\mathcal{Q}}^{MC}(ATS)$	100	6.04	8.62	20.76	49.77	79.63	95.78	99.57	99.99	100.00
$\varphi_{\mathcal{Q}}^*(ATS)$	100	4.93	7.24	18.48	46.69	77.33	95.07	99.44	100.00	100.00
$\varphi_{\mathcal{Q}}^\dagger(ATs)$	100	4.96	7.41	18.22	47.52	78.39	95.46	99.59	99.97	100.00
$\varphi_{\mathcal{Q}}^*(WTS)$	100	5.26	7.45	18.02	46.48	78.19	95.89	99.69	99.98	100.00
$\varphi_{\mathcal{Q}}^\dagger(WTS)$	100	5.04	7.35	18.26	47.07	79.14	96.15	99.62	100.00	100.00
ψ_{ATS}^*	100	4.64	5.87	10.77	22.07	44.14	73.53	93.27	99.28	99.97
ψ_{ATS}^\dagger	100	5.19	6.32	10.93	23.54	46.56	75.27	94.33	99.34	99.97

Table 7: Power of different test statistics under an one-point-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a skew normal distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	15.29	17.17	28.06	43.94	61.04	75.89	88.33	94.93	98.24
φ^*	25	5.22	6.26	11.66	21.68	36.97	53.02	69.98	82.89	91.75
φ^\dagger	25	5.00	6.50	12.10	22.19	36.28	53.60	69.77	82.95	91.51
$\varphi_{\hat{Q}}^{MC}(ATS)$	25	9.70	11.87	20.61	34.92	52.55	70.10	84.53	93.29	97.79
$\varphi_{\hat{Q}}^*(ATS)$	25	4.95	6.47	11.64	22.48	38.24	55.74	73.92	86.06	94.45
$\varphi_{\hat{Q}}^\dagger(ATs)$	25	5.64	6.94	13.50	24.32	40.05	58.87	75.82	88.31	95.02
$\varphi_{\hat{Q}}^*(WTS)$	25	4.87	6.40	11.87	22.42	38.53	54.64	71.40	83.71	92.23
$\varphi_{\hat{Q}}^\dagger(WTS)$	25	4.96	6.68	12.44	22.70	37.28	54.56	70.89	83.79	91.94
ψ_{ATS}^*	25	3.95	5.98	13.01	27.04	45.51	65.28	82.81	92.43	97.83
ψ_{ATS}^\dagger	25	5.26	7.65	15.32	29.26	49.54	70.14	84.85	94.47	98.16
φ	50	8.99	14.73	31.72	57.69	81.82	94.55	98.89	99.86	100.00
φ^*	50	4.75	9.20	22.11	46.05	72.10	89.84	97.44	99.59	99.97
φ^\dagger	50	5.12	8.96	21.59	45.53	71.97	89.38	97.12	99.68	99.95
$\varphi_{\hat{Q}}^{MC}(ATS)$	50	7.09	12.40	28.02	54.12	79.52	93.94	98.78	99.87	100.00
$\varphi_{\hat{Q}}^*(ATS)$	50	4.93	9.21	22.99	47.08	74.18	91.42	98.09	99.77	100.00
$\varphi_{\hat{Q}}^\dagger(ATs)$	50	5.55	8.65	22.68	47.43	74.81	91.30	98.12	99.84	100.00
$\varphi_{\hat{Q}}^*(WTS)$	50	4.67	9.27	22.82	46.51	72.75	90.34	97.63	99.67	99.97
$\varphi_{\hat{Q}}^\dagger(WTS)$	50	5.10	8.78	21.97	46.20	72.60	89.61	97.39	99.68	99.96
ψ_{ATS}^*	50	4.58	9.81	26.68	56.21	82.39	95.66	99.30	99.94	99.99
ψ_{ATS}^\dagger	50	5.08	10.00	28.52	57.64	83.01	96.03	99.47	99.97	100.00
φ	100	6.82	16.97	48.73	84.42	98.06	99.86	100.00	100.00	100.00
φ^*	100	5.07	13.70	43.27	80.11	97.05	99.75	100.00	100.00	100.00
φ^\dagger	100	4.87	13.76	42.45	80.18	96.92	99.83	100.00	100.00	100.00
$\varphi_{\hat{Q}}^{MC}(ATS)$	100	6.14	15.82	47.35	83.48	98.08	99.90	100.00	100.00	100.00
$\varphi_{\hat{Q}}^*(ATS)$	100	5.22	14.16	44.35	81.38	97.33	99.85	100.00	100.00	100.00
$\varphi_{\hat{Q}}^\dagger(ATs)$	100	5.13	14.13	43.97	81.60	97.63	99.92	100.00	100.00	100.00
$\varphi_{\hat{Q}}^*(WTS)$	100	4.97	13.82	43.57	80.68	97.17	99.80	100.00	100.00	100.00
$\varphi_{\hat{Q}}^\dagger(WTS)$	100	5.05	13.67	42.88	80.96	97.20	99.86	100.00	100.00	100.00
ψ_{ATS}^*	100	4.91	15.90	52.51	88.92	99.05	100.00	100.00	100.00	100.00
ψ_{ATS}^\dagger	100	5.15	16.66	53.34	89.26	99.02	99.98	100.00	100.00	100.00

