Testing linear hypotheses of mean vectors for high-dimension data with unequal covariance matrices

Takahiro Nishiyama[a,*], Masashi Hyodo[b], Takashi Seo[a], Tatjana Pavlenko[c]

[a] *Department of Mathematical Information Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan*
[b] *Graduate School of Economics, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, 113-0033, Japan, JSPS Research Fellow*
[c] *Department of Mathematics, KTH Royal Institute of Technology, SE-100, 44, Stockholm, Sweden*

Abstract

We propose a new test procedure for testing linear hypothesis on the mean vectors of normal populations with unequal covariance matrices when the dimensionality, p exceeds the sample size N, i.e. $p/N \to c < \infty$. Our procedure is based on the Dempster trace criterion and is shown to be power and size consistent in high dimensions.

The asymptotic null and non-null distributions of the proposed test statistic are established in the high dimensional setting and improved estimator of the critical point of the test is derived using Cornish-Fisher expansion. As a special case, our testing procedure is applied to multivariate Behrens-Fisher problem. We illustrate the relevance and benefits of the proposed approach via Monte-Carlo simulation which show that our new test is comparable to, and in many cases is more powerful than, the tests for equality of means presented in the recent literature.

Keywords: Cornish-Fisher transform, Dempster trace criterion, High dimensionality, Multivariate Behrens-Fisher problem, (N, p)-asymptotics.

*Corresponding author. Tel.: +81-3-3260-4271; Fax: +81-3-3260-4293.
Email address: nishiyam@rs.kagu.tus.ac.jp (Takahiro Nishiyama)

1. Introduction

The problem of testing mean vectors is a part of many procedures of multivariate statistical analysis, such as multiple comparisons, MANOVA and classification. The standard testing technique is based on classical Hotelling's T^2 test which is known to have optimal performance properties in a large sample case, i.e assuming that the number of feature variables, p is fixed and is much smaller the sample size, N. However, in many practical applications of modern multivariate statistics (e.g. DNA microarray data) the number of feature p exceeds N, so that a straightforward use of T^2 statistics is impossible due to singularity of the sample covariance matrix. Thus, to cope with this high dimensional situation, it would be desirable to develop new tests for $N \leq p$, and investigate their asymptotic properties when both N and p are going to infinity; this asymptotic framework is also known as (N, p)-*asymptotics*.

There have been a series of important results on this testing problem. In particular, Bai and Saranadasa (1996) focus on (normal) the two-sample case with equal covariance matrix, and propose to use an estimator of Euclidean norm of the shift vector instead of T^2 statistics; they also establish the asymptotic normality of the test statistics assuming that p and N is of the same order. Later, by using the same approach as Bai and Saranadasa (1996), Aoshima and Yata (2011) derive the test for which the significance levels can be controlled with or without the assumption of normality of the data, that is robust for the model assumption. Unlike the above approach, Srivastava (2007) suggests F-type test statistics based on the Moore-Penrose inverse of the singular sample covariance matrix, and Srivastava and Du (2008) have developed the test procedure based on the Dempster's trace criterion (D-criterion) (Dempster (1958,1960)) under the assumption of variable independence. Also it is important to note, that both these procedures are less restrictive than that of Bai and Saranadasa (1996) in a sense that they allow p grow faster than N according to (N, p)-asymptotic framework.

Motivated by the previous literature and as part of effort in developing testing procedures with stable characteristics in high-dimensions, we focus on testing linear hypotheses of mean vectors for high-dimensional data with unequal covariance matrices. Our main objective in this paper is to show that our newly derived test statistics based on the Dempster trace criterion has a number of attractive asymptotic properties and demonstrates good performance in large (N, p). We state the asymptotic distribution of the derived test statistics under both the null and the local alternative hypotheses, and provide the explicit expressions for asymptotic power of the test in terms of δ when $N = O(p^\delta)$ and $0 < \delta < 1$. To further improve the test performance, Cornish-Fisher approximation of the upper 100α percentile of the test is provided in (N, p)-asymptotics. We also apply our new test procedure to the multivariate Behrens-Fisher problem and compare its performance with the above-mentioned testing procedures from resent literature.

