Accurate Inference for Repeated Measures in High Dimensions

Xiaoli Kong
Loyola University Chicago, xkong1@luc.edu

Solomon W. Harrar

Follow this and additional works at: https://ecommons.luc.edu/math_facpubs

Part of the Mathematics Commons, and the Statistics and Probability Commons

Recommended Citation
Kong, Xiaoli and Harrar, Solomon W. Accurate Inference for Repeated Measures in High Dimensions. Electronic Journal of Statistics, 13, 2: 4916-4944, 2019. Retrieved from Loyola eCommons, Mathematics and Statistics: Faculty Publications and Other Works, http://dx.doi.org/10.1214/19-EJS1644

This Article is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Mathematics and Statistics: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution 4.0 International License. © Kong and Harrar, 2019.
Accurate inference for repeated measures in high dimensions*

Xiaoli Kong and Solomon W. Harrar

Department of Mathematics and Statistics, Loyola University Chicago
e-mail: xkong1@luc.edu
Department of Statistics, University of Kentucky
e-mail: solomon.harrar@uky.edu

Abstract: This paper proposes inferential methods for high-dimensional repeated measures in factorial designs. High-dimensional refers to the situation where the dimension is growing with sample size such that either one could be larger than the other. The most important contribution relates to high-accuracy of the methods in the sense that p-values, for example, are accurate up to the second-order. Second-order accuracy in sample size as well as dimension is achieved by obtaining asymptotic expansion of the distribution of the test statistics, and estimation of the parameters of the approximate distribution with second-order consistency. The methods are presented in a unified and succinct manner that it covers general factorial designs as well as any comparisons among the cell means. Expression for asymptotic powers are derived under two reasonable local alternatives. A simulation study provides evidence for a gain in accuracy and power compared to limiting distribution approximations and other competing methods for high-dimensional repeated measures analysis. The application of the methods are illustrated with a real-data from Electroencephalogram (EEG) study of alcoholic and control subjects.

MSC 2010 subject classifications: Primary 62H10; secondary 62H15.

Keywords and phrases: Asymptotic expansion, characteristic function, covariance estimation, quadratic form, multivariate test, profile analysis.

Received April 2019.

Contents

1 Introduction .................................. 4917
2 Model and hypotheses .......................... 4919
3 Higher-order asymptotic tests ..................... 4920
4 Power under local alternatives .................... 4925
5 Simulation study ................................ 4926
6 Real data analysis ................................ 4934
7 Discussion and conclusion ........................ 4937
A Proofs ........................................ 4938
Acknowledgements .................................. 4942
References ....................................... 4942

*This is an original methodology paper.
1. Introduction

Repeated measures data arise in various disciplines of the sciences, social sciences, engineering and humanities. Study designs such as time course studies, cross-over designs and split-plot designs naturally lead to repeated measures data. The distinctive feature of repeated measures data is that the observations from the same study unit (observational or experimental) are commensurate and exhibit correlations. Analysis of continuous repeated measures data to make inference on the effects of one or many between- or within-subject crossed or nested factor effects fall into three broad categories: multivariate analysis, univariate analysis and mixed model analyses. Mixed model analyses involve some assumption concerning the correlations of the repeated measures. Despite its generality in modeling the correlation and leading to exact inference, classic multivariate methods are not applicable when the number of repeated measures is larger than the error degrees of freedom.

Univariate methods on the other hand focus on adjusting the analysis of variance (ANOVA) for the within-unit correlation. It is well known that when all observations are independent, ANOVA test statistics have exact F distribution. In the presence of the within-unit correlation, the ANOVA tests are valid only if these correlations satisfy a condition known as sphericity (Bock, 1963; Huynh and Feldt, 1970). Box (1954) suggested a correction which involves adjustment of the numerator and denominator degrees of freedom of the F-distribution by a constant multiplying factor, commonly referred to as Box’s ε. Since the constant factor ε depends on the unknown within-unit covariance matrix, solution such as using a lower bound for ε (Geisser and Greenhouse, 1958) or estimates of the within-unit covariance matrix in the calculation of ε (Huynh and Feldt, 1976; Huynh, 1978; Lecoutre, 1991) have been implemented in practical applications. These solutions have been shown to work satisfactorily in terms of controlling type-I error rate when the number of repeated measures is low compared to the degrees of freedom for estimating the covariance matrix. However, the univariate approach is obviously approximate and Rencher and Christensen (2012, p. 219) argue that it has no power advantage over the exact multivariate test. They continue to say “... The only case in which we need to fall back on univariate test is when there are insufficient degrees of freedom to perform multivariate test...”, i.e., when the number of repeated measures is larger than the error degrees of freedom to estimate the within-unit covariance.

Well before most researchers embarked on the high-dimension-low-sample-size (HDLSS) problem, Collier et al. (1967), Stoloff (1970) and Maxwell and Arvey (1982) have numerically demonstrated that the univariate approaches for repeated measures tend to be very conservative. In an attempt to improve the estimation of ε in the high dimensional situation, Chi et al. (2012) used “dual” forms of the within sum-of-squares and cross-products matrices. They claim that, besides giving stable estimates of ε, the use of the “dual” version has computational advantage. The approaches of Brunner et al. (2012) and Happ et al. (2015), on the other hand, overcome the high dimensional problem by using the so-called ANOVA-type statistic (Brunner et al., 1997, 1999) and then use
F-approximation to their null distribution by matching first two moments of the numerator and denominator quadratic forms with that of a scaled-gamma distribution, an approach shown to be successful in a related problem for independent observations by Brunner et al. (1997). Again, these approaches although were shown to be numerically satisfactory, they are only approximate solutions. On the other hand, by deriving asymptotic distributions of some suitable statistics in the high-dimensional asymptotic framework, Pauly et al. (2015), Takahashi and Shutoh (2016) and Harrar and Kong (2016) devised asymptotically-valid tests. Pauly et al. (2015) consider high-dimensional repeated measures analysis for one sample situation but with the possibility of several within subject factors. The two-sample situation was considered by Takahashi and Shutoh (2016) assuming equal covariance matrices for the two populations. More generally, Harrar and Kong (2016) addressed the multi-group as well as the unequal covariance cases. Other works such as Wang and Akritas (2010a,b) and Wang et al. (2010) are also high-dimensional asymptotic results applicable for repeated measures but additionally assume that the repeated measurements are inherently ordered and the dependence between the measurements decays as the separation between them increases. High-dimensional asymptotic mean vector comparisons have recently received attention in the statistics literature (see Bai and Saranadasa, 1996; Chen and Qin, 2010; Katayama et al., 2013; Cai et al., 2014, and the references there in) under assumptions different from that of Pauly et al. (2015), Takahashi and Shutoh (2016), and Harrar and Kong (2016). These recent results are based on limiting distributions (i.e. first-order asymptotics) and are not applicable in the multi-factor repeated measures settings.

More specifically, Harrar and Kong (2016) have proven asymptotic normality for their test statistics under certain assumptions on the covariances. In their simulation study, Harrar and Kong (2016) noticed that the error of approximation from these asymptotic distributions could be considerable unless both the number of repeated measurements and replication sizes are large. The present paper aims to derive second order asymptotics for the tests considered in Harrar and Kong (2016). In addition, the results in the current manuscript are more general in the sense that they are applicable in situations where there are multiple within and/or between unit crossed and/or nested factors. However, to get second-order asymptotic results that are valid for general design, we need to assume equality of covariance matrices in this paper.

This paper is organized as follows. Section 2 introduces the statistical model, hypotheses and notations used in the remainder of the paper. Test statistics for the various effects are presented in Section 3 together with asymptotic expansions for their null distributions. The asymptotic power are derived in Section 4. Numerical studies are carried out in Section 5. First, Monte Carlo simulations are used to show the gain in accuracy from the asymptotic expansions for a selection of covariance matrices and wide choices of values for the number of repeated measures and replication sizes. Data from a large Electroencephalogram (EEG) study of alcoholic and control subjects is used to illustrate the application of the results in Section 6. Also, simulation results by generating data with similar design parameters as the real data is considered later in the section. Sec-
tion 7 contains discussions and conclusions. All proofs and preliminary results are placed in the Appendix.

2. Model and hypotheses

Suppose $n_i$ independent $b$-dimensional observations; for $i = 1, \ldots, a$; are available from multivariate normal populations $N_b(\mu_i, \Sigma)$ denoted by $X_{i1}, \ldots, X_{in_i}$ and assume that the $a$ samples are mutually independent. The number of samples $a$ is assumed fixed. The aim of this paper is to derive second order asymptotic result for testing hypotheses in repeated measures analysis when both the total sample size $\sum_{i=1}^a n_i$ and the number of repeated measurements $b$ tend to infinity.

Let $X = (X'_{11}, \ldots, X'_{n1}, X'_{21}, \ldots, X'_{n2}, \ldots, X'_{a1}, \ldots, X'_{an_a})'$, where $X_{ik} = (X_{i1k}, \ldots, X_{ibk})'$. Further, let $\bar{X} = (\bar{X}'_1, \ldots, \bar{X}'_a)'$, where $\bar{X}_i = n_i^{-1} \sum_{k=1}^{n_i} X_{ik}$. The usual setting gives the interpretation that $X_{ijk}$ is the response from the $k$th subject treated with the $i$th level of factor $A$ (e.g., treatment) and the $j$th level of factor $B$ (e.g. time). In this model $X_{ijk}$ and $X_{i'j'k'}$ are assumed to be independent only if $i \neq i'$ or $k \neq k'$. Otherwise the dependence is completely unspecified.

Throughout this paper, $0$ will denote a matrix of all zeros where the dimension will be clear from the context, and $1_k$ denotes the $k$-dimensional vector consisting of all ones. The matrix $I_k$ is the identity matrix, whereas $J_k$ and $P_k$ are defined as $J_k = 1_k1_k'$ and $P_k = I_k - k^{-1}J_k$, respectively. We will use extensively the Kronecker (or direct) product $A \otimes B$ and the direct sum $A \oplus B$ of matrices $A$ and $B$. The symbol $\xrightarrow{D}$ stands as an abbreviation for “converges in distribution to” and $\xrightarrow{P}$ for “converges in probability to”.

In estimating a sequence of parameters $\theta_n = O(1)$ by a sequence of estimators $T_{n,b}$, consistency is meant in the sense of $E(T_{n,b} - \theta_n)^2 \to 0$ as $(n, b)$ go to infinity.

Note that from the distributional assumption made above $E[X_{ik}] = \mu_i = (\mu_{i1}, \ldots, \mu_{ib})'$ and $\text{Var}(X_{ik}) = \Sigma$, where $\Sigma$ is a $b \times b$ positive definite matrix. Let $\mu = (\mu_{11}, \ldots, \mu_{1b}, \ldots, \mu_{a1}, \ldots, \mu_{ab})'$ and $\Sigma = \sum_{i=1}^{a} \Sigma_i/n_i = D \otimes \Sigma$ where $D = \text{diag}(1/n_1, \ldots, 1/n_a)$. Then we have $E[X] = \mu$ and $\text{Var}(X) = \Sigma$.

