Analysis of High-Dimensional One Group Repeated Measures Designs

Markus Pauly*, David Ellenberger** and Edgar Brunner**

May 19, 2014

Abstract

We propose a novel one sample test for repeated measures designs and derive its limit distribution for the situation where both the sample size n as well as the dimension d of the observations go to infinity. This covers the high-dimensional case with $d > n$. The tests are based on a quadratic form which involve new unbiased and dimension-stable estimators of different traces of the underlying unrestricted covariance structure. It is shown that the asymptotic distribution of the statistic may be standard normal, standardized χ^2-distributed, or even of weighted χ^2-form in some situations. To this end, we suggest an approximation technique which is asymptotically valid in the first two cases and provides an accurate approximation for the latter. We motivate and illustrate the application with a sleep lab data set and also discuss the practical meaning of $d \to \infty$ in case of repeated measures designs. It turns out that the limit behaviour depends on how the number of repeated measures is increased which is crucial for application.

Keywords: Approximations; High-dimensional Data; Quadratic Forms; Repeated Measures.

* Heinrich-Heine University of Düsseldorf, Mathematical Institute, Germany
  email: pauly@math.uni-duesseldorf.de

** University Medical Center Göttingen, Institute of Medical Statistics, Germany

1   Introduction

To motivate the methods derived in this paper, we present the sleep-laboratory trial reported by Jordan et al. (2004) where the activity of prostaglandin-D-synthase (β–trace) was observed on 10 young healthy men. The observations were taken at the time points 24h, 4h, 8h, 12h, 16h, and 20h for 3 consecutive nights and days. Thus, we have 18 repeated measures on 10 independent subjects. The data are displayed in Figure 1.
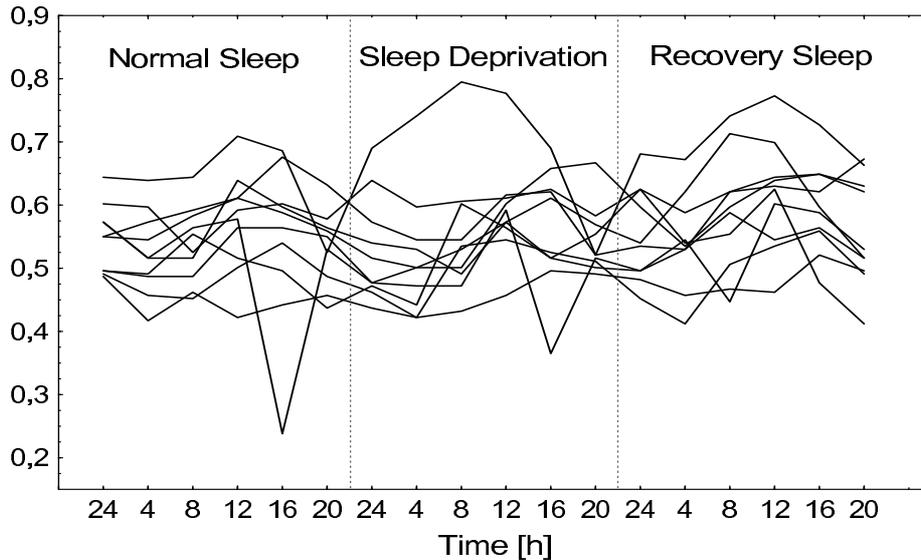
Figure 1: Prostaglandin-D-synthase (β-trace) levels in 10 healthy male subjects in a sleep-laboratory experiment during three days under different sleep conditions.

The questions to be answered in this trial are (1) whether the time profiles of the three interventions are flat or (2) if not, whether they are different, and (3) if the time profiles are different then it would be of interest to see whether the profile of the normal sleep differs from that of the sleep deprivation or from that of the recovery sleep.

In this experiment, we have 10 independent observation vectors $\boldsymbol{X}_k = (X_{k1}, \ldots, X_{kd})'$ of $k = 1, \ldots, n$ individuals with $d = 18$ repeated measures which are structured by two crossed fixed factors A (intervention) and B (time points). Factor A has 3 levels (normal sleep / sleep deprivation / recovery sleep) and factor B has 6 levels ($24h, 4h, 8h, 12h, 16h, 20h$). We do not assume any particular structure of the covariance matrix $\boldsymbol{V} = Cov(\boldsymbol{X}_k)$ since the results might be biased in case of a misspecification of the unknown structure of $\boldsymbol{V}$. Also, the patterns of the covariances within the three interventions might be different because of potential circadian rhythms which might be dependent on the interventions. Thus, the only reasonable option is to admit an unstructured $d \times d$ covariance matrix $\boldsymbol{V}$. As we have $d = 18$ repeated measures and

$n = 10$ independent subjects, the high-dimensional case $d > n$ invalidates the application of classical multivariate methods, such as Hotelling's T^2 or Wilk's Λ, see Davis (2002).

For a fixed number d of repeated measures many approximate procedures have been developed in the past to answer the questions listed above. These procedures are mainly based on Box's approximation (Box, 1954) and on different estimations of the so-called Box $\epsilon = \mathrm{tr}^2(\boldsymbol{V})\big/[(d-1)\,\mathrm{tr}(\boldsymbol{V}^2)]$ (Geisser and Greenhouse, 1958; Greenhouse and Geisser, 1959; Huynh and Feldt, 1976; Lecoutre, 1991; Werner, 2002, 2004; Ahmad, Werner, Brunner, 2008; Brunner, 2009; Brunner, Becker, and Werner, 2009; Chi et al., 2012).

These papers only discuss approximations which might become degenerate if the dimension $d \to \infty$. In repeated measures designs, the dimension can become quite large if the measurements are automatically recorded by some technical equipment which measures the observations within short time intervals. The accuracy of the approximations considered in the above quoted papers has only been investigated by simulations for small, medium, and moderate large sample sizes. The question of a limit distribution for $d \to \infty$ or for both $n, d \to \infty$ remained open until Bai and Saranadasa (1996) derived the limit distribution of a studentized quadratic form in a multivariate model under certain assumptions, *e.g.*, on the ratio d/n and on the eigenvalues $\lambda_1, \ldots, \lambda_d$ of the covariance matrix $\boldsymbol{V}$. In particular, they required that $\tau_{BS} = \lambda_{\max}\big/\sqrt{\sum_{i=1}^d \lambda_i^2} \to 0$ as $d \to \infty$, where $\lambda_{\max}$ denotes the largest eigenvalue. This seminal paper triggered many subsequent papers relaxing the assumptions and generalizing the ideas of Bai and Saranadasa to more general multivariate models (Srivastava, 2007 and 2009; Srivastava and Du, 2008; Chen and Qin, 2010; Park and Ayyala, 2013 and references cited therein). In particular, Chen and Qin (2010) derived the asymptotic normality of a studentized quadratic form under the assumption that the ratio $\tau_{CQ} = \mathrm{tr}(\boldsymbol{V}^4)\big/\mathrm{tr}^2(\boldsymbol{V}^2) \to 0$ for $d \to \infty$ allowing that n and d independently tend to ∞. We note that $0 \le \tau_{CQ} \le 1$, and that the above mentioned papers only investigated the asymptotic distribution of a studentized quadratic form if $\tau_{CQ} \to 0$. However, this condition will in general not be fulfilled if a random subject effect is present which is quite often the case in repeated measures models. This will be discussed in detail in Sections 3 and 4. Examples where τ_{CQ} might converge to some $\kappa \in (0, 1]$ are also presented there. Thus, it is also our aim to investigate the limit distribution for the case where τ_{CQ} does not converge to 0. It is worth to mention that the conditions $\tau_{BS} \to 0$ and $\tau_{CQ} \to 0$ as well as $\tau_{BS} \to 1$ and $\tau_{CQ} \to 1$ are equivalent. This is shown in the Appendix.

It should be noted that Bai and Saranadasa (1996), Srivastava (2007), and Chen and Qin (2010) derived procedures in a multivariate model where the observations in the different components may be measured on totally different measurement scales and thus, may not be commensurate. This means that the result of the analysis could be altered simply by changing the measurement scale in a single component, *e.g.*, from [min] to [sec]. Thus it is reasonable to studentize each component by the observed standard deviation of the observations within this component and then to work with the correlation matrix instead of the covariance matrix. This was first observed by Srivastava and Du (2008). Also the recent paper by Park and Ayyala (2013) improves the Chen-Qin Statistic in this regard. In models for repeated measures, however, the observations are typically measured on the same scale and thus, they are commensurate. Hence,

repeated measures models do not require a scaling of the observations. Scaling has the undesirable effect that in case of small or medium sample sizes, the variance of the statistic is inflated by involving additional d estimated quantities. In repeated measures designs it is therefore reasonable to consider statistics similar to those suggested by Bai and Saranadasa (1996) or by Chen and Qin (2010). We note that our considered statistic given in (2.11) is invariant under the same linear transformations of all components of the observation vectors.

Another issue what has not been considered in the above mentioned papers is the quality of the approximation by the standard normal distribution. In many practical situations in repeated measures designs, the number d of the dimension may be larger than the number n of independent vectors $\boldsymbol{X}_k$ but it may still not be large enough to provide an accurate approximation by a fixed limit distribution . As this is an important issue in applications, we will also investigate this problem in detail. In Section 3 we suggest to approximate the distribution of the studentized quadratic form in (2.11) by a sequence of distributions which for $n, d \to \infty$ converges to the correct limit distribution irrespective whether $\tau_{CQ} \to 0$ or $\tau_{CQ} \to 1$. In these asymptotic considerations we do not make any particular assumptions on the dependence between n and d.

To keep the presentation of the paper simple and not to overload it with lengthy derivations we restrict our considerations to the case of the multivariate normal distribution. Basically, the same results can be derived under weaker conditions. This is briefly discussed at the end of Section 7.

The main aims of this paper are

(1) the derivation of inference procedures for low- and high-dimensional one-sample repeated measures designs,

(2) the derivation of an accurate approximation for these procedures in case of small and medium sample sizes,

(3) the investigation of a statistic and the development of an approximation of its distribution which converges to the correct limit distribution irrespective whether $\tau_{CQ} \to 0$ or $\tau_{CQ} \to 1$, and finally,

(4) a detailed discussion of the practical meaning of $d \to \infty$.