The rest of the paper is organized as follows. Section 2 provides description of the new test and main asymptotic results. Section 3 considers the application to the Behrens-Fisher problem. In Section 4, the level hypothesis is tested and the attained significance levels of the newly derived test is analyzed for a number of high dimensional scenarios. At last, we provide some concluding remarks. The proofs of theorems and lemmas stated in Section 2 are given in the Appendix A and the Appendix B.

2. Description of the new test and asymptotic properties

Let $\boldsymbol{x}_1^{(i)}, \ldots, \boldsymbol{x}_{N_i}^{(i)}$, $i = 1, \ldots, k$ be N_i samples from $N_p(\boldsymbol{\mu}_i, \Sigma_i)$, respectively. We are interesting in the linear testing the hypothesis

$$H_0 : \sum_{i=1}^{k} \beta_i \boldsymbol{\mu}_i = \mathbf{0} \quad \text{vs.} \quad H_1 :\neq H_0, \tag{2.1}$$

where $\beta_1, \ldots, \beta_k$ are given scalars and covariance matrices Σ_i's are assumed to be unequal. In this study, we consider Bennett (1951)'s transform derived Anderson (2003) for k-sample case (see, Bennett (1951), Anderson (2003)). For convenience, let N_1 be the smallest. Then, for $j = 1, \ldots, N_1$, we define

$$\boldsymbol{y}_j = \beta_1 \boldsymbol{x}_j^{(1)} + \sum_{\ell=2}^{k} \beta_\ell \sqrt{\frac{N_1}{N_\ell}} \left(\boldsymbol{x}_j^{(\ell)} - \frac{1}{N_1} \sum_{m=1}^{N_1} \boldsymbol{x}_m^{(\ell)} + \frac{1}{\sqrt{N_1 N_\ell}} \sum_{n=1}^{N_\ell} \boldsymbol{x}_n^{(\ell)} \right).$$

Especially, when $N_1 = \cdots = N_k$, $\boldsymbol{y}_j = \sum_{\ell=1}^{k} \beta_\ell \boldsymbol{x}_j^{(\ell)}$. The expected value of $\boldsymbol{y}_j$ and the covariance matrix of $\boldsymbol{y}_\ell$ and $\boldsymbol{y}_m$ are

$$\begin{aligned} \mathrm{E}\left(\boldsymbol{y}_j\right) &= \sum_{i=1}^{k} \beta_i \boldsymbol{\mu}_i, \\ \mathrm{Cov}\left(\boldsymbol{y}_\ell, \boldsymbol{y}_m\right) &= \delta_{\ell m} \left(\sum_{i=1}^{k} \frac{\beta_i^2 N_1}{N_i} \Sigma_i \right), \end{aligned}$$

respectively, where $\delta_{\ell m}$ is Kronecker's delta. We further set

$$\sum_{i=1}^{k} \beta_i \boldsymbol{\mu}_i \equiv \boldsymbol{\mu}, \quad \sum_{i=1}^{k} \frac{\beta_i^2 N_1}{N_i} \Sigma_i \equiv \Sigma,$$

respectively, and note that $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N_1}$ are independent and identically distributed as $N_p(\boldsymbol{\mu}, \Sigma)$. Thus, we may consider testing

$$H_0 : \boldsymbol{\mu} = \mathbf{0} \quad \text{vs.} \quad H_1 : \boldsymbol{\mu} \neq \mathbf{0}, \tag{2.2}$$

which is equivalent to (2.1).

For testing (2.2) under the assumption that $p > n_1 = N_1 - 1$, we define a new test statistic as follows

$$T_D = \frac{N_1 \overline{\boldsymbol{y}}' \overline{\boldsymbol{y}}}{\text{tr} S}, \tag{2.3}$$

where

$$\overline{\boldsymbol{y}} = \frac{1}{N_1} \sum_{i=1}^{N_1} \boldsymbol{y}_i, \quad S = \frac{1}{n_1} \sum_{i=1}^{N_1} (\boldsymbol{y}_i - \overline{\boldsymbol{y}})(\boldsymbol{y}_i - \overline{\boldsymbol{y}})'.$$

T_D is based on Dempster trace criterion and does not require any restrictions on the relation between the dimension and sample size.