The hypotheses of interest can be expressed as

$$H_0 : K\mu = 0 \quad \text{VS} \quad H_1 : K\mu \neq 0 \quad (1)$$

with $K = T_1 \otimes T_2$, where $T_1$ and $T_2$ are $a \times a$ and $b \times b$ matrices respectively. We require that the two matrices $T_1$ and $T_2$ are symmetric and there exist positive definite matrices $G_i$, such that $T_i G_i T_i' = T_i$ (i.e. $G_i = I$ if $T_i$ is idempotent).

Actually, we can apply the linear transformation $G_i^{-1/2} \otimes G_i^{-1/2}$ on the data $\bar{X}$ and use the symmetric and idempotent matrix $G_i^{-1/2} T_i G_i^{1/2}$ instead of $T_i$. With this manipulation, the new $K$ still defines the same hypotheses as in (1).

Without loss of generality, we can assume that $T_1$, $T_2$ and, therefore, $K$ are symmetric and idempotent matrices. For such $T_i$, $K$ is positive semidefinite.
matrix. In the following Theorem we establish an equivalent quadratic form expressions for the hypotheses (1).

**Theorem 2.1.** The null hypotheses (1) are equivalent to

\[ H'_0 : \mu' K \mu = 0 \text{ vs. } H'_1 : \mu' K \mu > 0. \]

The above setup may give the impression that the paper is dealing with one between-subject and one within-subject factor with levels \( a \) and \( b \), respectively. However, the indices \( i = 1, \ldots, a \) and \( j = 1, \ldots, b \) are to be viewed as lexicographic order of the between-subject factor level combinations and within-subject factor level combinations, respectively. Therefore, the setup covers repeated measures in factorial designs with crossed and nested factors. The factors \( T_1 \) and \( T_2 \) of matrix \( K \) can be viewed as parts of the contrast matrix concerning the between-subject factors and the within-subject factors, respectively. More specifically, suitable choices of \( T_1 \) and \( T_2 \) allow between-subject and within-subject mean comparisons. For a concrete example, consider a factorial design in which there are two between-subject crossed factors, say \( A \) and \( C \) with \( a \) and \( c \) levels, respectively, and two within-subject factors, say \( B \) and \( D \), where the levels of \( D \) are nested within that of \( B \) (see also other specific designs considered in Section 5). Suppose \( B \) has \( b \) levels and the \( j \)th level of \( B \) has \( d_j \) levels of \( D \) nested within it. The mean vector in this setup would be \( \mu = (\mu_{11}^\prime, \ldots, \mu_{1c}^\prime, \ldots, \mu_{a1}^\prime, \ldots, \mu_{ac}^\prime) \) where \( \mu_{ik} = (\mu_{1k11}^\prime, \ldots, \mu_{1k1d}^\prime, \ldots, \mu_{1khd}^\prime, \ldots, \mu_{akbd}^\prime) \). To test the interaction effect of \( A \) and \( B \), for instance, we would use \( T_1 = P_a \otimes J_c \) and \( T_2 = Q(\Sigma Q')^{-1} Q' \) where \( Q = P_b (\bigoplus_{j=1}^b d_j^{-1} 1_{d_j}) \) and \( (\Sigma Q')^{-1} \) is the generalized inverse of \( \Sigma Q' \). Further, the setup above can be reset accordingly. For example \( a \) would be replaced by \( ac \), and \( D = \text{diag} \{ n_{11}, \ldots, n_{1c}, \ldots, n_{a1}, \ldots, n_{ac} \}^{-1} \). See also Brunner et al. (2017) and Konietschke et al. (2015) for other covered designs. Another example is found in the data analysis section (Section 6).

### 3. Higher-order asymptotic tests

We have seen in Theorem 2.1 that \( K \mu = 0 \) if and only if \( \mu' K \mu = 0 \). A reasonable estimator of \( \mu' K \mu \) is given by \( H = X' K X \). The mean and variance the quadratic form \( H \) are

\[
E(H) = \text{tr}(T_i D) c_1 \mu' K \mu \quad \text{and} \quad Var(H) = 2\text{tr}(T_i D)^2 c_2 + 4\mu' K \Sigma K \mu
\]

where

\[
c_k := \text{tr}(T_2 \Sigma)^k = \text{tr}(\Sigma^{1/2} T_2 \Sigma^{1/2})^k,
\]

for \( k = 1, \ldots, 8 \). Let \( n = \Sigma_{i=1}^a n_i - a \).

In this section, we will devise a test statistic for testing (1) under the following high-dimensional asymptotic framework,
Accurate inference in high dimensions

\[ A_1 : \frac{c_8}{c_2^4} = O(b^{-3}) \text{ as } b \to \infty. \]

\[ A_2 : n \to \infty \text{ and } b \to \infty \text{ such that } \frac{b}{n} = \gamma \to \gamma_0 \in (0, \infty). \]

\[ A_3 : \frac{\mu'K\mu}{\sqrt{\text{tr}(T_1D)c_2}} = O(b^{-3/2}) \text{ as } b \to \infty. \]

It is apparent that assumption \( A_1 \) is a sparsity condition on the covariance matrix. Using the well known trace inequality (e.g. Yang et al., 2001)

\[ \text{tr}(AB)^m \leq \left\{ \text{tr}(A^m)\text{tr}(B^m) \right\}^{1/2}, \]

for any positive semidefinite matrices \( A \) and \( B \), we have

\[ \frac{c_k}{c_2^{k/2}} = O(b^{-k/2+1}) \text{ as } b \to \infty \text{ for any } 1 \leq k \leq 7. \]  

(2)

To see this, for example, when \( k = 4 \) the inequality yields

\[ c_4 = \text{tr}(T_2\Sigma)^4 = \text{tr}(T_2\Sigma^2I)^4 \leq \left\{ \text{tr}(T_2\Sigma^2)^8\text{tr}(I^8) \right\}^{1/2} \leq (c_8 \cdot b)^{1/2}. \]

Now, using condition \( A_1 \),

\[ c_4/c_2^2 \leq (c_8 \cdot b)^{1/2}/c_2^2 = O(b^{-1}). \]

Similarly, when \( k = 6 \), \( c_6 \leq (c_8 c_4)^{1/2} \). Thus,

\[ c_6/c_2^3 \leq (c_8 c_4/c_2^2)^{1/2} = (c_8/c_2^4)^{1/2}(c_4/c_2^2)^{1/2} = O(b^{-2}). \]

The conditions used in Pauly et al. (2015), Takahashi and Shutoh (2016), and Harrar and Kong (2016) are \( c_k = O(b) \) for \( k = 1, 2, 3, 4 \). These conditions are stronger compared to (2) for \( k = 1, 2, 3, 4 \). To get higher order accuracy, however, we would in addition need \( c_k = O(b) \) for \( k = 5, 6, 7, 8 \), and condition \( A_1 \) is satisfied under these assumptions.

Assumption \( A_2 \) states that the sample size and dimension grow at the same rate but otherwise either one can be larger than the other. It is weaker than the usual requirement that each of sample sizes diverge and have the same relation with \( b \). Condition \( A_3 \) is local alternative assumption which, according to Theorem 4.1, implies that the power of the test can be accurately calculated for alternatives approach to the null at a rate slower than \( b^{-3/2} \).

Assume \( \Sigma \) is known. A centered and suitably-scaled version of \( H = \bar{X}'K\bar{X} \) given by:

\[ T = \frac{H - \text{tr}(T_1D)c_1 - \mu'K\mu}{\sqrt{2\text{tr}(T_1D)^2c_2}} \]

yields a reasonable test statistic for testing \( H_0. \) In order to generalize the test statistic \( T \) for the unknown covariance case, we need to estimate \( c_1 \) and \( c_2 \) to the appropriate order. The estimators \( \hat{c}_1 \) and \( \hat{c}_2 \) are defined by

\[ \hat{c}_1 = \text{tr}(T_2S) \quad \text{and} \quad \hat{c}_2 = \frac{n^2}{(n-1)(n+2)} \left\{ \text{tr}(T_2S)^2 - \frac{1}{n}\left\{ \text{tr}(T_2S) \right\}^2 \right\}, \]  

(3)
where
\[ S = \frac{1}{n} \sum_{k=1}^{n} (n_i - 1) S_i \quad \text{and} \quad S_i = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_i)(X_{ik} - \bar{X}_i)', \]
have desirable asymptotic properties given in Theorem 3.1. Here, it should be noted that \( \hat{c}_1 \) and \( \hat{c}_2 \) are unbiased estimators of \( c_1 \) and \( c_2 \), respectively (Srivastava, 2005; Harrar and Kong, 2016).

**Theorem 3.1.** Under the high-dimensional asymptotic frameworks \( A_1 \) and \( A_2 \), the estimators \( \hat{c}_1 \) and \( \hat{c}_2 \) have the following asymptotic properties:

(i) Asymptotic equivalence: \( (\hat{c}_1 - c_1) / \sqrt{c_2} = O_p(b^{-1/2}) \) and \( (\hat{c}_2 - c_2) / c_2 = O_p(b^{-1}) \).

(ii) Ratio consistency: \( \hat{c}_2 / c_2 \xrightarrow{p} 1 \).

Let \( \delta_k = tr(T_1D)^k / [tr(T_1D)]^k \) for \( k = 1, 2, 3, 4 \) (\( \delta_1 = 1 \)). Since \( T_1 \) is a symmetric and idempotent matrix, one can see that \( 0 < \delta_k \leq 1 \) and, hence, \( \delta_k = O(1) \) as \( n \to \infty \). Next we study the asymptotic sampling distribution of the test-statistic,

\[ \widehat{T} = \frac{\{\text{tr}(T_1D)\}^{-1} H - \hat{c}_1}{\sqrt{2\delta_2c_2}} \]

which is obtained from \( T \) by replacing \( c_1 \) and \( c_2 \) by their empirical counterparts under \( H_0 \). It is shown in the Appendix that \( \widehat{T} \) can be expanded as

\[ \widehat{T} = T - \frac{V}{\sqrt{b}} \frac{1}{b} TW + O_p(b^{-3/2}), \]

where
\[ V = \frac{\sqrt{b}(\hat{c}_1 - c_1)}{\sqrt{2\delta_2c_2}} \quad \text{and} \quad W = \frac{b(\hat{c}_2 - c_2)}{c_2}, \]

are \( O_p(1) \) by Theorem 3.1. The characteristic function of \( \widehat{T} \) can be expanded as given in the following Theorem.

**Theorem 3.2.** If the null hypothesis \( H_0 \) holds, then under the high-dimensional asymptotic frameworks \( A_1 \) and \( A_2 \), the characteristic function of \( \widehat{T} \) can be expanded as

\[ \phi_{\widehat{T}}(t) = e^{\frac{1}{2} t^2 \gamma} \left\{ 1 + \frac{1}{\sqrt{b}} t^3 \eta_3 + \frac{1}{b} \left( t^2 \frac{\gamma}{2\delta_2} + t^4 \eta_4 + \frac{1}{b} \frac{\gamma^2}{2\delta_2} \right) + O(b^{-3/2}) \right\}, \]

where \( \eta_3 = \frac{4b^{1/2} \delta_3 c_3}{3(2\delta_2c_2)^{3/2}} \) and \( \eta_4 = \frac{2b\delta_4 c_4}{(2\delta_2c_2)^2} \).