Especially the last point is relevant in practice since the data set at hand must be regarded as representative for an asymptotic model. Typically, this asymptotic model should not depend on the way how the number of repeated measures is increased. It will turn out however, that this is not the case (see Section 4). This issue has not been investigated in the literature so far.

Moreover, we consider unbiased estimators of $\text{tr}(\boldsymbol{V}^2)$ and $\text{tr}(\boldsymbol{V}^3)$ needed for variance estimation (Werner, 2002) and to compute the degrees of freedom f of an approximation by a sequence of standardized χ_f^2-distributions (Ellenberger, 2011). These estimators are U-statistics which are shown to be ratio-consistent uniformly in the dimension d.

The paper is organized as follows. The statistical model, the hypotheses and the test statistic are introduced in Section 2 where also its limit behavior is discussed under different conditions. In Section 3, a new finite sample approximation is suggested. The meaning of letting $d \to \infty$ in the context of repeated measures designs is discussed in Section 4. Section 5 investigates the performance of the suggested approximation through simulation studies in different settings and the sleep-laboratory trial from Figure 1 is analyzed in Section 6. Finally, the results presented in this paper and in particular the conclusions in the context of repeated measures designs are discussed in Section 7. All proofs of the theoretical results are shifted to the Appendix whereas some technical details as well as additional simulation results are sourced out to the supplementary material of this paper (see Pauly, Ellenberger, and Brunner, 2013).

2   Statistical Model, Hypotheses and Statistics

We consider n independent d-dimensional random vectors

$$\boldsymbol{X}_k \quad = \quad (X_{k1}, \ldots, X_{kd})' \sim N(\boldsymbol{\mu}, \boldsymbol{V}), \ k = 1, \ldots, n, \tag{2.1}$$

where $\boldsymbol{\mu} = E(\boldsymbol{X}_1)$ and $\boldsymbol{V} = Cov(\boldsymbol{X}_1) > 0$. The components X_{ks} of the observation vector $\boldsymbol{X}_k$ of subject k refer to the treatments $s = 1, \ldots, d$ of the repeated measures which may be either time points or levels of a treatment which is allowed to have a factorial structure. In the motivating example we have 3 levels of an intervention (normal sleep / sleep deprivation / recovery sleep) which is crossed with the factor time involving 6 levels $(24h, 4h, 8h, 12h, 16h, 20h)$. Thus, we have $d = 3 \times 6 = 18$ repeated measures under different treatment combinations.

In many biological experiments it is reasonable to also include a random additive subject effect. Then the vector $\boldsymbol{X}_k$ of the observations is decomposed as

$$\boldsymbol{X}_k \quad = \quad (X_{k1}, \ldots, X_{kd})' \quad = \quad \boldsymbol{\mu} + cS_k \boldsymbol{1}_d + \boldsymbol{\epsilon}_k, \quad k = 1, \ldots, n, \tag{2.2}$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)'$ are the fixed effects of the repeated measures, $S_k \sim N(0, 1)$ denotes a random subject effect, and the error terms $\boldsymbol{\epsilon}_k = (\epsilon_{11}, \ldots, \epsilon_{kd})' \sim N(\boldsymbol{0}, \boldsymbol{A})$ are independent of S_k. The variance of the random subject effect can be modeled by the constant $c > 0$ and $c = 0$ refers to a model without a random subject effect. The covariance matrix $\boldsymbol{A}$ of $\boldsymbol{\epsilon}_k$ can be modeled in an appropriate way where different structures of $\boldsymbol{A}$ may be considered as reasonable depending on the experiment at hand. Such a structure may be the compound symmetry of $\boldsymbol{A}$, i.e., $\boldsymbol{A} = \sigma^2 \boldsymbol{I}_d + b\boldsymbol{J}_d$, where $\sigma^2 > 0$ and b are appropriate unknown constants, $\boldsymbol{I}_d$ denotes the d-dimensional unit matrix and $\boldsymbol{J}_d = \boldsymbol{1}_d \boldsymbol{1}_d'$ denotes the $d \times d$ matrix of 1s. Also, an autoregressive structure of $\boldsymbol{A} = (a_{ij})_{i,j=1,\ldots,d}$ may be assumed, i.e., $a_{ij} = \rho^{|i-j|}$ for $0 < \rho < 1$. Furthermore, a linearly decreasing symmetric Toeplitz structure of $\boldsymbol{A}$ can be modeled by $a_{ij} = 1 - |i - j|/d$. The most general form of $\boldsymbol{A}$ is an unstructured covariance matrix where it is only assumed that $\boldsymbol{A} = \boldsymbol{A}'$ and its elements a_{ij} fulfill the Cauchy-Schwarz inequality. Particular models will be discussed later in Sections 3 and 4.

5

In the context of biostatistical experiments, usually the hypotheses of no time effect, of no intervention effect, and of no interaction between time and intervention (parallelism or identity of the time curves) are of primary interest. Such hypotheses can be formulated by using an appropriate contrast matrix $\boldsymbol{H}$ and the most general formulation of a linear hypothesis about the expectation $\boldsymbol{\mu}$ is written as $H_0 : \boldsymbol{H}\boldsymbol{\mu} = \boldsymbol{0}$. For convenience, we will work with the unique projection matrix $\boldsymbol{T} = \boldsymbol{H}'(\boldsymbol{H}\boldsymbol{H}')^-\boldsymbol{H}$ where $(\cdot)^-$ denotes some generalized inverse of a matrix. We note that $H_0 : \boldsymbol{H}\boldsymbol{\mu} = \boldsymbol{0} \iff H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$ and that $\boldsymbol{T} = \boldsymbol{T}'$ and $\boldsymbol{T}^2 = \boldsymbol{T}$. Typical examples for testing hypotheses in factorial repeated measures designs are, *e.g.*, $\boldsymbol{T} = \boldsymbol{I}_d$ or $\boldsymbol{T} = \boldsymbol{P}_d = \boldsymbol{I}_d - \frac{1}{d}\boldsymbol{J}_d$ which correspond to testing whether the time profiles are equal or flat, respectively.

For testing the general linear hypothesis $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, we consider the quadratic form

$$Q_n = n\overline{\boldsymbol{X}}_{\cdot}'\boldsymbol{T}\overline{\boldsymbol{X}}_{\cdot}, \tag{2.3}$$

where $\overline{\boldsymbol{X}}_{\cdot}$ denotes the mean of the observation vectors $\boldsymbol{X}_k$, $k = 1, \ldots, n$. The classical approaches of Geisser and Greenhouse (1958), Greenhouse and Geisser (1959) or Huynh and Feldt (1976) depend on approximations of its scaled version $Q_n / \operatorname{tr}(\widehat{\boldsymbol{\Sigma}})$, where $\widehat{\boldsymbol{\Sigma}}$ is the empirical covariance matrix. However, these approximations become degenerate if $d \to \infty$ which may lead to inconsistent inference procedures. Hence, to study the asymptotic distribution of Q_n under $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, for $n, d \to \infty$ we consider the standardized version of Q_n, namely

$$\widetilde{W}_n = \frac{Q_n - E_{H_0}(Q_n)}{\sqrt{Var_{H_0}(Q_n)}}. \tag{2.4}$$

By using well-known theorems on the moments of quadratic forms (see, *e.g.*, Matthai and Provost, 1992) it follows that $E_{H_0}(Q_n) = \operatorname{tr}(\boldsymbol{\Sigma})$ and $Var_{H_0}(Q_n) = 2\operatorname{tr}(\boldsymbol{\Sigma}^2)$, where $\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{V}\boldsymbol{T}$. The asymptotic distribution of $\widetilde{W}_n$ under H_0 is established in the following theorem.

THEOREM 2.1 *Let λ_s denote the eigenvalues of $\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{V}\boldsymbol{T}$ in decreasing order and let $\lambda_{\max} = \max_s(\lambda_s) = \lambda_1$ as well as $a_s = \lambda_s / \sqrt{\sum_{r=1}^d \lambda_r^2}$. Then, under $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, the standardized quadratic form $\widetilde{W}_n$ in (2.4) has, asymptotically,*

(a) a standard normal distribution if

$$\tau_{BS} = a_1 = \to 0 \quad as \ n, d \to \infty, \tag{2.5}$$

(b) a standardized $(\chi_1^2 - 1)/\sqrt{2}$ distribution if

$$\tau_{BS} = a_1 \to 1 \quad as \ n, d \to \infty, \tag{2.6}$$

(c) a weighted χ^2-distribution of the form $\sum_{s=1}^{\infty} b_s(C_s^2 - 1)/\sqrt{2}$, where C_s^2 are independent χ_1^2-distributed random variables if in addition

$$a_s \to b_s \quad for \ all \quad s \quad as \ n, d \to \infty, \tag{2.7}$$

where b_s is a decreasing sequence in $[0, 1]$ fulfilling $\sum_s b_s^2 = 1$ and $b_s > 0$.

6

All proofs are given in the Appendix.

REMARKS 2.1

(a) Condition (2.5) has first been given by Bai and Saranandasa (1996). Later other conditions have been given by Srivastava (2007, 2009) while Chen and Qin (2010) have given the condition

$$\tau_{CQ} \;\;=\;\; \frac{\text{tr}(\boldsymbol{\Sigma}^4)}{\text{tr}^2(\boldsymbol{\Sigma}^2)} \;\;\to\;\; 0 \quad \text{for } d \to \infty. \tag{2.8}$$

It is shown in Lemma 8.1 in the Appendix that this condition (2.8) and the condition (2.5) are equivalent. Similar to Chen and Qin (2010), Park and Ayyala (2013) or Paindaveine and Verdebout (2013) we have not postulated any assumption on the rate d/n.
(b) Recently, Park and Ayyala (2013) have given a similar condition for the corresponding correlation matrix in a multivariate model. For repeated measures, however as discussed already in the introduction, it suffices to work with the covariance matrix $\boldsymbol{\Sigma}$.