Table 8: Power of different test statistics under a shift-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a standard normal distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	14.69	17.36	28.09	43.69	61.29	77.80	90.12	95.60	98.56
φ^*	25	4.76	5.33	11.99	21.71	36.45	54.82	72.00	84.79	92.59
φ^\dagger	25	4.50	6.47	12.96	23.49	38.43	56.91	73.35	85.73	93.26
$\varphi_{\hat{Q}}^{MC}(ATS)$	25	9.89	11.62	20.14	34.78	52.47	70.06	86.21	93.90	97.80
$\varphi_{\hat{Q}}^*(ATS)$	25	4.93	5.76	11.37	22.27	37.71	56.48	75.07	87.41	94.45
$\varphi_{\hat{Q}}^\dagger(ATs)$	25	5.08	7.42	13.18	24.12	42.05	61.01	77.45	88.90	95.61
$\varphi_{\hat{Q}}^*(WTS)$	25	4.78	5.61	12.48	22.38	38.28	56.42	73.74	85.59	92.76
$\varphi_{\hat{Q}}^\dagger(WTS)$	25	4.49	6.51	12.99	23.94	39.91	58.23	74.24	86.16	93.19
ψ_{ATS}^*	25	4.27	6.04	13.20	25.82	44.44	65.23	82.34	92.40	97.28
ψ_{ATS}^\dagger	25	5.47	7.49	15.83	29.15	49.87	69.74	85.00	94.03	97.77
φ	50	8.54	14.81	32.96	59.97	82.25	95.01	99.04	99.84	100.00
φ^*	50	4.59	9.09	22.88	47.39	72.88	90.45	97.82	99.60	99.97
φ^\dagger	50	4.81	8.87	22.86	48.14	73.87	91.09	97.81	99.63	99.95
$\varphi_{\hat{Q}}^{MC}(ATS)$	50	6.78	11.91	28.13	54.59	79.70	94.14	98.94	99.84	100.00
$\varphi_{\hat{Q}}^*(ATS)$	50	4.62	8.98	22.79	47.21	74.24	91.49	98.27	99.67	99.99
$\varphi_{\hat{Q}}^\dagger(ATs)$	50	4.92	8.88	23.67	49.39	75.35	92.22	98.31	99.77	99.99
$\varphi_{\hat{Q}}^*(WTS)$	50	4.63	9.02	23.48	48.13	73.50	91.15	97.99	99.65	99.98
$\varphi_{\hat{Q}}^\dagger(WTS)$	50	4.66	8.73	23.23	49.39	74.64	91.33	97.91	99.71	99.96
ψ_{ATS}^*	50	4.71	9.38	26.96	56.21	82.42	95.55	99.24	99.93	99.99
ψ_{ATS}^\dagger	50	5.13	10.05	28.96	57.62	83.89	96.22	99.36	99.97	99.99
φ	100	6.89	17.16	50.39	84.19	98.24	99.83	100.00	100.00	100.00
φ^*	100	5.07	13.58	44.59	80.13	97.15	99.74	100.00	100.00	100.00
φ^\dagger	100	5.23	14.41	44.32	80.60	97.09	99.77	99.99	100.00	100.00
$\varphi_{\hat{Q}}^{MC}(ATS)$	100	6.30	15.79	48.19	83.16	98.24	99.84	100.00	100.00	100.00
$\varphi_{\hat{Q}}^*(ATS)$	100	5.14	13.92	44.86	81.31	97.63	99.80	100.00	100.00	100.00
$\varphi_{\hat{Q}}^\dagger(ATs)$	100	5.28	14.87	44.65	81.94	97.64	99.82	100.00	100.00	100.00
$\varphi_{\hat{Q}}^*(WTS)$	100	4.79	13.63	44.83	80.93	97.49	99.77	100.00	100.00	100.00
$\varphi_{\hat{Q}}^\dagger(WTS)$	100	5.02	14.45	44.31	81.32	97.43	99.74	99.99	100.00	100.00
ψ_{ATS}^*	100	5.16	15.51	53.26	88.27	98.99	99.96	100.00	100.00	100.00
ψ_{ATS}^\dagger	100	5.08	16.86	53.25	89.12	99.12	99.97	100.00	100.00	100.00

Table 9: Power of different test statistics under a shift-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a t_9 -distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.