Note that, by Assumption \( A_1 \), \( \eta_3 \) and \( \eta_4 \) are \( O(1) \). Inverting the characteristic function term by term, we get an asymptotic expansion for the distribution function of \( \widehat{T} \) as follows.
Theorem 3.3. If the null hypothesis $H_0$ holds, then under the high-dimensional asymptotic frameworks $A_1$ and $A_2$, the distribution function of $T$ can be expanded as

$$ F_T(x) = G_T(x) + O(b^{-3/2}), $$

uniformly in $x$ where

$$ G_T(x) = \Phi(x) - \frac{1}{\sqrt{b}} \eta_3 \Phi^{(3)}(x) + \frac{1}{b} \left\{ \frac{\gamma}{2 \delta_2} \Phi^{(2)}(x) + \eta_4 \Phi^{(4)}(x) + \frac{\eta_2^2}{2} \Phi^{(6)}(x) \right\} $$

and $\Phi^{(j)}(x)$ is the $j$th derivative of the standard normal cumulative distribution function $\Phi(x)$.

More specifically, Theorem 3.3 states that $\sup_{z \in \mathbb{R}} |F_T(x) - G_T(x)| = O(b^{-3/2})$. The function $G_T(x)$ can alternatively be expressed as

$$ G_T(x) = \Phi(x) - \phi(x) \left[ \frac{1}{\sqrt{b}} \eta_3 h_2(x) + \frac{1}{b} \left\{ \frac{\gamma}{2 \delta_2} h_1(x) + \eta_4 h_3(x) + \frac{\eta_2^2}{2} h_5(x) \right\} \right] $$

where $\phi(x)$ is the standard normal density functions and $h_i(x)$ is the $i$th Hermite polynomial. The first five Hermite polynomials are:

$$ h_1(x) = x, \quad h_2(x) = x^2 - 1, \quad h_3(x) = x^3 - 3x, $$
$$ h_4(x) = x^4 - 6x^2 + 3 \quad \text{and} \quad h_5(x) = x^5 - 10x^3 + 15x. $$

It should be emphasized that when the terms of orders $b^{-1/2}$ and $b^{-1}$ are ignored, assumptions $A_1$ and $A_2$ can be relaxed as: (i) the assumption of proportional divergence of $n$ and $b$ in $A_2$ is not needed (Harrar and Kong, 2016) and (ii) the sparsity condition on the covariance matrix (assumption $A_1$) is needed only for $c_4/c_2^2 = o(1)$, in which case, the assumption reduces to that of Chen and Qin (2010).

Let $u(z)$ be defined by $P(\hat{T} \leq u(z)) = P(Z \leq z)$ where $Z$ is a standard normal random variable. In what follows, asymptotic expansion of $u(z)$ in terms of $z$ known as Cornish-Fisher expansion (Hill and Davis, 1968) is given in Corollary 3.4.

Corollary 3.4. If the null hypothesis $H_0$ holds, then under the high-dimensional asymptotic frameworks $A_1$ and $A_2$, $u(z) = u_A(z) + O(b^{-3/2})$ where

$$ u_A(z) = z + \frac{1}{\sqrt{b}} \eta_3 h_2(z) + \frac{1}{b} \left\{ \frac{\gamma}{2 \delta_2} h_1(z) + \eta_4 h_3(z) + \frac{\eta_2^2}{2} h_5(z) - z \eta_4^2 h_2(z) \left( \frac{1}{2} h_2(z) - 2 \right) \right\}. $$

The expansions $G_T(x)$ and $u_A(z)$ are approximations for the CDF and quantile, respectively, of $\hat{T}$ under the null hypothesis. In these approximations, $\eta_3$ and $\eta_4$ depend on $c_2$, $c_3$, and $c_4$ which are unknown quantities. Therefore, for practical applications, we need an estimated version of the expansions which is uniformly correct up to $O_p(b^{-3/2})$ in the sense that $\sup_{z \in \mathbb{R}} |F_T(x) - \hat{F}_T(x)| =$
and $\hat{c}_o$ and $\hat{c}_r$ be as defined in (3) and define $\hat{c}_3$ and $\hat{c}_4$ as

$$\hat{c}_3 = \frac{n^4}{m_1} \left[ \text{tr}(T_2S)^3 - \frac{3}{n} \text{tr}(T_2S)^2 \text{tr}(T_2S) + \frac{2}{n^2} \{\text{tr}(T_2S)^4\} \right]$$

and

$$\hat{c}_4 = \frac{n^5(n^2 + n + 2)}{m_2} \left[ \text{tr}(T_2S)^4 - \frac{4}{n} \text{tr}(T_2S)^3 \text{tr}(T_2S) \right.$$

$$- \frac{2n^2 + 3n - 6}{n(n^2 + n + 2)} \{\text{tr}(T_2S)^2\}^2 + \frac{2(5n + 6)}{n(n^2 + n + 2)} \text{tr}(T_2S)^2 \{\text{tr}(T_2S)^2\}^2$$

$$- \frac{5n + 6}{n^2(n^2 + n + 2)} \{\text{tr}(T_2S)\}^4 \right],$$

where $m_1 = (n-2)(n-1)(n+2)(n+4)$ and $m_2 = m_1(n+1)(n-3)(n+6)$. These estimators are unbiased and enjoy some higher order asymptotic properties that makes them suitable for use in asymptotic expansions.

**Theorem 3.5.** Under the high-dimensional asymptotic frameworks $A_1$ and $A_2$, the estimators $\hat{c}_2, \hat{c}_3$ and $\hat{c}_4$ have the following properties:

(i) **Unbiasedness:** $E[\hat{c}_k] = c_k$ for $k = 2, 3, 4$.

(ii) **Asymptotic equivalence:**

$$\frac{b^{1/2} \hat{c}_3}{c_2^{3/2}} = \frac{b^{1/2}c_3}{c_2^{3/2}} + O_p(b^{-1})$$

and

$$\frac{b^{1/2} \hat{c}_4}{c_2^{3/2}} = \frac{b^{1/2}c_4}{c_2^{3/2}} + O_p(b^{-1}).$$

Now, by using Theorems 3.1 and 3.5, we know that

$$\hat{\eta}_3 := \frac{4b^{1/2} \delta \hat{c}_3}{3(2\delta c_2)^{3/2}} = \eta_3 + O_p(b^{-1}),$$

$$\hat{\eta}_3^2 = \eta_3^2 + O_p(b^{-1})$$

and

$$\hat{\eta}_4 := \frac{2b \delta_2 \delta \hat{c}_4}{(2\delta c_2)^{3/2}} = \eta_4 + O_p(b^{-1}).$$

Therefore, we can define the estimated version $\hat{G}_\pi(x)$ of $G_\pi(x)$ of Theorem 3.3 by replacing $\hat{\eta}_3$ and $\hat{\eta}_4$ with $\hat{\eta}_3$ and $\hat{\eta}_4$, respectively.

Before we close this section we shed some light on how the asymptotics $(b, n) \to \infty$ is working. Let $R = \text{rank}(T_2)$ be the number of nonzero eigenvalues of $T_2 \Sigma$ and denote the eigenvalues by $\lambda_{b,k}$ in decreasing order, i.e. $\lambda_{b,1} \geq \lambda_{b,2} \geq \cdots \geq \lambda_{b,R} > 0 = \lambda_{b,R+1} = \cdots = \lambda_{b,b}$. Furthermore, assume there are $r$ dominant eigenvalues. That is, there exists $0 < M \leq 1$ such that

$$\liminf_{b \to \infty} \frac{\lambda_{b,r}}{\lambda_{b,1}} = M$$

and

$$\limsup_{b \to \infty} \frac{\lambda_{b,r+1}}{\lambda_{b,1}} = 0.$$

If the rank of $T_2$ is finite, then $r$ is bounded and, consequently, condition $A_1$ is violated. Thus, for condition $A_1$ to be satisfied, the rank of $T_2$ needs to grow with $b$. Especially, when all eigenvalues are bounded, $A_1$ holds only if $r/b \to 0$. If $r$ grows with $b$ at a slower rate, a limiting normal distribution still holds for $T$.
but not the second order accuracy results. This case is interesting and we plan to investigate asymptotic expansion under this framework in future researches.

Next we provide an approximate solution in the situation where \( r \) is bounded. In this case, \( T \) has the same distribution as a linear combination of centered and independent Chi-square random variables with one degrees of freedom (see also Pauly et al., 2015). More precisely,

\[
T = \frac{1}{\sqrt{c_2}} \sum_{k=1}^{r} \frac{\lambda_{b,k} (Y_k - 1)}{\sqrt{c_2}},
\]

where \( Y_k \) are iid \( \chi^2_1 \) random variables. When \( n, p \to \infty \), \( \hat{T} = T + o_p(1) \), if the limit of \( \lambda_{b,k}/\sqrt{c_2} \) exists, for all \( k = 1, \ldots, r \).

The exact distribution cannot be used directly as the eigenvalues are unknown and their consistent estimators are difficult to come by. We may use the following approximations

\[
H \approx \frac{\text{tr}(T_1D)\delta_2 c_2}{c_1} \lambda_{c_1/(c_2 b_2)}^2 \quad \text{and} \quad \hat{c}_1 \approx n^{-1} c_2 \lambda_{nc_1/(c_2)}^2.
\]

These approximations are obtained by matching the first two moments with that of a scaled Chi-Square distribution under \( \mathcal{H}_0 \). Further, it is known that \( H \) is independent of \( \hat{c}_1 \). Thus, a reasonable test statistic is,

\[
\hat{T} = \frac{H}{\text{tr}(T_1D)}\hat{c}_1.
\]

Under \( \mathcal{H}_0 \), the distribution of \( \hat{T} \) can be approximated by \( \mathcal{F}_{c_1/(c_2 b_2), nc_1/(c_2)} \) where \( \mathcal{F}_{\nu_1, \nu_2} \) is the \( F \) distribution with degrees of freedom \( \nu_1 \) and \( \nu_2 \). When the rank of \( T_2 \) is 1, it turns out that \( T_2 \Sigma \) has only one nonzero eigenvalue \( c_1 \), and \( c_2 = c_1^2 \). Then \( \hat{T} \) has exact \( \mathcal{F}_{1/(b_2 n)} \) distribution. For example, for testing the main effect of \( A \) or \( C \) or their interaction in the four-factor design mentioned in the last paragraph of Section 2, we shall use \( Q = J_6 / b(\bigoplus_{j=1}^{b} d_j^{-1} Y_j) \), in which \( \text{Rank}(T_2) = 1 \). For another example, let \( T_1 = \text{diag}(n_1, \ldots, n_a) \) \( - \) \( (n_1, \ldots, n_a)(n_1, \ldots, n_a)/(n + a) \) and \( T_2 = J_6 / b \). These are matrices of special interest in testing the equality of mean vectors given that they are parallel (see Harrar and Kong, 2016). Under this set up, \( \hat{T} \) has exact \( \mathcal{F}_{a-1, n} \) distribution. In the more general case, \( \hat{c}_1^2/\hat{c}_2 \) is a consistent estimate of \( c_1^2/c_2 \). From the proof of Theorem 3.1, we know that \( E[\hat{c}_k] = c_k, k = 1, 2 \) and \( \text{Var}(\hat{c}_1) = 2c_2/n \). Thus, \( \hat{c}_1/\sqrt{c_2} \to c_1/\sqrt{c_2} \) and \( c_2/\hat{c}_2 \to 1 \) as \( n, b \to \infty \). Therefore, by Slutsky theorem and continuous mapping theorem, \( \hat{c}_1^2/\hat{c}_2 \to c_1^2/c_2 \), as \( n, b \to \infty \).