To estimate the unknown two quantities $\text{tr}(\boldsymbol{\Sigma})$ and $\text{tr}(\boldsymbol{\Sigma}^2)$ under $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, we prefer to directly estimate the traces instead of first estimating the total covariance matrix and then computing the traces. Thus, we use the U-statistics B_0 and B_2 investigated by Werner (2002) and later considered in more detail by Brunner, Becker, and Werner (2009). The estimators B_0 and B_2 are based on the quadratic forms $A_k = \boldsymbol{X}_k'\boldsymbol{T}\boldsymbol{X}_k, k = 1, \ldots, n$, and on the bilinear forms $A_{k\ell} = \boldsymbol{X}_k'\boldsymbol{T}\boldsymbol{X}_\ell$, for $k \neq \ell$ and are given by

$$B_0 \;\;=\;\; \frac{1}{n} \sum_{k=1}^{n} A_k \qquad \text{and} \qquad B_2 \;\;=\;\; \frac{1}{n(n-1)} \sum_{k \neq \ell} A_{k\ell}^2. \tag{2.9}$$

We note that under $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, the U-statistics B_0 and B_2 are unbiased estimators of $\text{tr}(\boldsymbol{\Sigma})$ and $\text{tr}(\boldsymbol{\Sigma}^2)$, respectively. Moreover, they are L_2-ratio-consistent uniformly in the dimension d in the following sense

$$Var_{H_0}\left(\frac{B_0}{\text{tr}(\boldsymbol{\Sigma})}\right) \leq Var_{H_0}\left(\frac{B_0}{\sqrt{\text{tr}(\boldsymbol{\Sigma}^2)}}\right) \leq \frac{2}{n} \quad \text{and} \quad Var_{H_0}\left(\frac{B_2}{\text{tr}(\boldsymbol{\Sigma}^2)}\right) \leq \frac{8}{n-1}. \tag{2.10}$$

For a detailed derivation see, *e.g.*, Brunner, Becker, and Werner (2009).

Finally, using the result in (2.10) and Slutsky's Lemma we obtain the following theorem.

THEOREM 2.2 *Under the hypothesis $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, the statistic*

$$W_n \;\;=\;\; \frac{Q_n - B_0}{\sqrt{2B_2}} \tag{2.11}$$

has the same limit distributions as $\widetilde{W}_n$ under the different conditions $(a) - (c)$ given in Theorem 2.1.

The above results now motivate us to approximate the distribution of W_n not by one fixed limit distribution but by a sequence of distributions which provide adequate approximations in all three situation $(a) - (c)$. In particular, we will present an adequate approximation technique which covers the cases (a) and (b) in Theorem 2.2 and works also for a small number of samples n and dimensions d. For the case that $\tau_{BS} \to \kappa \in (0, 1)$ it provides a quite accurate three moments approximation in the same sense as suggested by Zhang (2005).

3   An Approximation for Small Samples

To investigate the quality and speed of the approximation by an $N(0, 1)$- or by the standardized $(\chi_1^2 - 1)/\sqrt{2}$ distribution, we demonstrate by a simple simulation that a large sample size n as well as a large dimension d are required to obtain a satisfactory approximation. To this end, we consider the simple design where $\boldsymbol{X}_k = (X_{k1}, \ldots, X_{kd})' \sim N(\boldsymbol{\mu}, \boldsymbol{V})$ with $\boldsymbol{V} = Cov(\boldsymbol{X}_1) > 0$. We investigate the behavior of W_n under the hypothesis $H_0 : \mu_1 = \cdots = \mu_d$, *i.e.* there is no effect of the repeated measures (*e.g.*, flat time profile) and $H_0 : \boldsymbol{\mu} = \boldsymbol{0}$ (*e.g.*, equal time profiles). In matrix notation the former hypothesis can be written as $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$ where $\boldsymbol{T} = \boldsymbol{I}_d - \frac{1}{d}\boldsymbol{J}_d$ is the so-called centering matrix while $\boldsymbol{T} = \boldsymbol{I}_d$ is used for the latter hypothesis. Again, $\boldsymbol{I}_d$ denotes the d-dimensional unit matrix and $\boldsymbol{J}_d = \boldsymbol{1}_d \boldsymbol{1}_d'$ denotes the $d \times d$ matrix of 1s.

We consider the cases of $n = 10, 20, 50$ for $d = 3, 5, 10, 20, 30, 50, 100, 200, 300, 500, 1000$ and $\boldsymbol{X}_k = \boldsymbol{\mu} + cS_k\boldsymbol{1}_d + \boldsymbol{\epsilon}_k$, $k = 1, \ldots, n$, as defined in (2.2). For the covariance matrix $\boldsymbol{A}$ we select an autoregressive structure $a_{ij} = \rho^{|i-j|}$ for $\rho = 0.6$. The choice of the constant $c = 0$ refers to a model without a random subject effect and $c = 1$ refers to a model involving a random subject effect with variance 1. Thus, $\boldsymbol{V} = \boldsymbol{A} + c\boldsymbol{J}_d$. Then for $\boldsymbol{T} = \boldsymbol{I}_d$ it follows that $\tau_{BS} \to 0$ for $c = 0$ while $\tau_{BS} \to 1$ for $c = 1$, both as $d \to \infty$. (The derivation is given in Example 1.1 in the supplementary material). The simulated type-I error rates in these settings when compared with the two-sided 95%-quantile of the standard normal ($\tau_{BS} \to 0$) and the $(\chi_1^2 - 1)/\sqrt{2}$ distribution ($\tau_{BS} \to 1$) are displayed in Figure 2.

From Figure 2 it appears that at least in this particular case the quality of the approximations is quite bad for a small or even a medium number of sample sizes and dimensions. Therefore, we intend to develop a better approximation which is also valid for a small sample size n and a small number of dimensions d. The derivation of this different approximation is based on the fact that under the hypothesis $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, the quadratic form Q_n in (2.3) is equal to $\sum_{s=1}^d \lambda_s C_s^2$, where the quantities $C_s^2 \sim \chi_1^2$ are independent and the λ_s are the eigenvalues of $\boldsymbol{\Sigma} = \boldsymbol{TVT}$. It is well known that the distribution of $\sum_{s=1}^d \lambda_s C_s^2$ can be approximated by a scaled χ^2-distribution, *e.g.* by $g \cdot \chi_f^2$ such that the first two moments coincide (cf. Box, 1954, or Geisser and Greenhouse, 1958, and Greenhouse and Geisser, 1959). This leads to approximating the distribution of $Q_n / \operatorname{tr}(\boldsymbol{\Sigma})$ by a χ_f^2/f-distribution where $f = \operatorname{tr}^2(\boldsymbol{\Sigma})/\operatorname{tr}(\boldsymbol{\Sigma}^2)$ is proportional to Box's ϵ. However, as already explained in Section 2, χ_f^2/f converges to a one-point distribution if $f \to \infty$ as $n, d \to \infty$. But this holds under condition (2.5), so that we cannot use this approach.
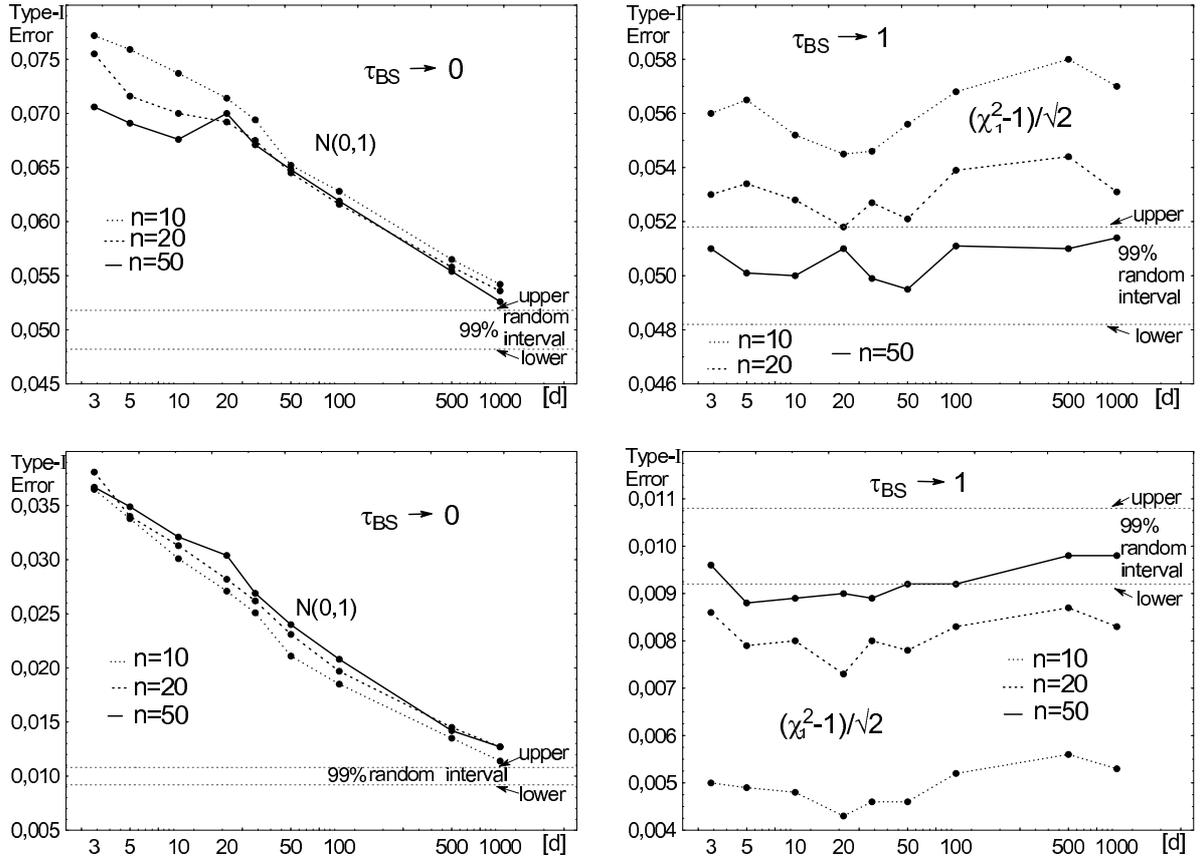
Figure 2: Simulated type-I error rates of the Statistic W_n in 10^5 simulation runs when compared with the asymptotic standard normal distribution ($\tau_{BS} \to 0$) and the asymptotic $(\chi_1^2 - 1)/\sqrt{2}$ distribution ($\tau_{BS} \to 1$) for different dimensions and sample sizes $n = 10, 20, 50$. The horizontal grid lines denote the 99% random interval for $\alpha = 5\%$ (upper row) and $\alpha = 1\%$ (lower row).