At first, we derive the null distribution of the statistic (2.3) under the following (N, p)-asymptotic framework:

$$(\text{A.1}) : p, n_1 \to \infty, \quad \frac{p}{n_1} \to c \in (0, \infty).$$

Further, in addition to $(A1)$, we assume that

$$(\text{A.2}) : \lim_{p \to \infty} \frac{\text{tr} \Sigma^i}{p} \to a_i \in (0, \infty), \ i = 1, \ldots, 4,$$
$$(\text{A.3}) : \lim_{p \to \infty} \frac{\text{tr} \Sigma^i}{p} \to a_i \in (0, \infty), \ i = 1, \ldots, 8.$$

Let

$$\widetilde{T}_D = \sqrt{p} \left\{ \frac{N_1 \overline{\boldsymbol{y}}' \overline{\boldsymbol{y}}}{\text{tr} S} - 1 \right\}. \tag{2.4}$$

The following theorem provides asymptotic null distribution of $\widetilde{T}_D/\sigma_D$ where

$$\sigma_D = \sqrt{\frac{2a_2}{a_1^2}} = \frac{\sqrt{2\mathrm{tr}\Sigma^2/p}}{\mathrm{tr}\Sigma/p}.$$

Theorem 2.1. *When the null hypothesis $H_0 : \boldsymbol{\mu} = \mathbf{0}$ is true, the distribution function of $\widetilde{T}_D/\widehat{\sigma}_D$ can be expanded in the asymptotic framework* (A.1) *and under assumption* (A.3) *as*

$$\begin{aligned} P\left(\frac{\widetilde{T}_D}{\widehat{\sigma}_D} \le z\right) &= \Phi(z) - \phi(z)\Big[\frac{1}{\sqrt{p}}c_3h_2(z) \\ &+\frac{1}{p}\{c_4h_3(z) + c_6h_5(z)\} + \frac{1}{n_1}c_2h_1(z)\Big] + O(p^{-3/2}), \end{aligned} \tag{2.5}$$

where $\widehat{\sigma}_D$ is obtained by replacing a_1 and a_2 with their estimators, $\Phi(z)$ is the distribution function of the standard normal distribution,

$$c_2 = \frac{1}{2}, \quad c_3 = \frac{\sqrt{2}a_3}{3\sqrt{a_2^3}}, \quad c_4 = \frac{a_4}{2a_2^2}, \quad c_6 = \frac{a_3^2}{9a_2^3},$$

and $h_i(z)$'s $(i = 1, \dots, 6)$ are the Hermite polynomials given by

$$\begin{aligned} &h_1(z) = z, \quad h_2(z) = z^2 - 1, \quad h_3(z) = z^3 - 3z, \quad h_4(z) = z^4 - 6z^2 + 3, \\ &h_5(z) = z^5 - 10z^3 + 15z, \quad h_6(z) = z^6 - 15z^4 + 45z^2 - 15. \end{aligned}$$

Proof. See, Appendix A.1.

In practice, c_i's and a_i's are unknown. Hence, to use the result of Theorem 2, we need their estimators that are expected to be good in high-dimension setting. As sample counterparts of a_i's, we use their (N, p)-consistent and unbiased estimators derived in Srivastava(2005), Srivastava and Yanagihara (2010) and Hyodo, Takahashi and Nishiyama (2012) as