4. Power under local alternatives

It turns out that the power functions under the local alternatives \( A_3 \) depend on the mean vectors through \( \Delta = \mu' K \mu \). Specifically, define the power function of \( \hat{T} \) by

\[
\beta(\Delta) = P(\hat{T} > u(z_0)).
\]
where $z_\alpha$ is the upper $\alpha$ (significant level) quantile of the standard normal distribution. Then, we can obtain the power functions under the local alternative $A_3$ as given in Theorem 4.1.

**Theorem 4.1.** Under the assumption $A_1, A_2$ and $A_3$, the function $\beta(\Delta)$ can be asymptotically expanded as

$$\beta(\Delta) = 1 - G_T\left( u(z_\alpha) - \frac{\{\text{tr}(T_1D)\}^{-1}\Delta}{\sqrt{2\delta_2c_2}} \right),$$

where $G_T(\cdot)$ is as defined in Theorem 3.3.

It follows from Corollary 3.4 that $u(z_\alpha) - \hat{u}_A(z_\alpha) = O_p(b^{-3/2})$. Therefore, Theorem 4.1 could still provide accurate approximation to the power function of the test that rejects $H_0$ if $\hat{T} > \hat{u}_A(z_\alpha)$. It can be easily seen that when terms of orders $b^{-1/2}$ and $b^{-1}$ are ignored from the quantiles $u_A(z_\alpha)$ and the power function $\beta(\Delta)$ reduces to

$$\beta(\Delta) = \Phi\left( \frac{\{\text{tr}(T_1D)\}^{-1}\Delta}{\sqrt{2\delta_2c_2}} - z_\alpha \right).$$

When $a = 2, T_1 = P_a$ and $T_2 = I_b$, it has the same form as Bai and Saranadasa (1996) and Chen and Qin (2010). It should be noted, however, that for this later test, Theorem 4.1 gives a second order approximation to the power function which is more accurate than Bai and Saranadasa (1996) and Chen and Qin (2010).

5. Simulation study

To exhibit the improvement resulting from the asymptotic expansion and, hence, facilitate comparison with the limiting distributions in Harrar and Kong (2016), the simulation study will mainly focus on the model where there is one between- and one within-subject factors. We generate 10,000 replications of data from $N_b(\mu, \Sigma)$. Although the assumed asymptotic frameworks stipulate $n$ to grow proportionally with $b$, in reality the actual ratio varies from application to application. To investigate the effect of various proportionality of growth, we look at values of several combinations of $a$, $b$ and $n_i$s. For practical reasons, we also consider small $b$ and large $n_1, \ldots, n_a$ (and vice-versa) combinations with balanced as well as unbalanced designs. For the number of groups (number of levels of factor A), we will consider $a = 2, 3, 4, 6$. and we set significant level $\alpha$ at 0.01 and 0.05.

Tables 2–7 present actual type I error rates (test sizes) for the covariance structures $\Sigma = \rho I_b + (1 - \rho)J_b$, $\Sigma = (\rho^{|j-j'|})$ and $\Sigma = (\sigma_{jj'})$ where $\sigma_{jj'} = 1$ and, for $j \neq j'$, $\sigma_{jj'} = \rho/(j - j')^{1/4}$. We consider a range of values for $\rho$. For the first covariance structure, the assumptions in $A_1$ are satisfied uniformly in $\rho$ because $c_i = (1 - \rho)^i(b - 1)$. However, for the other two covariance structures these assumptions do not hold except for $\rho = 0$. Especially, when $\rho$ is close
to 1, the quantities $b^2c_6/c_2^3$ and $b^3c_8/c_2^4$ could diverge very fast. To illustrate this concretely, Table 1 shows the values of $b^{3/2}c_5/c_2^{3/2}$ and $b^3c_8/c_2^4$ for different values of $b$ and $\rho$.

| $\rho$ | $b = 12$ | $b = 25$ | $b = 50$ | $b = 100$ | $b = 200$ | $b = 400$ |
|--------|----------|----------|----------|----------|----------|----------|
| 0.5    | 2.98     | 3.58     | 15.85    | 46.14    | 133.50   | 392.58   |
| 0.5    | 16.69    | 99.25    | 694.56   | 4683.55  | 497817.62|
| 0.9    | 14.84    | 29.66    | 61.35    | 135.15   | 314.17   | 763.47   |
| 0.9    | 302.75   | 1349.69  | 6311.86  | 33055.11 | 190299.63| 1182933.54|

It should be noted that the covariance matrix structure $\Sigma = \rho I_b + (1 - \rho)J_b$ will be positive definite if and only if $-1/(b - 1) < \rho < 1$. The empty cells in Tables 2–7 correspond to the cases where $b$ and $\rho$ combinations do not yield positive definitive covariance matrices. In all the three tables we consider the contrast matrices $T_1 = P_a$ or $J_a/a$ and $T_2 = P_b$. Another matrix of particular interest in repeated measures analysis is $T_2 = J_b/b$ which is useful to test the group effect. However, the distribution of $\hat{T}$ in this case does not depend on $b$. Hence, we do not carry out simulation for this contrast matrix.

First and foremost, comparing Table 2, 4 and 6 (results for $\alpha = 0.05$) with the results in Harrar and Kong (2016), one can clearly see a marked gain in accuracy resulting from the inclusion of higher-order terms in the asymptotic expansion. We can see in Table 2 that for both tests (i.e. $T_1 = P_a$ and $J_a/a$), a large number of the achieved error rates are within a tenth of the actual values. This phenomena happens more for weaker correlations than for stronger ones. Further, it is clear from the tables that the performance of the tests in controlling type I error rates is excellent when either the sample sizes or the dimension is large. This might be due to the homoscedastic nature. It seems also the case that when $n_i$’s are small, the tests control type I error rate better as $a$ gets larger. For example, looking at the rows for $a = 6$, performance appear to be satisfactory for small sample sizes but large dimension. Tables 4 and 6 seem to exhibit similar patterns and behaviors. Likewise, for $\alpha = 0.01$, the asymptotic expansion provides a gain in accuracy in controlling type-I error rates (see Tables 3, 5 and 7).

To investigate performance in terms of power, we compare the power of the method by Chi et al. (2012) with the methods proposed in this paper taking $T_2 = P_b$, and setting $T_1$ to either $P_a$ or $J_a/a$. The power comparison is restricted to Chi et al. (2012) because, to the best of our knowledge, this is the only work for high-dimensional repeated measures data that covers the general factorial design. Their approach uses a property of the Wishart random matrix to avoid the singularity problem that arises due to high dimensionality. This manipulation could, however, potentially compromise the power properties of the test. That is one objective of the power comparison in addition to evaluating the power performance of the proposed method. To keep the comparison
Table 2
Achieved Type I error rates (×100%) for the testing procedures when $T_2 = P_0$ and sampling from $N_0(\mu, \Sigma)$ where $\Sigma = (1 - \rho)I + \rho J$. The nominal size is $\alpha = 0.05$.

| $a$ | $b$ | $(n_1, \ldots, n_9)$ | $T_1 = P_0$ | $T_1 = J_{a/b}$ |
|-----|-----|----------------------|-------------|-----------------|
|     |     |                      | $a = 0.01$  | $a = 0$         |
| 1   | 2   | (10,10)              | 5.60        | 5.30            |
| 2   | 10  | (10,10)              | 5.70        | 6.00            |
| 2   | 15  | (10,10)              | 6.30        | 6.00            |
| 2   | 12  | (50,100)             | 5.19        | 5.10            |
| 2   | 12  | (100,100)            | 4.96        | 4.65            |
| 2   | 25  | (100,100)            | 5.12        | 4.66            |
| 2   | 50  | (50,100)             | 5.20        | 4.89            |
| 2   | 50  | (100,100)            | 4.73        | 4.57            |
| 2   | 100 | (10,13)              | 5.11        | 4.86            |
| 2   | 100 | (12,15)              | 5.10        | 4.86            |
| 2   | 100 | (25,25)              | 4.96        | 4.89            |
| 2   | 200 | (25,25)              | 4.87        | 4.82            |
| 2   | 200 | (50,50)              | 5.08        | 4.98            |
| 2   | 200 | (100,100)            | 5.20        | 4.97            |
| 2   | 200 | (50,50)              | 4.98        | 4.98            |
| 2   | 400 | (100,100)            | 4.81        | 4.95            |
| 2   | 400 | (100,200)            | 5.04        | 5.19            |
| 2   | 500 | (100,100)            | 5.04        | 5.19            |
| 2   | 400 | (100,200)            | 5.04        | 5.19            |
| 3   | 5   | (10,10,10)           | 5.90        | 5.70            |
| 3   | 10  | (10,10,10)           | 6.70        | 5.80            |
| 3   | 15  | (10,10,10)           | 5.70        | 5.60            |
| 3   | 12  | (100,100,100)        | 5.15        | 4.81            |
| 3   | 25  | (50,100,100)         | 5.15        | 4.81            |
| 3   | 25  | (100,100,100)        | 4.80        | 4.88            |
| 3   | 50  | (50,100,100)         | 5.15        | 4.81            |
| 3   | 50  | (100,100,100)        | 4.79        | 4.78            |
| 3   | 100 | (16,17,17)           | 4.95        | 4.99            |
| 3   | 100 | (16,17,33)           | 4.86        | 4.80            |
| 3   | 200 | (33,34,34)           | 5.18        | 5.09            |
| 3   | 200 | (33,34,67)           | 4.63        | 4.62            |
| 3   | 250 | (45,50,50)           | 4.62        | 4.62            |
| 3   | 250 | (50,50,100)          | 4.85        | 4.90            |
| 4   | 12  | (50,50,100,100)      | 5.13        | 4.93            |
| 4   | 12  | (100,100,100,100)    | 5.01        | 5.03            |
| 4   | 25  | (50,50,100,100)      | 4.81        | 4.82            |
| 4   | 25  | (100,100,100,100)    | 5.19        | 5.14            |
| 4   | 50  | (50,50,100,100)      | 5.36        | 5.22            |
| 4   | 50  | (100,100,100,100)    | 5.11        | 5.06            |
| 4   | 100 | (12,13,13)           | 5.02        | 5.06            |
| 4   | 100 | (12,13,25,25)        | 5.02        | 5.06            |
| 4   | 200 | (25,50,50)           | 4.96        | 4.89            |
| 4   | 200 | (50,50,50)           | 5.31        | 5.07            |
| 4   | 200 | (100,100,100,100)    | 4.97        | 4.71            |
| 6   | 100 | (8,8,8,8,8,9,9)      | 4.98        | 5.07            |
| 6   | 100 | (8,8,16,17,17)       | 5.13        | 5.07            |
| 6   | 200 | (16,17,17,17,17)     | 4.97        | 4.71            |
Table 3
Achieved Type I error rates (×100%) for the testing procedures when $T_2 = P_1$, and sampling from $N(p, \Sigma)$ where $\Sigma = (1 - \rho)I_6 + \rho J_6$. The nominal size is $\alpha = 0.01$.