Therefore, similar to Bai and Saranadasa (1996) or Chen and Qin (2010) a standardized version $\widetilde{W}_n$ of Q_n in (2.4) is used as a statistic. But unlike as in those papers we do not approximate this standardized statistic by a fixed limit distribution but consider a sequence of standardized χ^2-distributions to develop a more accurate approximation of the distribution of W_n under $H_0 : \boldsymbol{T\mu} = \boldsymbol{0}$. In particular, applying the representation theorem for quadratic forms (which is also valid if $\boldsymbol{\Sigma}$ is singular) we re-write $\widetilde{W}_n$ as

$$\widetilde{W}_n = \frac{Q_n - \operatorname{tr}(\boldsymbol{\Sigma})}{\sqrt{2\operatorname{tr}(\boldsymbol{\Sigma}^2)}} = \frac{Q_n - \sum_{s=1}^d \lambda_s}{\sqrt{2\sum_{s=1}^d \lambda_s^2}} = \sum_{s=1}^d a_s \frac{C_s^2 - 1}{\sqrt{2}}, \qquad (3.12)$$

where $a_s = \lambda_s / \sqrt{\sum_{r=1}^d \lambda_r^2}$. Thus, we suggest to approximate the distribution of $\widetilde{W}_n$ by a

9

sequence of standardized χ^2-distributions

$$K_f \;=\; \frac{\chi_f^2 - f}{\sqrt{2f}}. \tag{3.13}$$

It is shown in Zhang (2005, Theorem 1) that even in case (a) of Theorem 2.1 this should be pre-ferred compared to the normal approximation if f is selected such that the first three moments of $\widetilde{W}_n$ and K_f coincide. This is achieved by choosing

$$f_P = (\tau_P)^{-1} \;=\; \frac{\text{tr}^3(\boldsymbol{\Sigma}^2)}{\text{tr}^2(\boldsymbol{\Sigma}^3)}, \tag{3.14}$$

see formula (5) in Zhang (2005). We call this $f = f_P$ the Pearson approximation, see Pearson (1959), Imhof (1961) as well as Katayama, Kano, and Srivastava (2013). We note that the condition $\tau_P \to 0$ is equivalent to the condition given in (2.5) and to the condition $\tau_{CQ} = \text{tr}(\boldsymbol{\Sigma}^4)/\text{tr}^2(\boldsymbol{\Sigma}^2) \to 0$ given by Chen and Qin (2010) and adapted to the one-sample repeated measures design considered in this paper. Moreover, the equivalences $\tau_{BS} \to 1 \iff \tau_P \to 1 \iff \tau_{CQ} \to 1$ are also true, see Lemma 8.1 in the Appendix. From this it follows that the K_f-approximation is also asymptotically correct in case (b) of Theorem 2.1. Moreover, in case (c) of this theorem it provides a good small sample approximation as discussed in Zhang (2005), see also Zhang and Liang (2014) for the case of a two moment approximation. For the latter two moment approach one would obtain $f_B = \tau_B^{-1} = \text{tr}^2(\boldsymbol{\Sigma})/\text{tr}(\boldsymbol{\Sigma}^2)$. For our asymptotic con-siderations, however, the conditions in (2.5) or (2.6) have to be verified. Simple computations show that $1/f_B \leq \lambda_{max}/\sqrt{\sum_{s=1}^d \lambda_s^2}$ such that (2.5) implies that $f_B \to \infty$. On the other hand, however, $f_B \to \infty$ does not imply (2.5) as can be seen by the simple counter example where $\lambda_{max} = \sqrt{d}$ and all other eigenvalues $\lambda_s \equiv 1$. Also the condition $\tau_B \to 1$ is not equivalent to condition (2.6). This is easily seen by letting $\boldsymbol{V} = \boldsymbol{A} + c\boldsymbol{J}_d$, where $\boldsymbol{A}$ has an autoregressive structure and $\boldsymbol{T} = \boldsymbol{I}_d$. In this case $\tau_{BS} \to 1$ but $\tau_B \to c^2/(c+1)^2$ as shown in Example 1.1(b) in the supplementary material to this paper, see Pauly, Ellenberger, and Brunner (2014). Thus, we only consider the choice $f = f_P$.

In order to apply the approximation suggested by Zhang (2005) a consistent estimator of τ_P is required. For the high-dimensional case considered in the present paper we prefer to directly estimate the trace $\text{tr}(\boldsymbol{\Sigma}^3)$ similar to the derivation of B_2 in Section 2. Let $A_{k\ell}$ be as given prior to (2.9). Then,

$$B_3 \;=\; \binom{n}{3}^{-1} \sum_{k<\ell<r} A_{k\ell} A_{\ell r} A_{rk} \tag{3.15}$$

is an unbiased estimator of $\text{tr}(\boldsymbol{\Sigma}^3)$. Moreover, it is shown in Theorem 8.2 in the Appendix that B_3^2/B_2^3 is a consistent estimator of $\tau_P = 1/f_P$ uniformly in the dimension d. We note that it is not reasonable to consider the ratio-consistency of B_2^3/B_3^2 for $f_P = 1/\tau_P$ since $1/\tau_P \to \infty$ under condition (2.5). Thus we consider the inverse τ_P of f_P since $0 \leq \tau_P \leq 1$. We summarize the foregoing considerations in the following theorem.

THEOREM 3.1 *Under the hypothesis* $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$*, the sequence of standardized* χ^2*-distributions* K_{f_P} *defined in (3.13) has, asymptotically,*

(a) a standard normal distribution under condition (2.5),

(b) a standardized $(\chi_1^2 - 1)/\sqrt{2}$ *distribution under condition (2.6).*

The same statements remain true if f_P *is estimated by*

$$\widehat{f}_P \;\;=\;\; B_2^3/B_3^2 \tag{3.16}$$

where B_2 *and* B_3 *are given in (2.9) and (3.15), respectively.*

REMARK 3.1 *(a) The simulations shown in Section 5 demonstrate that this approximation technique performs quite well in all considered situations covering the cases* $(a) - (c)$ *of Theorem 2.1.*
(b) Another approach would be to plug in consistent estimates of the unknown standardized eigenvalues a_s *in (3.12) and to simulate from this distribution. In a high-dimensional setting, however, the consistent estimation of the eigenvalues requires additional assumptions (see e.g. Yata and Aosima, 2012). Since regarding the structure of the covariance matrix we like to be as general as possible we will not consider this approach. Moreover, in practice it is unclear whether such conditions might hold. Also, bootstrap methods will not be used as they are extremely time consuming (see e.g. Zhang and Liang, 2014). Furthermore, it is unclear how to bootstrap our statistic in such a way that all different limit distributions are covered.*

4   Some Considerations on $d \to \infty$

The classical assumption that the sample size $n \to \infty$ simply means that more observations are sampled while maintaining the same conditions under which the observations are obtained. This means that the n observations in the experiment at hand are considered as representative for a theoretical experiment with an infinite number of observations. The dimension d, however, is a fixed quantity of a particular experiment. Letting $d \to \infty$ may therefore be interpreted as considering the experiment at hand as representative for an experiment with an infinite number of dimensions. In a repeated measures design, where the repeated measurements represent time points, for example, the number of time points can be increased in several ways.

First, one may observe more time points by extending the observation time without changing the distances between the time points of the measurements. For an autoregressive model with $\boldsymbol{A} = (a_{ij})_{i,j=1,\ldots,d}$, where $a_{ij} = \rho^{|i-j|}$ for $\rho \in (0,1)$, this means that $\boldsymbol{A}$ becomes larger when d is increased without changing the elements of the basic $d \times d$ matrix, *i.e.* for $d' > d$ only rows and columns are added where the elements are still given by $a_{ij} = \rho^{|i-j|}$. In this case, $\tau_{CQ} = \mathrm{tr}^2(\boldsymbol{A}^2)/\mathrm{tr}(\boldsymbol{A}^4) \to 0$ or equivalently, $\tau_P \to 0$ and $\tau_{BS} \to 0$, so that W_n has, asymptotically, a

standard normal distribution. This follows from Example 1.1(a) in the supplementary material together with Theorem 2.11 and Lemma 8.1.

The experimenter, however, may argue that s/he is not interested in observing a longer time, and that only the time between the first and the last observation is of practical relevance. In this situation, the number of time points d could only be increased by decreasing the distances between the time points. To describe this with an autoregressive structure, it is required that the correlation between two successive time points increases with increasing d, whereas the correlation between the first and the last time point remains the same. This may be modeled by an autoregressive matrix $\boldsymbol{A}$ with $a_{ij} = \rho^{|i-j|/(d-1)}$. Note that in this setting the correlation between the first and the last measurement is given by $a_{1,d} = \rho$, which does not depend on d. In this setting, however, it is shown in Example 1.2 in the supplementary material that τ_B converges to a constant $g(\rho)$ which only depends on ρ and is bounded away from 0 and 1 for each choice of $\rho \in (0,1)$. Since $\tau_B \leq \tau_P \leq 1$ by Cauchy-Schwarz inequality, this implies that neither τ_P nor τ_{CQ} nor τ_{BS} converge to 0 as in the former case with increasing observation times but fixed distances. For example, we have $g(0.6) \approx 0.73$ which corresponds to the simulations from Table 2 in Section 5. Now, in view of Theorems 2.2, this means that the asymptotic distribution of the test statistic depends on how d tends to ∞, $i.e.$, on how the number of time points is increased.

A similar problem appears when adding an additive subject effect S_k as considered in model (2.2). It is easily seen that a random subject effect may change the validity of the assumptions that $\tau_{BS} \to 0$ or $\tau_{BS} \to 1$ in (2.5) and (2.6) depending on the hypothesis matrix $\boldsymbol{T}$. For a multivariate hypothesis, for example, we use the hypothesis matrix $\boldsymbol{T} = \boldsymbol{I}_d$ and thus, $\boldsymbol{\Sigma} = \boldsymbol{TVT} = \boldsymbol{T}(\boldsymbol{A} + c^2\boldsymbol{J}_d)\boldsymbol{T} = \boldsymbol{A} + c^2\boldsymbol{J}_d$. Now let $\boldsymbol{A} = \sigma^2\boldsymbol{I}_d$ and $c = 0$. Then $\tau_{CQ} = \mathrm{tr}(\boldsymbol{\Sigma}^4)/\mathrm{tr}^2(\boldsymbol{\Sigma}^2) = \frac{1}{d} \to 0$ as $d \to \infty$ and thus, (2.5) holds. If $c = 1$, however, then $\tau_{CQ} = [d^4 + \mathcal{O}(d^3)]/[d^4 + \mathcal{O}(d^3)] \to 1$ as $d \to \infty$ and thus, (2.6) holds. For a hypothesis involving the repeated measures, $e.g.$, $\boldsymbol{T} = \boldsymbol{P}_d$ it follows that $\boldsymbol{\Sigma} = \boldsymbol{P}_d\boldsymbol{A}\boldsymbol{P}_d$ both for $c = 0$ and $c = 1$ since $\boldsymbol{P}_d\boldsymbol{J}_d = \boldsymbol{0}$. Then, $\tau_{CQ} = \frac{1}{d-1} \to 0$ as $d \to \infty$.