$$\widehat{a}_1 = \frac{\mathrm{tr}S}{p},$$

$$
\begin{aligned}
\widehat{a}_2 &= \frac{n_1^2}{p(n_1-1)(n_1+2)}\left\{\mathrm{tr}S^2 - \frac{(\mathrm{tr}S)^2}{n_1}\right\}, \\
\widehat{a}_3 &= \frac{n_1}{(n_1-1)(n_1-2)(n_1+2)(n_1+4)} \\
&\qquad\qquad \times\left\{\frac{\mathrm{tr}S^3}{p} - 3(n_1+2)(n_1-1)\widehat{a}_1\widehat{a}_2 - n_1p^2\widehat{a}_1^3\right\}, \\
\widehat{a}_4 &= \frac{1}{b_0}\left(\frac{\mathrm{tr}S^4}{p} - pb_1\widehat{a}_1 - p^2b_2\widehat{a}_1^2\widehat{a}_2 - pb_3\widehat{a}_2^2 - n_1p^3\widehat{a}_1^4\right),
\end{aligned}
$$

where

$$
b_0 = n_1(n_1^3+6n_1^2+21n_1+18), \quad b_1 = 2n_1(2n_1^2+6n_1+9),
$$
$$
b_2 = 2n_1(3n_1+2), \quad b_3 = n_1(2n_1^2+5n_1+7).
$$

Type 1 error of the asymptotic test based on using the main term of (2.5) can be essentially improved by the corrected estimator of the upper 100α-percentile of $\widetilde{T}_D/\widehat{\sigma}$. This correction is achieved by using Cornish-Fisher expansion.

Corollary 2.2. *Under the asymptotic framework* (A.1) *and assumption* (A.3)*, Cornish-Fisher expansion of the estimated upper percentile of* $\widetilde{T}_D/\widehat{\sigma}$ *is derived as*

$$
\begin{aligned}
\widehat{z}(\alpha) &= z_\alpha + \frac{1}{\sqrt{p}}\frac{\sqrt{2}\widehat{a}_3}{3\sqrt{\widehat{a}_2^3}}(z_\alpha^2-1) \\
&+\frac{1}{p}\left\{\frac{\widehat{a}_4}{2\widehat{a}_2^2}z_\alpha(z_\alpha^2-3) - \frac{2\widehat{a}_3^2}{9\widehat{a}_2^3}z_\alpha(2z_\alpha^2-5)\right\} + \frac{1}{2n_1}z_\alpha + O_p(p^{-3/2}), \quad (2.6)
\end{aligned}
$$

where z_α *is the upper* $100\alpha\%$ *percentile of the standard normal distribution.*

Proof. See, Appendix A.2.

Next, we state the asymptotic distribution of the test statistic T_D under the local alternative. We state

$$
H_1^{L(\delta)} : N\boldsymbol{\mu}'\boldsymbol{\mu} = O(p^\delta),\ N\boldsymbol{\mu}'\Sigma\boldsymbol{\mu} = O(p^\delta),\ 0<\delta<1, \quad (2.7)
$$

where $N = \sum_{i=1}^{k} N_i$ and assume that $N_1/N_i \to c_i \in (0, \infty)$ $(i = 1, \ldots, k)$. Then, under the hypothesis $H_1^{L(\delta)}$, the test statistic can be expressed as

$$T_D^* = \frac{N_1 \overline{\boldsymbol{y}}' \overline{\boldsymbol{y}}}{\mathrm{tr} S} - 1 - \frac{N_1 \boldsymbol{\mu}' \boldsymbol{\mu}}{\mathrm{tr} \Sigma}. \tag{2.8}$$

Theorem 2.3 provides the limiting distribution of T_D^*/σ_D^*, where

$$\sigma_D^* = \sqrt{\frac{2\mathrm{tr}\Sigma^2 + 4N_1 \boldsymbol{\mu}' \Sigma \boldsymbol{\mu}}{(\mathrm{tr}\Sigma)^2}},$$

under the hypothesis $H_1^{L(\delta)}$.

Theorem 2.3. *When the local alternative hypothesis $H_1^{L(\delta)}$ is true, the distribution of T_D^*/σ_D^* is asymptotically normal, i.e.*

$$\frac{T_D^*}{\sigma_D^*} = \frac{N_1 \overline{\boldsymbol{y}}' \overline{\boldsymbol{y}}/trS - 1 - N_1 \boldsymbol{\mu}' \boldsymbol{\mu}/tr\Sigma}{\sqrt{(2tr\Sigma^2 + 4N_1 \boldsymbol{\mu}' \Sigma \boldsymbol{\mu})/tr\Sigma^2}} \xrightarrow{d} N(0, 1),$$

in the asymptotic framework (A.1) *and under assumption* (A.2).