| $a$ | $b$ | $(n_1, \ldots, n_6)$ | $T_1 = J_6 / \alpha$ | $T_2 = P_1$ |
|-----|-----|----------------------|----------------------|-------------|
| 1   | 1   | (10, 10)             | 2.06 1.40 1.30 1.40 1.60 1.60 | 1.50 1.50 1.40 1.40 |
| 1   | 10  | (10, 10)             | 1.78 1.78 1.95 1.78 1.60 1.60 | 1.40 1.40 1.40 1.40 |
| 2   | 15  | (10, 10)             | 1.00 1.60 1.60 1.60 1.60 1.60 | 1.00 1.00 1.00 1.00 |
| 2   | 25  | (10, 10)             | 0.95 0.95 0.95 0.95 0.95 0.95 | 0.95 0.95 0.95 0.95 |
| 3   | 30  | (10, 10)             | 1.27 1.02 1.02 1.02 1.02 1.02 | 0.84 0.84 0.84 0.84 |
| 4   | 10  | (10, 10, 10)         | 0.88 0.91 0.91 0.91 0.91 0.91 | 0.96 0.96 0.96 0.96 |
| 4   | 12  | (10, 10, 10)         | 1.17 1.02 1.02 1.02 1.02 1.02 | 0.96 0.96 0.96 0.96 |
| 4   | 15  | (10, 10, 10)         | 1.06 0.94 0.94 0.94 0.94 0.94 | 0.88 0.88 0.88 0.88 |
| 5   | 25  | (10, 10, 10)         | 1.07 1.04 1.04 1.04 1.04 1.04 | 0.98 0.98 0.98 0.98 |
| 6   | 40  | (10, 10, 10)         | 1.04 1.04 1.04 1.04 1.04 1.04 | 0.98 0.98 0.98 0.98 |

**Note:** The table continues with similar data for different values of $a$ and $b$.
Table 4
Achieved Type I error rates (×100%) for the testing procedures when $T_2 = P_j$ and sampling from $N_p(\mu, \Sigma)$ where $\Sigma = (\rho^{j-i})$. The nominal size is $\alpha = 0.05$.

| $\alpha$ | $b$ | $(n_1, \ldots, n_2)$ | $T_1 = P_j$ | $T_1 = J_{\rho}/a$ |
|-------|-----|----------------------|-------------|---------------------|
|       |     |                      | $\rho = 0.01$ | $0$ 0.2 0.5         | $\rho = 0.01$ | $0$ 0.2 0.5         |
| 2     | 5   | (10,10)              | 5.50 5.60 6.50 7.90 | 6.90 6.10 6.40 7.00 |
| 2     | 10  | (10,10)              | 5.50 5.70 6.50 4.50 | 7.20 5.90 6.60 6.80 |
| 2     | 15  | (10,10)              | 5.60 5.80 6.50 5.40 | 4.50 4.00 6.30 6.20 |
| 2     | 25  | (10,10)              | 5.70 5.90 6.50 6.50 | 6.90 5.70 6.40 6.40 |
| 2     | 50  | (50,100)             | 5.60 5.80 6.50 5.30 | 6.40 5.60 6.60 6.90 |
| 2     | 100 | (50,100)             | 5.60 5.60 6.50 6.90 | 6.90 6.20 6.50 6.80 |
| 2     | 250 | (50,100)             | 5.70 5.90 6.50 6.90 | 6.70 6.00 6.70 6.90 |
| 2     | 500 | (50,100)             | 5.70 5.90 6.50 6.90 | 6.70 6.00 6.70 6.90 |

This table shows the achieved Type I error rates for different combinations of $\alpha$, $b$, and $(n_1, \ldots, n_2)$, with the testing procedures $T_1 = P_j$ and $T_1 = J_{\rho}/a$, under the null hypothesis $H_0: \rho = 0.01$. The nominal size $\alpha = 0.05$ is used in the testing procedures when $T_2 = P_j$.
### Table 5
Achieved Type I error rates (×100%) for the testing procedures when $T_2 = P_b$ and sampling from $N_b(\mu, \Sigma)$ where $\Sigma = (\rho^{j-j'})$. The nominal size is $\alpha = 0.01$.

| $n$ | $b$ | $T_1 = P_1$ | $T_1 = J_0/\alpha$ |
|-----|-----|-------------|-------------------|
| 2   | 5   | $T_1 = P_1$ | $T_1 = J_0/\alpha$ |
|     |     | $\rho = 0$ | $\rho = 0$ |
|     |     | 0.01 0.02 0.05 | 0.01 0.02 0.05 |
| 2   | 10  | 1.30 1.20 1.10 | 1.90 1.80 1.70 |
| 2   | 15  | 1.70 1.60 1.50 | 2.30 2.20 2.10 |
| 2   | 20  | 1.00 0.90 0.80 | 1.70 1.60 1.50 |
| 2   | 25  | 1.00 0.90 0.80 | 1.60 1.50 1.40 |
| 2   | 30  | 0.70 0.60 0.50 | 1.70 1.60 1.50 |
| 2   | 35  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 40  | 0.70 0.60 0.50 | 2.00 1.90 1.80 |
| 2   | 45  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 50  | 0.70 0.60 0.50 | 2.00 1.90 1.80 |
| 2   | 55  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 60  | 0.70 0.60 0.50 | 2.00 1.90 1.80 |
| 2   | 65  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 70  | 0.70 0.60 0.50 | 2.00 1.90 1.80 |
| 2   | 75  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 80  | 0.70 0.60 0.50 | 2.00 1.90 1.80 |
| 2   | 85  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 90  | 0.70 0.60 0.50 | 2.00 1.90 1.80 |
| 2   | 95  | 0.90 0.80 0.70 | 2.00 1.90 1.80 |
| 2   | 100 | 0.70 0.60 0.50 | 2.00 1.90 1.80 |

*Note: $\alpha = 0.01$.*
Table 6
Achieved Type I error rates (×100%) for the testing procedures when \( T_2 = P_j \) and sampling from \( N_j(\mu, \Sigma) \) where \( \Sigma = (\rho/(j + 1)^{1/2}) \). The nominal size is \( \alpha = 0.05 \).

| \( n \) | \( b \) | \( (n_1, \ldots, n_b) \) | \( T_1 = P_j \) | \( T_1 = J_{n-b} \) |
|---|---|---|---|---|
| 2 | 5 | (10,10) | 0.01 | 0.01 |
| 2 | 10 | (10,10) | 0.006 | 0.006 |
| 2 | 15 | (10,10) | 0.006 | 0.006 |
| 2 | 25 | (10,10) | 0.015 | 0.015 |
| 2 | 50 | (10,10) | 0.025 | 0.025 |
| 2 | 100 | (10,10) | 0.045 | 0.045 |
| 2 | 200 | (10,10) | 0.075 | 0.075 |
| 2 | 400 | (10,10) | 0.115 | 0.115 |
| 3 | 5 | (10,10,10) | 0.007 | 0.007 |
| 3 | 10 | (10,10,10) | 0.007 | 0.007 |
| 3 | 15 | (10,10,10) | 0.007 | 0.007 |
| 3 | 25 | (10,10,10) | 0.012 | 0.012 |
| 3 | 50 | (10,10,10) | 0.019 | 0.019 |
| 3 | 100 | (10,10,10) | 0.031 | 0.031 |
| 3 | 200 | (10,10,10) | 0.051 | 0.051 |
| 3 | 400 | (10,10,10) | 0.072 | 0.072 |
| 4 | 5 | (10,10,10,10) | 0.009 | 0.009 |
| 4 | 10 | (10,10,10,10) | 0.009 | 0.009 |
| 4 | 15 | (10,10,10,10) | 0.009 | 0.009 |
| 4 | 25 | (10,10,10,10) | 0.014 | 0.014 |
| 4 | 50 | (10,10,10,10) | 0.021 | 0.021 |
| 4 | 100 | (10,10,10,10) | 0.033 | 0.033 |
| 4 | 200 | (10,10,10,10) | 0.054 | 0.054 |
| 4 | 400 | (10,10,10,10) | 0.075 | 0.075 |
| 5 | 5 | (10,10,10,10,10) | 0.011 | 0.011 |
| 5 | 10 | (10,10,10,10,10) | 0.011 | 0.011 |
| 5 | 15 | (10,10,10,10,10) | 0.011 | 0.011 |
| 5 | 25 | (10,10,10,10,10) | 0.016 | 0.016 |
| 5 | 50 | (10,10,10,10,10) | 0.025 | 0.025 |
| 5 | 100 | (10,10,10,10,10) | 0.042 | 0.042 |
| 5 | 200 | (10,10,10,10,10) | 0.063 | 0.063 |
| 5 | 400 | (10,10,10,10,10) | 0.085 | 0.085 |
| 6 | 5 | (10,10,10,10,10,10) | 0.013 | 0.013 |
| 6 | 10 | (10,10,10,10,10,10) | 0.013 | 0.013 |
| 6 | 15 | (10,10,10,10,10,10) | 0.013 | 0.013 |
| 6 | 25 | (10,10,10,10,10,10) | 0.017 | 0.017 |
| 6 | 50 | (10,10,10,10,10,10) | 0.028 | 0.028 |
| 6 | 100 | (10,10,10,10,10,10) | 0.048 | 0.048 |

\[ T_2 = P_j \] and \( \Sigma = (\rho/(j + 1)^{1/2}) \).
Table 7
Achieved Type I error rates (×100%) for the testing procedures when \( T_2 = P_a \) and sampling from \( N_k(\mu, \Sigma) \) where \( \Sigma = (\rho / (j - 1)^{(j/4)}) \). The nominal size is \( \alpha = 0.01 \).