This means that the limits of τ_{CQ} or τ_{BS} for $d \to \infty$ depend on whether or not a random subject effect is present as well as on the way how the number of time points is increased. But all information we have in a practical experiment is the fixed sample at hand. Thus, τ_{CQ} or τ_{BS} may converge to 0, to 1, or to some $\kappa \in (0,1)$. In practice where the correct statistical model is unknown we do not know to which limits τ_{CQ} or τ_{BS} converge for $d \to \infty$. Thus it is not advisable to approximate the distribution of W_n by a fixed limit distribution such as the standard normal or the standardized $(\chi^2 - 1)/\sqrt{2}$ distribution. In contrast, our approach based on approximating the distribution of W_n by a sequence of standardized χ^2-distributions given by K_{f_P} in (3.13) seems to be more appropriate since, by Theorem 3.1, it accounts for the different asymptotic limit distributions of W_n as $d \to \infty$.

5 Simulations

In this section we present some simulation results for different structures of the covariance matrix $\boldsymbol{A} = (a_{ij})_{i,j=1,\dots,d}$ where we report the results for

(1) an autoregressive structure: AR(ρ) for $\rho = 0.6$, *i.e.*, $a_{ij} = \rho^{|i-j|}$,

(2) an autoregressive structure where the correlation depends on the dimension: AR(d, ρ) for $\rho = 0.6$, *i.e.*, $a_{ij} = \rho^{|i-j|/(d-1)}$,

(3) a compound symmetry structure: CS, *i.e.*, $\boldsymbol{A} = \boldsymbol{I}_d + \boldsymbol{J}_d$,

(4) a linearly decreasing (symmetric) Toeplitz structure TO, *i.e.*, $a_{ij} = 1 - |i-j|/d$.

We compare the type-I errors of the W_n statistic in (2.11) when approximated with the K_f-distribution based on $f = f_P$ and, if $\tau_{BS} \to 0$, also with its asymptotic standard normal distribution as well with the Bai and Saranadasa (1996) statistic. The results for the latter statistic are not shown here since the behaviour is similar to that for the statistic W_n with a fixed normal approximation. Moreover, a comparison with the inference procedure proposed in Ahmad, Werner and Brunner (2008) is not appropriate since their statistic becomes degenerate for $\tau_{BS} \to 0$.

5.1 Data Generation

The independent vectors $\boldsymbol{X}_k$ were generated as $\boldsymbol{X}_k = \boldsymbol{\mu} + cS_k \boldsymbol{1}_d + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}_k \sim N(\boldsymbol{0}, \boldsymbol{A})$ and $\boldsymbol{\mu} = \boldsymbol{0}$. A random subject effect $S_k \sim N(0,1)$ independent of $\boldsymbol{\epsilon}_k$ was included ($c = 1$) in the model or not ($c = 0$). Thus, the covariance matrix $\boldsymbol{V}$ of $\boldsymbol{X}_k$ is given by $\boldsymbol{V} = \boldsymbol{A} + c^2 \boldsymbol{J}_d$.

We present simulations for two types of hypotheses $H_0 : \boldsymbol{T\mu} = \boldsymbol{0}$ distinguishing between a multivariate question by letting $\boldsymbol{T} = \boldsymbol{I}_d$ and a question about the equality of the repeated measures by letting $\boldsymbol{T} = \boldsymbol{P}_d = \boldsymbol{I}_d - \frac{1}{d}\boldsymbol{J}_d$. We note that the cases $\boldsymbol{\Sigma} = \boldsymbol{P}_d \boldsymbol{V} \boldsymbol{P}_d = \boldsymbol{P}_d \boldsymbol{A} \boldsymbol{P}_d$ for $c = 0$ and $c = 1$ are identical since $\boldsymbol{P}_d \boldsymbol{J}_d = \boldsymbol{0}$.

To demonstrate the accuracy of the approximation by the $(\chi_f^2 - f)/\sqrt{2f}$-distribution for $f = f_P = 1/\tau_P$, we reconsider the example in Section 3 and let $\boldsymbol{T} = \boldsymbol{P}_d$ and $\boldsymbol{V} = \boldsymbol{A} + \boldsymbol{J}_d$, where $\boldsymbol{A}$ has an autoregressive structure with $\rho = 0.6$ as in (1). In Figure 3 we have added the simulated type-I errors from $N_{\text{sim}} = 10^5$ simulation runs displayed as solid lines for the approximation by the $(\chi_{f_P}^2 - f_P)/\sqrt{2f_P}$-distribution compared with an approximation by the $N(0,1)$-distribution displayed as dotted lines ($n = 10, 20, 50$). We note that in this case, $\tau_{BS} \to 0$ and $\tau_P \to 0$ as $d \to \infty$ as already discussed in Section 3 since $\tau_{BS} \to 0 \iff \tau_P \to 0 \iff \tau_{CQ} \to 0$.

From Figure 3 it is obvious that the approximation by the sequence of the $(\chi_{f_P}^2 - f_P)/\sqrt{2f_P}$-distributions is quite accurate, even for small d and $n \geq 20$. For $n = 10$ it is slightly liberal for $d \leq 30$ while the approximation by the standard normal distribution is considerably liberal for all n and $d \leq 500$.

13
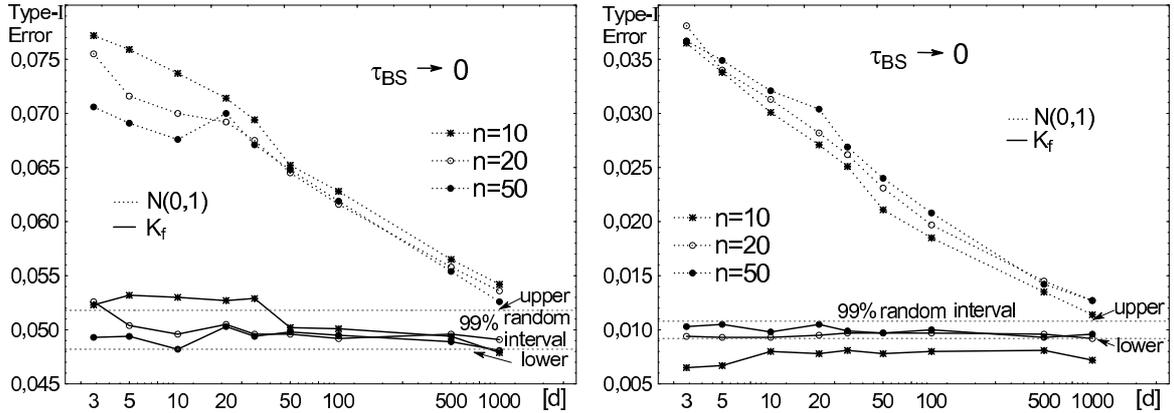
Figure 3: Simulated type-I error rates of the Statistic W_n when compared with the K_f-approximation based on $f = f_P$ and the asymptotic standard normal distribution ($\tau_P \to 0$) for different dimensions and sample sizes $n = 10, 20, 50$: similar as in Figure 2. The horizontal grid lines denote the 99% random interval for $\alpha = 5\%$ (left) and $\alpha = 1\%$ (right), respectively.

The simulations for the other settings of covariance matrices and hypotheses showed a similar quality of the approximation. All results are displayed in Tables 3 - 4 for the approximations by the sequence of the $(\chi^2_{f_P} - f_P)/\sqrt{2f_P}$-distributions. For brevity, we only show the results for $n = 20$ and note that for $n = 50$ the approximation was equally accurate. For the extremely small sample size of $n = 10$, the results are a bit less accurate, but still acceptable. The simulated 5% level ranged from 0.0471 to 0.0590 while the 1% level ranged from 0.0041 to 0.0091. The detailed results are shown in the supplementary material, see Pauly, Ellenberger, and Brunner (2013). A slight liberality of the test based on W_n in (2.11) and the approximation by the $(\chi^2_{f_P} - f_P)/\sqrt{2f_P}$-distribution was corrected by multiplying W_n with $\sqrt{(n-1)/n}$ which leads to the recommendation using $W_n\sqrt{(n-1)/n}$ as a statistic for testing the hypothesis $H_0 : \boldsymbol{T\mu} = \boldsymbol{0}$ in a high-dimensional one-group repeated measures design and approximating its distribution by the $(\chi^2_{f_P} - f_P)/\sqrt{2f_P}$-distribution. This is a similar correction as used for the F-approximation in the case of a two moment matching, see Brunner, Becker and Werner (2009) for a theoretical argument.

Table 1 Simulated type-I error rates for an autoregressive structure AR(ρ) for $\rho = 0.6$ where $\boldsymbol{A} = (a_{ij})_{i,j=1,\ldots,d}$ and $a_{ij} = \rho^{|i-j|}$.

	AR(0.6): $\boldsymbol{\Sigma} = \boldsymbol{T}(\boldsymbol{A} + c\boldsymbol{J}_d)\boldsymbol{T}$, $a_{ij} = \rho^{|i-j|}$								
	$c = 0$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 1$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 0$ and $\boldsymbol{T} = \boldsymbol{P}_d$		
d	τ_P	5%	1%	τ_P	5%	1%	τ_P	5%	1%
3	0.759	0.0509	0.0092	0.946	0.0530	0.0104	0.648	0.0523	0.0094
5	0.594	0.0497	0.0093	0.916	0.0537	0.0099	0.489	0.0502	0.0093
10	0.359	0.0502	0.0098	0.870	0.0530	0.0104	0.338	0.0490	0.0093
20	0.193	0.0490	0.0095	0.876	0.0521	0.0098	0.199	0.0507	0.0095
30	0.131	0.0495	0.0094	0.890	0.0529	0.0102	0.137	0.0499	0.0097
50	0.080	0.0496	0.0092	0.916	0.0522	0.0099	0.083	0.0492	0.0097
100	0.041	0.0501	0.0096	0.949	0.0539	0.0100	0.041	0.0490	0.0097
500	0.008	0.0496	0.0095	0.988	0.0554	0.0105	0.008	0.0492	0.0089
1000	0.004	0.0498	0.0096	0.994	0.0523	0.0097	0.004	0.0483	0.0091

Table 2 Simulated type-I error rates for an autoregressive structure AR(d, ρ) for $\rho = 0.6$ where the correlation depends on the dimension, *i.e.*, $\boldsymbol{A} = (a_{ij})_{i,j=1,\ldots,d}$ and $a_{ij} = \rho^{|i-j|/(d-1)}$.