	N	$\delta = 0$	$\delta = 0.25$	$\delta = 0.5$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 1.25$	$\delta = 1.5$	$\delta = 1.75$	$\delta = 2.00$
φ	25	14.40	17.66	28.37	43.78	62.29	77.46	90.57	96.13	99.18
φ^*	25	4.54	6.51	12.18	22.04	36.84	53.06	71.90	85.10	93.69
φ^\dagger	25	4.90	6.82	12.71	22.10	37.55	54.57	71.30	85.64	93.45
$\varphi_{\hat{Q}}^{MC}(ATS)$	25	9.12	11.38	19.72	33.44	51.13	68.60	84.21	93.71	98.04
$\varphi_{\hat{Q}}^*(ATS)$	25	4.42	5.78	11.21	20.82	35.70	52.44	71.68	86.24	94.23
$\varphi_{\hat{Q}}^\dagger(ATs)$	25	5.40	6.70	13.25	22.35	38.95	57.12	75.12	88.31	95.31
$\varphi_{\hat{Q}}^*(WTS)$	25	4.82	6.61	12.20	22.05	36.90	52.93	71.85	85.12	93.74
$\varphi_{\hat{Q}}^\dagger(WTS)$	25	5.17	6.75	12.55	21.85	36.97	53.77	70.92	85.38	93.14
ψ_{ATS}^*	25	4.25	6.03	12.29	25.36	45.04	64.57	83.24	93.44	98.01
ψ_{ATS}^\dagger	25	5.51	7.45	15.52	28.60	48.99	69.86	84.95	94.93	98.52
φ	50	9.31	14.09	31.97	58.54	81.95	95.36	99.33	99.92	100.00
φ^*	50	4.99	8.52	22.17	46.71	72.35	90.97	98.33	99.80	99.99
φ^\dagger	50	5.10	8.54	22.83	46.40	73.42	90.64	98.20	99.76	99.98
$\varphi_{\hat{Q}}^{MC}(ATS)$	50	7.11	11.23	26.88	53.04	78.03	93.75	99.14	99.87	100.00
$\varphi_{\hat{Q}}^*(ATS)$	50	5.02	8.02	21.58	45.88	72.20	91.34	98.25	99.77	100.00
$\varphi_{\hat{Q}}^\dagger(ATs)$	50	4.95	8.47	22.43	45.96	73.98	91.48	98.47	99.82	99.98
$\varphi_{\hat{Q}}^*(WTS)$	50	5.40	8.44	21.57	45.83	71.86	91.07	98.35	99.80	99.99
$\varphi_{\hat{Q}}^\dagger(WTS)$	50	5.18	8.54	22.21	45.29	72.68	90.71	98.29	99.78	99.98
ψ_{ATS}^*	50	4.87	9.76	26.71	55.99	82.13	95.94	99.55	99.98	100.00
ψ_{ATS}^\dagger	50	5.26	10.46	29.11	57.54	84.36	96.25	99.47	99.96	100.00
φ	100	6.86	16.67	50.10	84.79	98.53	99.98	100.00	100.00	100.00
φ^*	100	4.96	13.31	43.51	80.90	97.53	99.93	100.00	100.00	100.00
φ^\dagger	100	4.82	13.52	42.44	80.94	97.45	99.88	100.00	100.00	100.00
$\varphi_{\hat{Q}}^{MC}(ATS)$	100	6.04	15.32	47.32	83.40	98.00	99.91	100.00	100.00	100.00
$\varphi_{\hat{Q}}^*(ATS)$	100	4.93	13.23	43.62	81.04	97.63	99.92	100.00	100.00	100.00
$\varphi_{\hat{Q}}^\dagger(ATs)$	100	4.96	13.76	42.56	81.20	97.57	99.91	100.00	100.00	100.00
$\varphi_{\hat{Q}}^*(WTS)$	100	5.26	13.01	42.97	80.61	97.79	99.93	100.00	100.00	100.00
$\varphi_{\hat{Q}}^\dagger(WTS)$	100	5.04	13.23	41.64	80.93	97.64	99.93	100.00	100.00	100.00
ψ_{ATS}^*	100	4.64	15.96	53.29	89.11	99.23	99.99	100.00	100.00	100.00
ψ_{ATS}^\dagger	100	5.19	16.20	53.29	88.93	99.11	100.00	100.00	100.00	100.00

Table 10: Power of different test statistics under a shift-alternative for 3 groups with 5-dimensional observation vectors. The error terms are based on a skew normal distribution and have a compound symmetry covariance matrix with $\Sigma_1 = \Sigma_2 = \text{diag}(2, 3, 4, 5, 6) + \mathbf{1}_5 \mathbf{1}_5^\top$, resp. a autoregressive matrix $(\Sigma_3)_{\ell k} = 0.65^{|\ell-k|}$, while the groups are unbalanced with $n_1 = n_2 = 0.4 \cdot N$ and $n_3 = 0.2 \cdot N$.