| \( n \) | \( b \) | \( (n_1, \ldots, n_a) \) | \( T_1 = P_a \) | \( T_3 = J_a / \rho \) |
|---|---|---|---|---|
| 1 | 15 | (10,10) | 1.76 | 1.80 |
| 2 | 10 | (10,10) | 1.40 | 1.70 |
| 3 | 15 | (10,10) | 1.00 | 1.80 |
| 4 | 10 | (10,10) | 1.03 | 1.10 |
| 5 | 15 | (10,10) | 0.91 | 0.86 |
| 6 | 10 | (10,10) | 0.93 | 1.04 |
| 7 | 15 | (10,10) | 0.99 | 0.91 |
| 8 | 10 | (10,10) | 1.11 | 0.97 |
| 9 | 15 | (10,10) | 1.00 | 0.96 |
| 10 | 10 | (10,10) | 1.01 | 0.90 |
| 11 | 15 | (10,10) | 1.03 | 1.05 |
| 12 | 10 | (10,10) | 1.06 | 1.05 |
| 13 | 15 | (10,10) | 1.09 | 1.04 |
| 14 | 10 | (10,10) | 1.11 | 1.00 |
| 15 | 15 | (10,10) | 1.10 | 0.99 |

Accurate inference in high dimensions
manageable, we fix $a = 3$ and $\Sigma = \rho I_b + (1 - \rho)J_b$ where $\rho = 0.2$. In regards to sample sizes and dimension, we use the combinations $(b; n_1, n_2, n_3) = (10; 5, 10, 10), (10; 50, 100, 100), (100; 5, 10, 10)$ and $(100; 50, 100, 100)$. For the alternative hypotheses, when $T_1 = P_a$, we take $\mu_2 = \mu_3 = 0$ and consider two structures for $\mu_1$. The first one represents a dense alternative, namely $\mu_{1i} = (1 + \delta)$ for $i$ odd and $\mu_{1i} = (1 - \delta)$ for $i$ even, and the other one represents a sparse alternative, namely $\mu_1 = (1 + \delta, 1_{b-1}')$. In both cases $\delta$ is made to vary from 0 to 1. When $T_1 = J_a/a$, we take $\mu_1 = 1_b + \mu_2$, $\mu_2 = \mu_3$ and consider two structures of $\mu_2$ representing dense and sparse alternatives. For the first one, we take $\mu_{2i} = \delta$ for $i$ odd and $\mu_{2i} = -\delta$ for $i$ even, and for the second one we take $\mu_2 = (0_{b-1}', \delta)'$. Here also, $\delta$ varies from 0 to 1. The latter structure for both values of $T_1$ represent departures that approach to the null hypotheses at the rate $b^{1/2}$. More precisely, the scaled departure from the null $||\mu_1 - 1_b||/\text{tr}(\Sigma)^{1/2}$ are $\delta$ and $|\delta|/\sqrt{b}$, respectively. Figures 1 and 2 show power results for $T_1 = P_a$ and $T_1 = J_a/a$, respectively. For dense alternatives (left panels), our methods has a clear advantage in all cases. More pronounced dominance is observed, in particular, when $n$ is small. On the other hand, Chi et al. (2012) has an advantage for sparse alternatives (right panels).

6. Real data analysis

In this section, we analyze a publicly available data obtained from the University of California-Irvine Machine Learning Repository.\(^1\) The data arose from a

\(^1\)web address: https://archive.ics.uci.edu/ml/datasets/EEG+Database
Accurate inference in high dimensions

Fig 2. Power comparison of the proposed methods and the test by Chi et al. (2012) for $T_1 = I_a/a$ and $T_2 = P_b$. Data is generated from $N_b(\mu, \Sigma)$, where $\Sigma = 0.8I_b + 0.2I_b$. In the both panel of the plot, $\mu_1 = 1_b + \mu_2$ and $\mu_2 = \mu_3$. In the left panel $\mu_{2i} = \delta$ for $i$ odd and $\mu_{2i} = -\delta$ for $i$ even and in the right panel $\mu_2 = (0_{p-1}', \delta)'$.

large study to examine Electroencephalograph (EEG) correlates of genetic predisposition to alcoholism. Measurements from 64 electrodes placed on subject’s scalps were recorded 256 times for 1 second. The study involved two groups of subjects: alcoholic ($n_1 = 77$) and control ($n_2 = 45$). Each subject was exposed to either a single stimulus (S1) or two stimuli (S1 and S2) which were pictures of objects chosen from a picture set. The sixty-four electrodes (channels) are divided into groups based on their location on the scalp (frontal, temporal, parietal and occipital lobes). To illustrate the application of the methods concisely, we focus the analysis on data from the stimulus S1 and the seventeen frontal-lobe channels. The outcome measurements are Event-Related Potentials (ERP) indicating the level of electrical activity (in volts) in the region of the brain where each of the electrodes is placed. This repeated measures data has two within-subject factors (time and channels) and one between-subject factor (alcohol use). The within-subject factors time and channel have 256 and 17 levels, respectively.

To assess the plausibility of the assumptions in the model, we checked channel-by-channel marginal normality after testing for equality of covariance matrices in the alcoholic and control groups. To date, there does not exist a satisfactory test of multivariate normality in the high-dimensional repeated measures. Neither of the high-dimensional tests Li and Chen (2012) nor Zhang et al. (2018) indicated violation of equality of covariance for any of the channels ($p$-values $\geq 0.512$). However, the tests did not give a clean bill of normality. The proportion of fail to rejections (not adjusted for multiplicity) ranged from 24% to 41% for all but five of the frontal periphery channels. For these later
channels, the proportions (unadjusted) ranged from 57% to 82%. Considering the multiplicity of the tests, this may only represent a mild violation of normality.

The main research questions of interest are: (H01) whether the ERP profiles over time differ between channels and groups (three-way interaction: alcohol × time × channel); (H02) whether ERP profiles are similar between the channels when averaged over groups (similar time profiles for all the channels); (H03) if the time profiles of ERP are similar between the two groups averaged over channels; (H04) whether the ERP profiles are constant (flat) when averaged over channels and groups. For describing the contrast matrices, we assume the data vectors from each subject are arranged by grouping the 17 channels within each time point, i.e. the data vector from the jth subject in the ith group is \( X_{ij} = (X_{i,j1,17}, \ldots, X_{i,j,256,1}, \ldots, X_{i,j,256,17})' \). In the notations of the paper, the four hypotheses of interest, \( H_0i \) for \( i = 1, 2, 3, 4 \), can be tested by using the contrast matrices \( T_1 = P_2 \) and \( T_2 = P_{256} \otimes P_{17} \); \( T_1 = J_{2}/2 \) and \( T_2 = P_{256} \otimes J_{17}/17 \); \( T_1 = P_2 \) and \( T_2 = P_{256} \otimes J_{17}/17 \); and \( T_1 = J_{2}/2 \) and \( T_2 = P_{256} \otimes J_{17}/17 \), respectively. The results of the analysis are presented in Table 8. Overall time-profile similarity across groups (averaged over channels) cannot be rejected (p-values = 0.205). In fact, channel-by-channel similarity of time profiles of ERP across groups cannot be rejected (p-values = 0.196). However, the flatness over time is rejected overall for all channels as well as channel-by-channel.

As a way of ascertaining the reproducibility and reliability of the results in Table 8, we conducted a simulation study using parameters similar to that of the EEG data. For table 9, we generate 1000 replications of data from \( \mathcal{N}_{bd}(0, \Sigma_i) \). We look at values of \( b = 256, d = 17, a = 2 \) and \( n_1 = 77 n_2 = 45 \) and take \( \alpha = 0.05 \). Table 9 present actual type I error rates (test sizes) for the covariance structures \( \Sigma_i = \rho I_b + (1 - \rho)J_b, \) for \( \rho = 0.2 \) and random matrices \( \Sigma_i \) for \( i = 2, 3, 4 \) defined as follows. Let \( \Sigma_i = Q_i’ \Lambda_i Q_i, \) where \( \Lambda_i \) is a diagonal matrix with diagonal entries taken from \( Unif(0, 1) \) and \( Q_i \) is orthogonal matrix. Indeed, \( Q_i \) can be defined from the QR decomposition of a random matrix \( Z_i = (Z_{i, jj'}) \) where \( Z_{i, jj'} \) are iid random variables. Here, we consider three distributions for \( Z_{i, jj'} \), namely \( Z_{2, jj'} = 1_{\{j = j’\}} \) with probability 1, \( Z_{3, jj'} \sim Exp(1) \) and \( Z_{4, jj'} \sim \mathcal{N}(0, 1) \). Here, it is clear that assumption \( A_1 \) is satisfied only for \( \Sigma_1 \). It is clear from Table 9 that the achieved type I error rates are satisfactorily close to 5% regardless of the covariance matrix assumed.
Table 9
Achieved Type I error rates (×100%) for the testing procedures with parameters similar to EEG data, i.e. $a = 2, b = 256, d = 17, n_1 = 77, n_2 = 45$.

| $T_1$ | $T_2$ | $\Sigma_1$ | $\Sigma_2$ | $\Sigma_3$ | $\Sigma_4$ |
|-------|-------|-------------|-------------|-------------|-------------|
| $P_2$ | $P_{256} \otimes J_{17}$ | 5.0 | 5.9 | 4.5 | 5.8 |
| $J_{2/2}$ | $P_{256} \otimes J_{17}$ | 4.3 | 4.0 | 5.4 | 5.8 |
| $P_2$ | $J_{2/2} \otimes P_{256} \otimes J_{17}$ | 4.6 | 4.6 | 5.4 | 5.4 |
| $J_{2/2}$ | $P_{256} \otimes J_{17}/17$ | 5.6 | 4.8 | 5.2 | 4.6 |

7. Discussion and conclusion

The paper derives approximations for the null distributions and quantiles of some test statistics in repeated measures. The approximations ensure the errors to be of order $O(b^{-3/2})$ where $b$ is the dimension, i.e. the number of repeated measures. Factorial designs are treated in a unified manner where multiple between- and within-subjects factors which may be crossed or nested are allowed. General covariance structure is allowed where no pre-determined sequence is assumed among the repeated measurements. Therefore, the repeated measurements could be over time or under different treatment conditions.

The asymptotic results require some regularity condition on the covariance matrix. Such assumption appears to be inevitable as long as one prefers to consider unstructured covariance matrix. Our observation from the simulation is that this assumption does not appear to restrict the utility of the results for application in more general situations. Nevertheless, we made somewhat milder requirements compared to similar works (see, for example, Bai and Saranadasa, 1996; Takahashi and Shutoh, 2016). The paper also assumes proportional divergence of the sample size and dimension, i.e. $n/b \to \gamma_0 \in (0, \infty)$, but either one can be larger than the other. We should point out that this assumption can be relaxed to cover other cases, namely $n = O(b^\epsilon)$ for $\epsilon > 1$ or $\epsilon < 1$. However, the expanded cumulative distribution function may have terms with order different from $b^{-j/2}$ for $j = 1, 2, \ldots$ in which case the standard Cornish-Fisher formula for the quantile will not apply. Non standard expansions will need to be derived for the quantiles. Regardless, our impression from the simulation is that the effect of these terms is likely to be insignificant. This is an open problem that needs further investigation.