	AR(d, 0.6): $\boldsymbol{\Sigma} = \boldsymbol{T}(\boldsymbol{A} + c\boldsymbol{J}_d)\boldsymbol{T}$, $a_{ij} = \rho^{|i-j|/(d-1)}$								
	$c = 0$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 1$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 0$ and $\boldsymbol{T} = \boldsymbol{P}_d$		
d	τ_P	5%	1%	τ_P	5%	1%	τ_P	5%	1%
3	0.920	0.0530	0.0094	0.982	0.0530	0.0099	0.710	0.0510	0.0068
5	0.949	0.0510	0.0095	0.989	0.0537	0.0101	0.691	0.0508	0.0182
10	0.962	0.0547	0.0101	0.992	0.0533	0.0102	0.731	0.0520	0.0085
20	0.966	0.0548	0.0106	0.993	0.0541	0.0099	0.749	0.0515	0.0186
30	0.968	0.0545	0.0102	0.993	0.0533	0.0102	0.753	0.0519	0.0182
50	0.968	0.0552	0.0102	0.993	0.0529	0.0096	0.755	0.0524	0.0184
100	0.969	0.0520	0.0100	0.993	0.0526	0.0104	0.756	0.0513	0.0082
500	0.970	0.0523	0.0095	0.993	0.0538	0.0103	0.757	0.0519	0.0102
1000	0.970	0.0544	0.0103	0.993	0.0531	0.0099	0.757	0.0509	0.0098

Tables 1-4 display the simulated type-I error rates (5% and 1% levels) for different covariance structures for the hypothesis $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$, where $\boldsymbol{T} = \boldsymbol{I}_d$ for $c = 0$ or $c = 1$, and $\boldsymbol{T} = \boldsymbol{P}_d$ for $c = 0$. For a sample size of $n = 20$, the statistic $W_n\sqrt{(n-1)/n}$ is compared with the $(\chi^2_{f_P} - f_P)/\sqrt{2f_P}$-distribution in $N_{\text{sim}} = 10^5$ simulation runs. The parameter f_P is estimated by $\widehat{f}_P = B_2^3/B_3^2$ where B_2 and B_3 are given in (2.9) and (3.15), respectively.

Table 3 Simulated type-I error rates for a compound symmetry structure CS, *i.e.*, $\boldsymbol{A} = \boldsymbol{I}_d + \boldsymbol{J}_d$.

	CS: $\boldsymbol{\Sigma} = \boldsymbol{T}^2 + c\boldsymbol{T}\boldsymbol{J}_d\boldsymbol{T}$								
	$c = 0$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 1$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 0$ and $\boldsymbol{T} = \boldsymbol{P}_d$		
d	τ_P	5%	1%	τ_P	5%	1%	τ_P	5%	1%
3	0.333	0.0495	0.0070	0.747	0.0557	0.0092	0.500	0.0505	0.0083
5	0.200	0.0494	0.0075	0.756	0.0561	0.0103	0.250	0.0496	0.0082
10	0.100	0.0479	0.0079	0.817	0.0573	0.0097	0.111	0.0492	0.0081
20	0.050	0.0495	0.0085	0.885	0.0562	0.0104	0.053	0.0483	0.0088
30	0.033	0.0490	0.0085	0.916	0.0580	0.0101	0.034	0.0489	0.0083
50	0.020	0.0493	0.0084	0.946	0.0571	0.0101	0.020	0.0484	0.0083
100	0.010	0.0486	0.0087	0.972	0.0580	0.0102	0.010	0.0487	0.0093
500	0.002	0.0485	0.0091	0.994	0.0538	0.0096	0.002	0.0497	0.0090
1000	0.001	0.0490	0.0092	0.997	0.0539	0.0105	0.001	0.0497	0.0089

Table 4 Simulated type-I error rates for a linearly decreasing (symmetric) Toeplitz structure TO where $\boldsymbol{A} = (a_{ij})_{i,j=1,\ldots,d}$ and $a_{ij} = 1 - |i - j|/d$.

	TO: $\boldsymbol{\Sigma} = \boldsymbol{T}(\boldsymbol{A} + c\boldsymbol{J}_d)\boldsymbol{T}$, $a_{ij} = 1 - |i - j|/d$								
	$c = 0$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 1$ and $\boldsymbol{T} = \boldsymbol{I}_d$			$c = 0$ and $\boldsymbol{T} = \boldsymbol{P}_d$		
d	τ_P	5%	1%	τ_P	5%	1%	τ_P	5%	1%
3	0.563	0.0509	0.0089	0.782	0.0513	0.0092	0.784	0.0519	0.0096
5	0.678	0.0517	0.0105	0.792	0.0515	0.0100	0.771	0.0520	0.0101
10	0.750	0.0521	0.0098	0.799	0.0523	0.0101	0.800	0.0534	0.0101
20	0.779	0.0521	0.0099	0.801	0.0518	0.0096	0.812	0.0513	0.0097
30	0.787	0.0520	0.0099	0.802	0.0529	0.0103	0.814	0.0525	0.0102
50	0.793	0.0527	0.0096	0.802	0.0518	0.0103	0.816	0.0522	0.0105
100	0.798	0.0522	0.0100	0.802	0.0507	0.0098	0.816	0.0514	0.0093
500	0.801	0.0514	0.0095	0.802	0.0536	0.0102	0.816	0.0513	0.0100
1000	0.802	0.0525	0.0093	0.802	0.0516	0.0097	0.816	0.0530	0.0106

Additional simulation results for the W_n statistic under different alternatives have been carried out in Ellenberger (2011) and are omitted for brevity. The main result is that the power depends on the type of the alternative. For a one point alternative the power becomes larger with decreasing dimension d, whereas for a trend alternatives it becomes larger with increasing dimension d.

6 Analysis of the Data Example

Finally, we analyze the sleep lab example from Section 1 by means of the above presented methods based on the test statistic W_n. Figure 1 shows the Prostaglandin-D-synthase levels of

$n = 10$ independent observations from healthy male subjects with $d = 18$ time points through 3 consecutive nights. Here we like to examine the following null hypotheses:

- $H_0(A)$: There is no factor effect A (intervention)

- $H_0(B)$: There is no factor effect B (time points)

- $H_0(AB)$: There is no interaction effect between intervention and time.

The results obtained by the statistic $W_n = \sqrt{n-1}(Q_n - B_0)/\sqrt{2nB_2}$ based on an approximation by the suggested sequence of the standardized χ^2-distributions K_{f_P} are shown in Table 5.

Table 5 Analysis of the sleep lab data from Figure 1: Shown are the p-values of the W_n-test based on the approximation by the sequence K_{f_P} of standardized χ^2-distributions ($W_n \sim K_f$).

Hypothesis	W_n	df	p-values
$H_0(A)$	1.862	1.29	0.0572
$H_0(B)$	3.420	5.00	0.0094
$H_0(AB)$	-0.706	10.00	0.7401

A highly significant time effect is detected while there is no evidence for an interaction. The result for the intervention is borderline. It may be noted that in Ahmad, Werner and Brunner (2008) the data for the female patients are analyzed, whereas here we analyze the data for the male patients.

7 Conclusions and Discussion

We have considered a one-group high-dimensional repeated measures setting where a factorial structure of the repeated measures is allowed. Assuming a multivariate normal distribution of the repeated measures, procedures for testing the common linear hypotheses about the expectations have been derived. We did not assume any particular structure of the covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{TVT}$ since a misspecification may lead to biased results. If the number n of the independent observation vectors is larger than the dimension d of these vectors then the classical procedures of one-group profile analysis can be applied. If, however, d is larger than n then these procedures are no longer valid. In this case, the quadratic form Q_n in (2.3) is used to test a global hypothesis $H_0 : \boldsymbol{T\mu} = \boldsymbol{0}$ about the factors involved in this repeated measures design. The scaled version of this quadratic form $Q_n/\operatorname{tr}(\boldsymbol{\Sigma})$, as suggested by Box (1954) and Geisser and Greenhouse (1958), cannot be used for asymptotic considerations since it becomes degenerate for $d \to \infty$. Therefore, the standardized version $\widetilde{W}_n = (Q_n - \operatorname{tr}(\boldsymbol{\Sigma}))/\sqrt{2\operatorname{tr}(\boldsymbol{\Sigma}^2)}$ of Q_n had been considered in the literature. Its asymptotic distribution for $n, d \to \infty$ had been derived by several authors under different assumptions on the eigenvalues of the covariance matrix $\boldsymbol{\Sigma}$. In practical data analysis, however, these assumptions are not always fulfilled and

moreover, they are difficult to verify. Therefore, it was our aim to study the behavior of $\widetilde{W}_n$ (or of its studentized form W_n) if these assumptions do not hold. It turned out that W_n can have different asymptotic distributions depending on the eigenvalues of Σ and it was discussed that these distributions may differ according as a random subject effect is included or not and on how the number of time points, *i.e.*, on how $d \to \infty$, is increased. Thus, the asymptotic distribution of W_n cannot be approximated by one distribution (as a standard normal or a standardized $(\chi_1^2 - 1)/\sqrt{2}$-distribution) since this may lead to inaccurate results. Instead we have suggested to approximate the asymptotic distribution by the sequence $K_f = (\chi_f^2 - f)/\sqrt{2f}$ of standardized χ^2-distributions. The parameter f was chosen such that mean, variance and skewness of K_f and $\widetilde{W}_n$ coincide. The unknown parameter f is estimated by $f_P = B_2^3/B_3^2$, where B_3 is a U-statistic for the parameter $\mathrm{tr}(\Sigma^3)$. For the derivation of the possible asymptotic distributions, the ratio-consistency of the involved unbiased estimators B_0, B_2, and B_3 for the traces $\mathrm{tr}(\Sigma)$, $\mathrm{tr}(\Sigma^2)$, and $\mathrm{tr}(\Sigma^3)$, respectively, was utilized as a key property. This approach leads to asymptotic valid testing procedures in case of asymptotically negligible ($\tau_{BS} \to 0$) or crucial ($\tau_{BS} \to 1$) impact of the largest eigenvalue of Σ^2. This is easily seen from the fact that $K_f \to N(0,1)$ weakly if $\tau_{BS} \to 0$ and $K_f \to (\chi_1^2 - 1)/\sqrt{2}$ weakly if $\tau_{BS} \to 1$. Moreover, in the cases between $\tau_{BS} \to 0$ and $\tau_{BS} \to 1$ it provides a quite good approximation of the unknown distribution of W_n.