Proof. See, Appendix A.3.

Using the asymptotic distribution under the alternative hypothesis, we are able to describe the (N, p)-asymptotic behavior of the power function of our test statistic which is collaborated in the following theorem.

Theorem 2.4. *Let*

$$Power_\alpha(\widetilde{T}_D, \delta) = P\left(\frac{\widetilde{T}_D}{\sigma_D} \geq z_\alpha \mid H_1^{L(\delta)}\right),$$

be the power function of $\widetilde{T}_D$. Then, in the asymptotic framework (A.1) *and assumptions* (A.2)*,*

$$\begin{aligned}
&\text{(i)} \quad Power_\alpha(\widetilde{T}_D, \delta) \longrightarrow \alpha, \quad \textit{if} \;\; 0 < \delta < 1/2, \\
&\text{(ii)} \quad Power_\alpha(\widetilde{T}_D, \delta) \longrightarrow \Phi\left(\frac{N_1 \boldsymbol{\mu}' \boldsymbol{\mu}}{\sqrt{2tr\Sigma^2}} - z_\alpha\right), \quad \textit{if} \;\; \delta = 1/2, \\
&\text{(iii)} \quad Power_\alpha(\widetilde{T}_D, \delta) \longrightarrow 1, \quad \textit{if} \;\; 1/2 < \delta < 1.
\end{aligned}$$

Proof. See, Appendix A.4.

In words, this theorem claims that the test statistic $\widetilde{T}_D$ is (N, p)-consistent.

3. New test procedure for multivariate Behrens-Fisher problem

In this section, we focus on an important special case of the testing problem (2.1). We consider testing equality of mean vectors of two normal populations with unequal covariance matrices, that is, we consider testing the hypothesis

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{vs.} \quad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2. \tag{3.1}$$

We note that (3.1) is the special case that $k = 2$, $\beta_1 = 1$ and $\beta_2 = -1$ in (2.1). This problem is known as the multivariate Behrens-Fisher problem, and many authors have discussed (see, e.g., Bennett (1951), Johnson and Weerahandi (1988) and Yanagihara and Yuan (2005)). Also, for high-dimensional data, testing equality of mean vectors of two populations have been discussed by Bai and Saranadasa (1996), Srivastava (2007), Chen and Qin (2010), Aoshima and Yata (2011), and so on. Especially, Chen and Qin (2010) and Aoshima and Yata (2011) gave a test statistic for the multivariate Behrens-Fisher problem in high-dimension setting.

We propose a new test statistic for this problem by using the idea stated in Section 2, that is, we consider the following statistic

$$T_D = \frac{N_1 \overline{\boldsymbol{y}}' \overline{\boldsymbol{y}}}{\mathrm{tr} S},$$

where, for $j = 1 \ldots, N_1$,

$$\boldsymbol{y}_j = \boldsymbol{x}_j^{(1)} - \sqrt{\frac{N_1}{N_2}} \boldsymbol{x}_j^{(2)} + \frac{1}{\sqrt{N_1 N_2}} \sum_{m=1}^{N_1} \boldsymbol{x}_m^{(2)} - \frac{1}{N_2} \sum_{n=1}^{N_2} \boldsymbol{x}_n^{(2)}. \tag{3.2}$$

Then we note that $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N_1}$ are independent and identically distributed as $N_p(\boldsymbol{\mu}, \Sigma)$ where

$$\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \ \Sigma = \Sigma_1 + \frac{N_1}{N_2} \Sigma_2, \tag{3.3}$$

respectively (see, Bennett (1951)), that is, (3.1) is equivalent to the following hypothesis:

$$H_0 : \boldsymbol{\mu} = \mathbf{0} \quad \text{vs.} \quad H_1 : \boldsymbol{\mu} \neq \mathbf{0}.$$

Therefore from Theorem 2.1, Corollary 2.2 and Theorem 2.4, we have following corollary.