The development of the paper is under normality. We recommend assessing the plausibility of this assumption before applying the methods. Transformation that improve normality could also be attempted in the event non-normality is detected or suspected. In the proofs, multivariate normality of the repeated measures is mostly needed for its nice property of independence up to correlation and independence of some linear and quadratic forms. Limiting distribution results under statistical models that include independence up to correlation assumption have been derived in Bai and Saranadasa (1996) and Chen and Qin (2010) for two-sample and Katayama et al. (2013) for multiple-sample comparison of mean vectors. In the interest of space, we opted to relegate the investigation of these models for limiting distribution as well as asymptotic expansion in the repeated measures context to a follow-up manuscript.
Appendix A: Proofs

Proof of Theorem 3.1 and Theorem 3.5. We know that $nS \sim W_b(\Sigma, n)$, where $W_b(\Sigma, n)$ stands for $b$-dimensional Wishart distribution with degrees of freedom $n$ and scale matrix $\Sigma$. Denote $a_k = \text{tr}(T_2S)^k$ for $k = 1, \ldots, 8$ and define

$$b_{T_2S} = (c_1^k, c_1^{k-2}c_2, \ldots, c_k)'$$

and

$$b_{T_2\Sigma} = (a_1^k, a_1^{k-2}a_2, \ldots, a_k)'$$

to be the vector of traces of the $k$th order moments. That means, for each partition of $k$, for example $k = \nu_1 + \cdots + \nu_q$, where $\nu_1 \leq \cdots \leq \nu_q$ and $q \leq k$, we include $c_{\nu_1} \cdots c_{\nu_q}$ and $a_{\nu_1} \cdots a_{\nu_q}$ to the vectors $b_{T_2S}$ and $b_{T_2\Sigma}$, respectively, at same position. It is known (Fujikoshi, 1973) that $E[\Sigma]^{(6)}$ stands for the matrices $F_k$ for $k \leq 6$. Using these result, it can be shown that

$$E[\hat{c}_k] = c_k, \text{ for } k = 1, 2, 3, 4.$$

Further, under the high-dimensional asymptotic frameworks $A_1$ and $A_2$ and after lengthy algebraic calculation, it can be seen that

$$\text{Var}\left(\frac{\hat{c}_1}{\sqrt{c_2}}\right) = \frac{2}{n} = O(n^{-1})$$

$$\text{Var}\left(\frac{\hat{c}_2}{c_2}\right) = \frac{4}{n(n-1)(n+2)c_2^2}\left[nc_2^2 + (2n^2 + 3n - 6)c_4\right] = O(n^{-2}),$$

$$\text{Var}\left(\frac{\sqrt{\hat{c}_3}}{c_2}\right) = \frac{6b}{nm_1c_2}\left[n^2c_2^3 + 3n(n-1)(n+4)c_3^2 + 3n(n^2 + 3n - 12)c_2^4 + (3n^4 + 15n^3 - 20n^2 - 120n + 160)c_6\right] = O(n^{-2})$$

and

$$\text{Var}\left(\frac{\sqrt{\hat{c}_4}}{c_2}\right) = \frac{8b}{nm_2c_2}\left[f_1c_4^2 + f_2c_2c_3^2 + f_3c_2^2c_4 + f_4c_2^4 + f_5c_3c_5 + f_6c_2c_6 + f_7c_8\right] = O(n^{-2}),$$

where

$$f_1 = n^2(n^2 + n + 2), \quad f_2 = 8n^2(n+1)(n-3)(n+6),$$
$$f_3 = 2n^2(2n^3 + 11n^2 - 47n + 54),$$
$$f_4 = n(6n^5 + 40n^4 - 85n^3 - 631n^2 + 726n + 1224),$$
$$f_5 = 8n(n+1)(n-3)(n+6)(n^2 + 4n - 16),$$
$$f_6 = 4n(n^5 + 10n^4 - 11n^3 - 220n^2 + 276n + 480)$$
and
$$f_7 = 2(2n^7 + 23n^6 + 38n^5 - 423n^4 - 992n^3 + 4066n^2 - 420n - 5040).$$

Obviously, $\hat{c}_2/c_2$ is $L^2$, so that in probability, convergent to 1. All the asymptotic equivalence results in Theorem 3.1 and Theorem 3.5 can be proved similarly, and
we only include the proof of $(\hat{C}_1 - c_1) / \sqrt{c_2} = O_p(b^{-1/2})$ here. By Chebyshev’s inequality and under the condition $A_2$, we have

$$\Pr \left( \frac{\sqrt{b}(\hat{C}_1 - c_1)}{\sqrt{c_2}} > M \right) \leq \frac{b \text{Var}(\hat{C}_1/c_2^{1/2})}{M^2} = \frac{2b}{nM^2} \rightarrow \frac{2\gamma_0}{M^2},$$

for any $M > 0$. Therefore, for any $\varepsilon > 0$, there exists $M(\sqrt{2\gamma_0/\varepsilon})$, such that the above probability is bounded by $\varepsilon$. For similar results see Srivastava (2005) and Hyodo et al. (2014).

**Theorem A.1.** Under the high-dimensional asymptotic frameworks $A_1$ and $A_3$, the characteristic function of $T$ can be expanded as

$$\phi_T(t) = e^{\frac{i}{2}t^2} \left\{ 1 + \frac{1}{\sqrt{b}} \sum_{j=3}^a t^j \eta_j + \frac{1}{b} \left( t^4 \eta_4 + t^6 \eta_6^2/2 \right) + O(b^{-3/2}) \right\},$$

as $n, b \rightarrow \infty$, where $i = \sqrt{-1}$.

**Proof of Theorem A.1.** Let $Z = \Sigma^{-1/2}(X - \mu) \sim N_{ab}(0, I_{ab})$. Then

$$T = \frac{\text{tr}(T_1 D)}{\sqrt{2b}c_2^{1/2}} \left( Z' \Sigma^{1/2} K \Sigma^{1/2} Z + 2\mu' K \Sigma^{1/2} Z - c_1 \right).$$

Denote $\sigma_1^2 = 2\delta_2 c_2$ with $\sigma_1 > 0$. The characteristic function of $T$ is derived as

$$\phi_T(t) = \mathbb{E} \left[ \exp \left\{ \frac{it}{\sigma_1} \left\{ \text{tr}(T_1 D) \right\}^{-1} \left( Z' \Sigma^{1/2} K \Sigma^{1/2} Z + 2\mu' K \Sigma^{1/2} Z - c_1 \right) \right\} \right]$$

$$= \exp \left( \frac{-itc_1}{\sigma_1} \right) \int_Z (2\pi)^{-ab/2} \exp \left\{ -\frac{1}{2} Z'MZ + \frac{2it}{\sigma_1} \{ \text{tr}(T_1 D) \}^{-1} \mu' K \Sigma^{1/2} Z \right\} dZ$$

$$= \exp \left( \frac{-itc_1}{\sigma_1} \right) \int_Z (2\pi)^{-ab/2} \exp \left\{ -\frac{1}{2} \left( Z - \frac{2it}{\sigma_1} \{ \text{tr}(T_1 D) \}^{-1} M^{-1} \Sigma^{1/2} Z \right) \cdot M \cdot \left( Z - \frac{2it}{\sigma_1} \{ \text{tr}(T_1 D) \}^{-1} M^{-1} \Sigma^{1/2} K \mu \right)\right.$$  

$$\left. + \frac{2\delta_2^2}{\sigma_1^2} \{ \text{tr}(T_1 D) \}^{-2} \mu' K \Sigma^{1/2} M^{-1} \Sigma^{1/2} K \mu \right\} dZ$$

$$= \exp \left( \frac{-itc_1}{\sigma_1} \right) |M|^{-1/2} \exp \left\{ \frac{2\delta_2^2}{\sigma_1^2} \{ \text{tr}(T_1 D) \}^{-2} \mu' K \Sigma^{1/2} M^{-1} \Sigma^{1/2} K \mu \right\},$$

where $M = I - \frac{2it}{\sigma_1} \{ \text{tr}(T_1 D) \}^{-1} \Sigma^{1/2} K \Sigma^{1/2}$ and $|M|$ is the determinant of $M$. Let $\alpha_i$’s be the eigenvalues of $T_1 D$, $\beta_j$’s be the eigenvalues of $\Sigma^{1/2} T_2 \Sigma^{1/2}$, then

$$|M| = \prod_{i=1}^a \prod_{j=1}^b \left( 1 - \frac{2it}{\sigma_1} \{ \text{tr}(T_1 D) \}^{-1} \alpha_i \beta_j \right).$$
Thus, by Taylor’s series expansion, we have

\[
\log |M|^{-1/2} = -\frac{1}{2} \sum_{i=1}^{a} \sum_{j=1}^{b} \log \left(1 - \frac{2it}{\sigma_1} \{\text{tr}(T_1 D)\}^{-1} \alpha_i \beta_j \right)
\]

\[
= \frac{1}{2} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{\infty} \frac{2^k (it)^k}{k \sigma_1^k} \cdot \frac{\alpha_i \beta_j^k}{\{\text{tr}(T_1 D)\}^k}
\]

\[
= \sum_{k=1}^{\infty} \frac{2^{k-1} (it)^k}{k} \frac{\delta_k c_k}{\delta_2^{k/2} c_2^{k/2}}
\]

\[
= \frac{itc_1}{\sigma_1} + \frac{1}{2} t^2 + \frac{1}{\sqrt{b}} \delta_1^2 t^3 \eta_3 + \frac{1}{b} t^4 \eta_4 + O(b^{-3/2}),
\]

where the last equality holds by condition $A_1$ and equation (2).