In an extensive simulation study, presented in the paper and the supplementary material, the behavior of the suggested approximation technique was evaluated for varying combinations of n and d as well as different covariance structures (covering the different limit distributions). It was observed that the procedure is quite accurate, even for a small size of the dimension d and the sample size n.

Finally, we like to note that the asymptotic results derived in the current paper are robust to the imposed assumption on multivariate normality. In particular, based on additional results from Ellenberger (2011) and Pauly (2009), it can be shown that similar asymptotics can be derived in the context of the more general model considered by Bai and Saranadasa (1996). However, the corresponding proofs involve lengthy derivations and need further technical results which are of its own interest and will therefore be deferred to a subsequent paper. Moreover, a corresponding two sample case where also a potential interaction between the gender and the cause of the time curve is of potential interest will be left to further investigations. In this case also permutation techniques similar to those developed by Pesarin and Salmaso (2010, 2012) or Pauly, Brunner and Konietschke (2015) will be left to future research.

Acknowledgement

8 Appendix

For the ease of convenience we partially suppress as above the dependencies on d or on n.

Proof of Theorem 2.1. We rewrite $\widetilde{W}_n$ as

$$\widetilde{W}_n \;=\; \frac{Q_n - \operatorname{tr}(\boldsymbol{\Sigma})}{\sqrt{2\operatorname{tr}(\boldsymbol{\Sigma}^2)}} \;=\; \frac{Q_n - \sum_{s=1}^{d}\lambda_s}{\sqrt{2\sum_{s=1}^{d}\lambda_s^2}} \;=\; \sum_{s=1}^{d} a_s \frac{C_s^2 - 1}{\sqrt{2}}\,,$$

with the array $a_s = \lambda_s / \sqrt{\sum_{r=1}^{d}\lambda_r^2}$ and row-wise independent and identically distributed random variables $(C_s^2 - 1)/\sqrt{2}$ with mean 0 and variance 1. Since the coefficients a_s fulfill the conditions

$$\max_{1\leq s\leq d} |a_s| \longrightarrow 0 \quad \text{and} \quad \sum_{s=1}^{d} a_s^2 = 1,$$

the result (a) follows from Theorem 1 in Hájek, Šidák, and Sen (1999, p. 184), see also Theorem 1(a) in Zhang (2005).

For case (b) recall that the eigenvalues are assumed to be ordered, *i.e.* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. Thus, the sequence a_s is ordered as well. We now write

$$\widetilde{W}_n \;=\; a_1 \frac{C_s^2 - 1}{\sqrt{2}} + \sum_{s=2}^{d} a_s \frac{C_s^2 - 1}{\sqrt{2}} \;=\; a_1 \frac{C_s^2 - 1}{\sqrt{2}} + R.$$

Here the first part is asymptotically standardized $(\chi^2 - 1)/\sqrt{2}$-distributed by Slutsky's Lemma. Moreover, the second part is asymptotically negligible since $E(R) = 0$ and

$$Var(R) \;=\; \sum_{s=2}^{d} a_s^2 = 1 - a_1^2 \to 0.$$

Hence (b) follows from another application of Slutsky's Lemma.

Finally, the prove of (c) follows in a similar way as in the proof of Theorem 3.1 in Jentsch and Pauly (2012) or Pauly (2009) by applying a truncation argument. First note that for each fixed $r \leq d$ we have convergence in distribution

$$\widetilde{W}_n^{(r)} = \sum_{s=1}^{r} a_s \frac{(C_s^2 - 1)}{\sqrt{2}} \;\xrightarrow{d}\; \sum_{s=1}^{r} b_s \frac{(C_s^2 - 1)}{\sqrt{2}} \;=\; V_1^{(r)}$$

as $n, d \to \infty$. Here the sequence of limit variables $(V_1^{(r)})_r$ is a Cauchy sequence in L_2 since

$$Var\left(\sum_{s=r_1}^{r_2} b_s \frac{(C_s^2 - 1)}{\sqrt{2}} \right) \;=\; \sum_{s=r_1}^{r_2} b_s^2 \;\leq\; \epsilon$$

19

for all $\epsilon > 0$ and sufficiently large $r_1 \leq r_2$. This implies that $V_1^{(r)}$ converges in distribution to $V_1^{(\infty)}$ as $r \to \infty$. Moreover, since

$$Var(\widetilde{W}_n - \widetilde{W}_n^{(r)}) = \sum_{s=r+1}^{\infty} a_s^2 = 1 - \sum_{s=1}^{r} a_s^2 \longrightarrow 1 - \sum_{s=1}^{r} b_s^2$$

as $n, d \to \infty$ it follows that $\lim_{r \to \infty} \limsup_{n,d \to \infty} Var(\widetilde{W}_n - \widetilde{W}_n^{(r)}) = 0$. The result for $\widetilde{W}_n$ follows then from Theorem 4.2 in Billingsley (1968). $\qquad\square$

Proof of Theorem 2.2. Consider the statistic W_n in (2.11). We have by (2.10) that

$$
\begin{aligned}
W_n &= \frac{Q_n - B_0}{\sqrt{2B_2}} = \left(\frac{Q_n - \mathrm{tr}(\boldsymbol{\Sigma})}{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}} - \frac{(B_0 - \mathrm{tr}(\boldsymbol{\Sigma}))}{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}} \right) \frac{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}}{\sqrt{2B_2}} \\
&= \left(\frac{Q_n - \mathrm{tr}(\boldsymbol{\Sigma})}{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}} - \frac{(B_0 - \mathrm{tr}(\boldsymbol{\Sigma}))}{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}} \right) (1 + o_p(1)) \\
&= \left(\frac{Q_n - \mathrm{tr}(\boldsymbol{\Sigma})}{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}} + o_p(1) \right) (1 + o_p(1)) \\
&= \frac{Q_n - \mathrm{tr}(\boldsymbol{\Sigma})}{\sqrt{2\,\mathrm{tr}(\boldsymbol{\Sigma}^2)}} + o_p(1) = \widetilde{W}_n + o_p(1).
\end{aligned}
$$

Here $o_p(1)$ denotes expressions that are asymptotically negligible in probability. The result now follows from Theorem 2.1 and Slutsky's Lemma. $\qquad\square$

LEMMA 8.1 *We have the following relationship between the different conditions on the covariance matrices*

$$0 \leq \tau_B \leq \tau_P \leq \tau_{CQ} \leq \tau_{BS}^2 \leq \tau_{BS} \leq 1, \tag{8.17}$$

and as $d \to \infty$

$$\tau_P \longrightarrow 0 \Leftrightarrow \tau_{CQ} \longrightarrow 0 \Leftrightarrow \tau_{BS} \longrightarrow 0, \tag{8.18}$$

$$\tau_P \longrightarrow 1 \Leftrightarrow \tau_{CQ} \longrightarrow 1 \Leftrightarrow \tau_{BS} \longrightarrow 1. \tag{8.19}$$

Proof. From Cauchy-Schwarz inequality it follows that $\left(\sum_{s=1}^d \lambda_s^{(a+b)} \right)^2 \leq \sum_{s=1}^d \lambda_s^{2a} \cdot \sum_{s=1}^d \lambda_s^{2b}$ for all $a, b > 0$. Hence the inequality (8.17) follows from

$$
\begin{aligned}
0 \quad \leq \quad \tau_{\mathrm{B}} &= \frac{\mathrm{tr}(\boldsymbol{\Sigma}^2)}{\mathrm{tr}^2(\boldsymbol{\Sigma})} \cdot \frac{\mathrm{tr}^2(\boldsymbol{\Sigma}^3)}{\mathrm{tr}^2(\boldsymbol{\Sigma}^3)} \\
&\leq \quad \frac{\mathrm{tr}(\boldsymbol{\Sigma}^2) \cdot \mathrm{tr}^2(\boldsymbol{\Sigma}^3)}{\mathrm{tr}^4(\boldsymbol{\Sigma}^2)} = \tau_{\mathrm{P}} \\
&\leq \quad \frac{\mathrm{tr}(\boldsymbol{\Sigma}^2) \cdot \mathrm{tr}(\boldsymbol{\Sigma}^4)}{\mathrm{tr}^3(\boldsymbol{\Sigma}^2)} = \tau_{\mathrm{CQ}} = \frac{\sum_s \lambda_s^4}{\mathrm{tr}^2(\boldsymbol{\Sigma}^2)} \\
&\leq \quad \frac{\max_i^2(\lambda_i) \sum_s \lambda_s^2}{\mathrm{tr}^2(\boldsymbol{\Sigma}^2)} = \tau_{\mathrm{BS}}^2 \leq \tau_{\mathrm{BS}} \leq 1.
\end{aligned}
$$

Moreover, the equivalences in (8.18) and (8.19) for $d \to \infty$ are derived by sandwiching

$$
\tau_{\mathrm{BS}}^6 \quad = \quad \frac{(\max_i \lambda_i)^6}{\mathrm{tr}^3(\boldsymbol{\Sigma}^2)} = \frac{(\max_i \lambda_i^3)^2}{\mathrm{tr}^3(\boldsymbol{\Sigma}^2)} \leq \frac{(\sum_s \lambda_s^3)^2}{\mathrm{tr}^3(\boldsymbol{\Sigma}^2)} = \tau_{\mathrm{P}}. \qquad \square
$$

In order to prove Theorem 3.1 we first have to summarize some properties of the estimator B_3 given in (3.15) which is defined as $B_3 = \binom{n}{3}^{-1} \sum_{k<\ell<r} A_{k\ell} A_{\ell r} A_{rk}$, where $A_{k\ell} = \boldsymbol{X}_k' \boldsymbol{T} \boldsymbol{X}_\ell$.