Corollary 3.1. *Suppose that $\widetilde{T}_D$, $\widehat{a}_i$'s, c_i's and $Power_\alpha(\widetilde{T}_D, \delta)$ are defined according to (3.2) and (3.3). Then, under the asymptotic framework* (A.1) *and assumption* (A.3) *the asymptotic distribution of $\widetilde{T}_D$ under the null hypothesis $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ and Cornish-Fisher expansion of the upper percentile are derived as follows*

$$\begin{aligned} P\left(\frac{\widetilde{T}_D}{\widehat{\sigma}_D} \leq z\right) &= \Phi(z) - \phi(z)\Big[\frac{1}{\sqrt{p}} c_3 h_2(z) \\ &\quad + \frac{1}{p}\{c_4 h_3(z) + c_6 h_5(z)\} + \frac{1}{n_1} c_2 h_1(z)\Big] + O(p^{-3/2}), \\ \widehat{z}(\alpha) &= z_\alpha + \frac{1}{\sqrt{p}} \frac{\sqrt{2}\widehat{a}_3}{3\sqrt{\widehat{a}_2^3}}(z_\alpha^2 - 1) \\ &+ \frac{1}{p}\Big\{\frac{\widehat{a}_4}{2\widehat{a}_2^2} z_\alpha(z_\alpha^2 - 3) - \frac{2\widehat{a}_3^2}{9\widehat{a}_2^3} z_\alpha(2z_\alpha^2 - 5)\Big\} + \frac{1}{2n_1} z_\alpha + O_p(p^{-3/2}), \end{aligned} \tag{3.4}$$

respectively, and for the asymptotic power of $\widetilde{T}_D$ we have

$$\begin{aligned} &\text{(i)} \quad Power_\alpha(\widetilde{T}_D, \delta) \longrightarrow \alpha, \quad \textit{if} \ \ 0 < \delta < 1/2, \\ &\text{(ii)} \quad Power_\alpha(\widetilde{T}_D, \delta) \longrightarrow \Phi\left(\frac{N_1 \boldsymbol{\mu}'\boldsymbol{\mu}}{\sqrt{2 tr \Sigma^2}} - z_\alpha\right), \quad \textit{if} \ \ \delta = 1/2, \\ &\text{(iii)} \quad Power_\alpha(\widetilde{T}_D, \delta) \longrightarrow 1, \quad \textit{if} \ \ 1/2 < \delta < 1. \end{aligned}$$

4. Simulation study

A simulation study shows the effectiveness of the suggested test statistics in high dimension. We first provide a study justifying accuracy of the approximation for the critical point derived in Corollary 2.2 of our testing procedure by simulating the Attained Significance Level (ASL), or size of the test. We draw k independent samples of size $N_i = 10(1+i)$ and $N_i = 20(1+i)$, where $i = 1, \ldots, k$ valid p-dimensional normal distributions under the null hypothesis (i.e. (2.1)). We replicate this $r = 10^5$ times, and using $\widetilde{T}_D$ from (2.4) calculate

$$ASL_\alpha^1\left(\widetilde{T}_D\right) = \frac{\sharp \text{ of } \left(\widetilde{T}_D/\widehat{\sigma} > z_\alpha \mid H_0 \text{ is true}\right)}{r},$$

and

$$ASL_\alpha^2\left(\widetilde{T}_D\right) = \frac{\sharp \text{ of } \left(\widetilde{T}_D/\widehat{\sigma} > \widehat{z}(\alpha) \mid H_0 \text{ is true}\right)}{r},$$

denoting the ASL of $\widetilde{T}_D$ where z_α is the upper 100α percentile of the standard normal distribution and $\widehat{z}(\alpha)$ is the corrected value of the upper 100α percentile given by (2.6).