Since

\[
M^{-1} = I + \frac{2it}{\sigma_1} \{\text{tr}(T_1 D)\}^{-1} \Sigma^{1/2} K \Sigma^{1/2} + \frac{4i^2 t^2}{\sigma_1^2} \{\text{tr}(T_1 D)\}^{-2} \Sigma^{1/2} K \Sigma \Sigma^{1/2} + \ldots
\]

Under assumption $A_3$, as $n, b \to \infty$, we have

\[
\mu' K \Sigma K \mu \leq \mu' K \mu \cdot \sqrt{\text{tr}(K \Sigma)^2} = \mu' K \mu \cdot \sqrt{\text{tr}(T_1 D)^2} \cdot c_2,
\]

and, therefore,

\[
\frac{1}{\sigma_1^2} \{\text{tr}(T_1 D)\}^{-2} \mu' K \Sigma K \mu \leq \frac{\mu' K \mu}{2 \sqrt{\text{tr}(T_1 D)^2} \cdot c_2} = O(b^{-3/2}).
\]

Similarly,

\[
\frac{1}{\sigma_1^3} \{\text{tr}(T_1 D)\}^{-3} \mu' K \Sigma K \Sigma K \mu \leq \frac{1}{\sigma_1^2} \{\text{tr}(T_1 D)\}^{-2} \mu' K \Sigma K \mu = O(b^{-3/2}),
\]

\[
\frac{1}{\sigma_1^4} \{\text{tr}(T_1 D)\}^{-4} \mu' K \Sigma K \Sigma K \Sigma K \mu \leq \frac{1}{\sigma_1^2} \{\text{tr}(T_1 D)\}^{-2} \mu' K \Sigma K \mu = O(b^{-3/2}),
\]

and so on. Thus, we have

\[
\frac{1}{\sigma_1^7} \{\text{tr}(T_1 D)\}^{-7} \mu' K \Sigma^{1/2} M^{-1} \Sigma^{1/2} K \mu \} = O(b^{-3/2}).
\]

By Taylor’s series expansion, the characteristic function of $T$ can be expanded as

\[
\phi_T(t) = e^{\frac{1}{2} t^2} \left(1 + \frac{1}{\sqrt{b}} t^3 \eta_3 + \frac{1}{b} t^4 \eta_4 + \frac{1}{b} t^5 \eta_5^2 / 2 + O(b^{-3/2})\right).
\]

\[\square\]

**Proof of Theorem 3.2.** Denote $H^* = \{\text{tr}(T_1 D)\}^{-1} H$. Note that

\[
\frac{c_2}{c_1} = 1 + \frac{1}{b} W \quad \text{and} \quad \frac{c_1 - c_2}{\sqrt{2} \delta_2 c_2} = \frac{V}{\sqrt{b}}.
\]
By Taylors’ expansion, we have
\[
\left( \frac{\hat{c}_2}{c_2} \right)^{-1/2} = 1 - \frac{1}{2b} W + O_p(b^{-2}).
\]

Then
\[
\hat{T} = \frac{H^* - \hat{c}_1}{\sqrt{2\delta_2c_2}} = \frac{(H^* - c_1) - (\hat{c}_1 - c_1)}{\sqrt{2\delta_2c_2}}
= \frac{(H^* - c_1) - (\hat{c}_1 - c_1)}{\sqrt{2\delta_2c_2}} \left\{ 1 - \frac{1}{2b} W + O_p(b^{-2}) \right\}
= T - \frac{V}{\sqrt{b}} - \frac{TW}{2b} + O_p(b^{-3/2}),
\]
where \( T = \frac{(H^* - c_1)}{\sqrt{2\delta_2c_2}} \). So the characteristic function of \( \hat{T} \) is
\[
\phi_{\hat{T}}(t) = \mathbb{E}[e^{it\hat{T}}] = \mathbb{E} \left[ e^{itT} \cdot e^{i(1 - \frac{1}{2b} W + O_p(b^{-3/2}))} \right]
= \mathbb{E} \left[ e^{itT} \cdot \left\{ 1 - \frac{itV}{\sqrt{b}} - \frac{itTW}{2b} + \frac{t^2V^2}{b} + O_p(b^{-3/2}) \right\} \right]
= \mathbb{E}[e^{itT}] + \frac{t^2V^2}{2b} \mathbb{E}[e^{itT}] \mathbb{E}[V^2] + O(b^{-3/2}),
\]
since \( T \) is independent with both \((V, W)\) and \( \mathbb{E}[V] = \mathbb{E}[W] = 0 \).

Finally, using Theorem A.1 and the fact (see Proof of Theorem 3.1) that
\[
\mathbb{E}[V^2] = \frac{b}{2\delta_2c_2} \text{Var}(\hat{c}_1) = \frac{b}{n\delta_2} = \frac{\gamma}{\delta_2},
\]
we have the desired result.

**Proof of Theorem 3.3.** In view of the expanded characteristic function in Theorem 3.2, the desired result follows from a term-by-term inversion by using the inversion formula
\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (-ut)e^{-ut}e^{-t^2/2} dtdu = \int_{-\infty}^{\infty} \frac{d}{du} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ut}e^{-t^2/2} dt \right\} du
= \int_{-\infty}^{\infty} \phi^{(j)}(u) du = \Phi^{(j)}(x)
\]
for any nonnegative integer \( j \). For more details see, for example, Theorem 3 in (Shiryaev and Chibisov, 2016, pp. 340).

**Proof of Theorem 4.1.**
\[
\hat{T} - \frac{\text{tr}(T_1D)}{\sqrt{2\delta_2c_2}} \Delta = T - \frac{V}{\sqrt{b}} - \frac{WU}{2b} + O_p(b^{-3/2}),
\]
where \( U = \frac{\text{tr}(T_1D)}{\sqrt{2\delta_2c_2}} \). Since \( U \) and \( W \) are independent and \( \mathbb{E}(U) = \mathbb{E}(V) = \mathbb{E}(W) = 0 \), application of Theorem A.1 leads to,
\[ \beta(\Delta) = P(\hat{T} > u(z_\alpha)) \]
\[ = P \left( \hat{T} - \frac{\{\text{tr}(T_1D)\}^{-1}\Delta}{\sqrt{2\delta_2c_2}} > u(z_\alpha) - \frac{\{\text{tr}(T_1D)\}^{-1}\Delta}{\sqrt{2\delta_2c_2}} \right) \]
\[ = 1 - G_{\hat{T}} \left( u(z_\alpha) - \frac{\{\text{tr}(T_1D)\}^{-1}\Delta}{\sqrt{2\delta_2c_2}} \right), \]
as to be proved. \hfill \Box

**Acknowledgements**

The authors are grateful to the two anonymous referees for critically reading the original and revised versions of the manuscript and their valuable suggestions that led to great improvements. Authors are also thankful to the editor and the associate editor for their constructive comments and orderly handling of the manuscript. The research of Xiaoli Kong was partially done while she was pursuing her Ph.D. She would like to express her gratitude to the Department of Statistics, University of Kentucky.

**References**

Bai, Z. and Saranadasa, H. (1996). Effect of high dimension: By an example of a two sample problem. *Statistica Sinica*, 6(2):311–329. MR1399305

Bock, R. (1963). *Multivariate analysis of variance of repeated measurements*. (Problems in measuring change. Editor C. W. Harris). University of Wisconsin Press, Madison, Wisconsin.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *The Annals of Mathematical Statistics*, 25(2):290–302. MR0061787

Brunner, E., Bathke, A. C., and Placzek, M. (2012). Estimation of box’s \( \varepsilon \) for low- and high-dimensional repeated measures designs with unequal covariance matrices. *Biometrical Journal*, 54(3):301–316. MR2931076

Brunner, E., Dette, H., and Munk, A. (1997). Box-type approximations in non-parametric factorial designs. *Journal of the American Statistical Association*, 92(440):1494–1502. MR1615259

Brunner, E., Konietzchke, 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. MR3731671

Brunner, E., Munzel, U., and Puri, M. L. (1999). Rank-score tests in factorial designs with repeated measures. *Journal of Multivariate Analysis*, 70(2):286–317. MR1711514

Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(2):349–372. MR3164870
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(2):808–835. MR2604697

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(8):724–742. MR2901796

Collier, R. O., Baker, F. B., Mandeville, G. K., and Hayes, T. F. (1967). Estimates of test size for several test procedures based on conventional variance ratios in the repeated measures design. *Psychometrika*, 32(3):339–353. MR0219171

Fujikoshi, Y. (1973). Asymptotic formulas for the distributions of three statistics for multivariate linear hypothesis. *Annals of the Institute of Statistical Mathematics*, 25(1):423–437. MR0331639

Geisser, S. and Greenhouse, S. W. (1958). An extension of box’s results on the use of the $F$ distribution in multivariate analysis. *The Annals of Mathematical Statistics*, 29(3):885–891. MR0096331

Happ, M., Harrar, S. W., and Bathke, A. C. (2015). Inference for low- and high-dimensional multi-group repeated measures designs with unequal covariance matrices. *Biometrical Journal*, 58(4):810–830. MR3527417

Harrar, S. W. and Kong, X. (2016). High-dimensional multivariate repeated measures analysis with unequal covariance matrices. *Journal of Multivariate Analysis*, 145:1–21. MR3459934

Hill, G. W. and Davis, A. W. (1968). Generalized asymptotic expansions of cornish-fisher type. *The Annals of Mathematical Statistics*, 39(4):1264–1273. MR0226762

Huynh, H. (1978). Some approximate tests for repeated measurement designs. *Psychometrika*, 43(2):161–175. MR0514724

Huynh, H. and Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact $F$-distributions. *Journal of the American Statistical Association*, 65(332):1582–1589.

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(1):69–82.

Hyodo, M., Takahashi, S., and Nishiyama, T. (2014). Multiple comparisons among mean vectors when the dimension is larger than the total sample size. *Communications in Statistics – Simulation and Computation*, 43(10):2283–2306. MR3223670

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. MR3049913

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. MR3372569

Lecoutre, B. (1991). A correction for the formula approximate test in repeated measures designs with two or more independent groups. *Journal of Educational and Behavioral Statistics*, 16(4):371–372.
Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940. MR2985938
Maxwell, S. E. and Arvey, R. D. (1982). Small sample profile analysis with many variables. *Psychological Bulletin*, 92(3):778–785.
Pauly, M., Ellenberger, D., and Brunner, E. (2015). Analysis of high-dimensional one group repeated measures designs. *Statistics*, 49(6):1243–1261. MR3415218
Rencher, A. C. and Christensen, W. F. (2012). *Methods of multivariate analysis*. John Wiley & Sons, United States, 3rd edition. MR2962097
Shiryaev, A. and Chibisov, D. (2016). *Probability-1*. Graduate Texts in Mathematics. Springer New York. MR3467826
Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *Journal of the Japan Statistical Society*, 35(2):251–272. MR2328427
Stoloff, P. H. (1970). Correcting for heterogeneity of covariance for repeated measures designs of the analysis of variance. *Educational and Psychological Measurement*, 30(4):909–924.
Takahashi, S. and Shutoh, N. (2016). Tests for parallelism and flatness hypotheses of two mean vectors in high-dimensional settings. *Journal of Statistical Computation and Simulation*, 86(6):1150–1165. MR3441561
Wang, H. and Akritas, M. G. (2010a). Inference from heteroscedastic functional data. *Journal of Nonparametric Statistics*, 22(2):149–168. MR2598959
Wang, H. and Akritas, M. G. (2010b). Rank test for heteroscedastic functional data. *Journal of Multivariate Analysis*, 101(8):1791–1805. MR2651955
Wang, H., Higgins, J., and Blasi, D. (2010). Distribution-free tests for no effect of treatment in heteroscedastic functional data under both weak and long range dependence. *Statistics & Probability Letters*, 80(5-6):390–402. MR2593579
Watamori, Y. (1990). On the moments of traces of wishart and inverted wishart matrices. *South African Statistical Journal*, 24(2):153–176. MR1128480
Yang, X. M., Yang, X. Q., and Teo, K. L. (2001). A matrix trace inequality. *Journal of Mathematical Analysis and Applications*, 263(1):327–331. MR1865285
Zhang, C., Bai, Z., Hu, J., and Wang, C. (2018). Multi-sample test for high-dimensional covariance matrices. *Communications in Statistics – Theory and Methods*, 48(13):3161–3177. MR3804216