THEOREM 8.2 *Under $H_0 : \boldsymbol{T}\boldsymbol{\mu} = \boldsymbol{0}$ it holds that*

(a) $E_{H_0}(B_3) = \mathrm{tr}(\boldsymbol{\Sigma}^3)$,

(b) $Var_{H_0}(B_3/(\mathrm{tr}(\boldsymbol{\Sigma}^2))^{3/2}) = \mathcal{O}(n^{-1})$, *and*

(c) B_3^2/B_2^3 *is a consistent estimator of $\tau_P = 1/f_P$ uniformly in the dimension d.*

Proof. Since B_3 is a U-statistic with kernel $h_3(\boldsymbol{X}_k, \boldsymbol{X}_\ell, \boldsymbol{X}_r) = A_{k\ell} A_{\ell r} A_{rk}$, it is an unbiased estimator of $E(A_{k\ell} A_{\ell r} A_{rk}) = \mathrm{tr}(\boldsymbol{\Sigma}^3)$ for $k < \ell < r$. Moreover, Theorem 12.3 in van der Vaart (1998) implies

$$
Var(B_3) = \binom{n}{3}^{-1} \sum_{c=0}^3 \binom{3}{c} \binom{n-3}{3-c} \zeta_c = \sum_{c=1}^3 \mathcal{O}(n^{-c}) \zeta_c, \qquad (8.20)
$$

where $\zeta_0 = 0$ and ζ_c is the covariance between $h_3\left(\boldsymbol{X}_{k_1}, \boldsymbol{X}_{k_2}, \boldsymbol{X}_{k_3}\right)$ and $h_3\left(\boldsymbol{X}_{\ell_1}, \boldsymbol{X}_{\ell_2}, \boldsymbol{X}_{\ell_3}\right)$ if $c > 0$ indices coincide. Here $\mathcal{O}$ is the commonly used Landau symbol. It is shown in Lemma 1.1 in the supplementary material that $\zeta_i = \mathcal{O}(\mathrm{tr}^3(\boldsymbol{\Sigma}^2))$ holds for all $i = 1, 2, 3$. Hence the variance formula follows from the inequality $\mathrm{tr}^2(\boldsymbol{\Sigma}^3)/\mathrm{tr}^3(\boldsymbol{\Sigma}^2) \leq 1$. Then the consistency of B_3^2/B_2^3 is a direct consequence of the ratio-consistency of B_2 which is known from Ahmad, Werner, and Brunner (2008). $\qquad \square$

Proof of Theorem 3.1. Note, that condition (2.5) implies $f_P \to \infty$ by Lemma 8.1. Thus part (a) is a direct consequence of the classical Central Limit Theorem. Moreover, again by Lemma 8.1, the convergence (2.6) yields $f_P \to 1$ and part (b) follows from Lévy's continuity theorem since the characteristic function of a $\Gamma(\nu/2, 2)$-distribution is continuous in ν. Finally, the second part is a consequence of the ratio consistency of $\widehat{f}_P$ stated above in Theorem 8.2.

$\square$

Supplement

In the supplement to the current paper (cf. Pauly, Ellenberger, and Brunner, 2013) we provide additional technical results as well as supporting simulations for the investigated approximation.

References

[1] Ahmad, M. R., Werner, C., and Brunner, E. (2008). Analysis of High Dimensional Repeated Measures Designs: The One Sample Case. *Computational Statistics and Data Analysis* **53**, 416–427.

[2] Bai, Z. and Saranadasa, H. (1996), Effect of High Dimension: by an Example of a Two Sample Problem. *Statistica Sinica* **6**, 311–329.

[3] Billingsley, P. (1968). *Convergence of probability measures.* John Wiley & Sons, New York.

[4] Box, G. E. P. (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification. *Annals of Mathematical Statistics* **25**, 484–498.

[5] Brunner, E. (2009). Repeated Measures under Non-Sphericity. *Proceedings of the 6th St.Petersburg Workshop on Simulation*, 605–609.

[6] Brunner E., Bathke A.C., Placzek M. (2012). Estimation of Box's ϵ for low- and high-dimensional repeated measures designs with unequal covariance matrices. *Biometrical Journal* **54**, 301–316.

[7] Brunner, E., Becker, B.M., and Werner, C. (2009). Approximate Distributions of Quadratic Forms in High-Dimensional Repeated-Measures Designs. *Technical Report*, Dept. Medical Statistics, University of Göttingen.

[8] Chen, S. X. and Qin, Y.-L. (2010). A Two-Sample Test for High-Dimensional Data with Applications to Gene-Set Testing. *The Annals of Statistics* **38**, 808–835.

[9] Chi, Y.-Y., Gribbin, M., Lamers, Y., Gregory, J.F., and Muller, K.E. (2012). Global hypothesis testing for high-dimensional repeated measures outcomes. *Statistics in Medicine* **31**, 724–742.

[10] Davis, C.S. (2002). *Statistical Methods for the Analysis of Repeated Measurements.* Springer, New York.

[11] Ellenberger, D. (2011). Verbesserung des Chen-Qin-Test im Ein-Gruppen-Design. *Diploma thesis*, Dept. Mathematical Stochastics, University of Göttingen.

[12] Geisser, S. and Greenhouse, S. W. (1958). An Extension of Box's Result on the Use of the F Distribution in Multivariate Analysis. *Annals of Mathematical Statistics* **29**, 885–891.

[13] Greenhouse, S. W. and Geisser, S. (1959). On Methods in the Analysis of Profile Data. *Psychometrika* **24**, 95–112.

[14] Hájek, J., Šidak, Z., and Sen, P.K. (1999). *Theory of Rank Tests.* Academic Press, San Diego.

[15] Huynh, H. (1978). Some approximate tests for repeated measures designs. *Psychometrika* **43**, 161–175.

[16] Huynh, H., and Feldt, L. S. (1976). Estimation of the Box Correction for Degrees of Freedom From Sample Data in Randomized Block and Split-Plot Designs. *Journal of Educational Statistics* **1**, 69–82.

[17] Imhof, J.P. (1961). Computing distribution of quadratic forms in normal variables, *Biometrika* **48**, 419–426.

[18] Jentsch, C., and Pauly, M. (2012). A note on using periodogram-based distances for comparing spectral densities. *Statistics and Probability Letters* **82**, 158–164.

[19] Jordan, W., Tumani, H., Cohrs, S., Eggert, S., Rodenbeck, A., Brunner, E., Rther, E., and Hajak, G. (2004). Prostaglandin D Synthase (β-trace) in Healthy Human Sleep. *Sleep* **27**, 867–874.

[20] Katayama, S., Kano, Y., and Srivastava, M.S. (2013). Asymptotic distributions of some test criteria for the mean vector with fewer observations than the dimension. *Journal of Multivariate Analysis* **116**, 410–421.

[21] Keselman, H. J., Algina, J., and Kowalchuk, R.K. (2001). The analysis of repeated measures designs. *British Journal of Mathematical and Statistical Psychology* **54**, 1–20.

[22] Lecoutre, B. (1991). A Correction for the $\widetilde{\epsilon}$ Approximative Test in Repeated Measures Designs With Two or More Independent Groups. *Journal of Educational Statistics* **16**, 371–372.

[23] Mathai, A.M. and Provost, S.B. (1992). *Quadratic Forms in Random Variables*. Marcel Dekker Inc., New York.

[24] Park, J., and Ayyala, D.N. (2013). A test for the mean vector in large dimension and small samples. *Journal of Statistical Planning and Inference* **143**, 929–943.

[25] Pauly, M. (2009). Eine Analyse bedinger Tests mit bedingten Zentralen Grenzwertsätzen für Resampling-Statistiken. *Phd thesis*, Mathematical Institute, University of Düesseldorf.

[26] Pauly, M., Ellenberger, D., and Brunner, E. (2014). Supplement to the paper "Analysis of High-Dimensional One Group Repeated Measures Designs".

[27] Pauly, M., Brunner, E. and Konietschke, F. (2015). Asymptotic Permutation Tests in General Factorial Designs. *Journal of the Royal Statistical Society - Series B*, DOI: 10.1111/rssb.12073.

[28] Paindaveine, D. and Verdebout, T. (2013). Universal asymptotics for high-dimensional sign tests. Preprint, arXiv:1311.3118.

[29] Pearson, E.S. (1959). Note on an approximation to the distribution of non-central χ^2. *Biometrika* **46**, 364.

[30] Pesarin, F. and Salmaso, L. (2010). Finite-sample consistency of combination-based permutation tests with application to repeated measures designs. *Journal of Nonparametric Statistics* **22**, 669–684.

[31] Pesarin, F. and Salmaso, L. (2012). A review and some new results on permutation testing for multivariate problems. *Statistics and Computing* **22**, 639–646.

[32] Srivastava, M.S.(2007). Multivariate Theory For Analyzing High Dimensional Data. *Journal of the Japan Statistical Society* **37**, 53–86.

[33] Srivastava, M. S. and Du, M. (2008). A test for the mean vector with fewer observations than the dimension. *Journal of Multivariate Analysis* **99**, 386–402.

[34] Srivastava, M.S. (2009). A test for the mean vector with fewer observations than the dimension under non-normality. *Journal of Multivariate Analysis* **100**, 518–532.

[35] Van der Vaart, A.W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

[36] Werner, C. (2002). Dimensionsstabile Approximation für Verteilungen von quadratischen Formen im Repeated-Measures-Design. *Diploma thesis*, Dept. Mathematical Stochastics, University of Göttingen.

[37] Werner, C. (2004). Dimensionsstabile Approximation für Verteilungen von quadratischen Formen im Repeated-Measures-Design. *Technical Report*, Dept. Medical Statistics, University of Göttingen.

[38] Yata, K. and Aoshima, M. (2012). Effective PCA for high-dimension, low-sample-size data with noise reduction via geometric representations. *Journal of Multivariate Analysis* **105**, 193–215.

[39] Zhang, J.T. (2005). Approximate and Asymptotic Distributions of Chi-Squared-Type Mixtures With Applications. *Journal of the American Statistical Association* **100**, 273–285.

[40] Zhang, J.T. and Liang, X. (2014). One-Way ANOVA for Functional Data via Globalizing the Pointwise Fâtest. *Scandinavian Journal of Statistics* **41**, 51–71.