Tables 1-4 provide the results for a number of high-dimensional scenarios and an assortment of null hypothesis H_0 specified by the choice of β_i's for $i = 1, \ldots, k$. Also, we set up the covariance structures $\Sigma_1 = I$, $\Sigma_2 = (0.2^{|i-j|})$, $\Sigma_3 = (0.5^{|i-j|})$ for the case $k = 3$ and $\Sigma_1 = I$, $\Sigma_2 = 2I$, $\Sigma_3 = (0.2^{|i-j|})$, $\Sigma_4 = (0.5^{|i-j|})$ for the case $k = 4$, respectively. For each table, the simulated size of $\widetilde{T}_D$ is systematically lower for the suggested corrected percentile $\widehat{z}(\alpha)$, and $\widehat{z}(\alpha)$ is closer to the selected nominal level α. Furthermore, in all the settings of H_0, the size of the test remains essentially the same when both p and k grows thereby validating our asymptotic results.

Further, we perform a series of power simulations to investigate consistency of our test and to demonstrate its improved performance under certain alternative hypotheses. From now on we focus on the simulations for the case of $k = 2$, representing multivariate Behrens-Fisher problem.

We provide examples for two cases of $H_1^{L(\delta)}$, $\Delta = 5$ and $\Delta = 10$ for the following settings of Σ_i:

$$\begin{aligned}
\text{Table 5} &: \ \Sigma_1 = I,\ \Sigma_2 = (0.5^{|i-j|}) \text{ and } \Delta = 5,\\
\text{Table 6} &: \ \Sigma_1 = I,\ \Sigma_2 = (0.5^{|i-j|}) \text{ and } \Delta = 10,\\
\text{Table 7} &: \ \Sigma_1 = (0.2^{|i-j|}),\ \Sigma_2 = (0.5^{|i-j|}) \text{ and } \Delta = 5,\\
\text{Table 8} &: \ \Sigma_1 = (0.2^{|i-j|}),\ \Sigma_2 = (0.5^{|i-j|}) \text{ and } \Delta = 10,
\end{aligned}$$

where $\Delta = \|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|^2$, and we recall that in our notations, see (2.7), $N\Delta = O(p^\delta)$.

Using the same number of replicates as above, we draw $k = 1, 2$ independent samples of size N_k under $H_1^{L(\delta)}$ and using T_D^* from (2.8) calculate the empirical power as

$$EP_\alpha\left(\widetilde{T}_D, \delta\right) = \frac{\sharp \text{ of } \left(\widetilde{T}_D/\widehat{\sigma}_D > \widehat{z}(\alpha) \mid H_1^{L(\delta)} \text{ is true}\right)}{r}.$$

To make comparisons of $\widetilde{T}_D$ with the test statistics defined in (Srivastava (2007) (S), Srivastava and Du (2008) (SD) and Aoshima and Yata (2011) (AY)), the same is repeated for the corresponding tests.

Srivastava (2007) and Srivastava and Du (2008) discussed one and two sample problem, but their procedures for testing equality of two mean vectors are based on the assumption of equal covariance structure. So, in order to adapt these procedures to our approach, we at first consider transformation (3.2), and adapted transformed $\boldsymbol{y}_j$ $(j = 1, \ldots, N_1)$ to their procedures for one sample problem.

Aoshima and Yata (2011) proposed the test procedure for the multivariate Behrens-Fisher problem with size α and power no less than $1 - \beta$ when $\Delta \geq \Delta_L$, where α, $\beta \in (0, 1/2)$ and $\Delta_L (> 0)$ are prespecified constant. As they assumed $\Delta_L = o(p^{1/2})$, we set up $\beta = 0.1$ and $\Delta_L = \Delta/2$. Also, procedure by Aoshima and Yata (2011) does not need the assumption of normality, but in order to compare it with other procedures we carry out the simulation study under normality.

Tables 5 to 8 show that our test statistic appear to be consistent as $(N, p) \to \infty$. Further, under simulated alternative hypothesis $H_1^{L(\delta)}$ with $\Delta = 5$ our statistic performs better than both S and SD, and is comparable to AY. For the alternative with $\Delta = 10$ our test turns out to be most powerful for almost all the settings of p ans N_i. Hence, as Δ increases the newly proposed test appears to dominate that of AY, and both are seen to be consistent as Δ grows. We also see that neither test perform particularly well for $\Delta = 5$, in combination with small N_i and large p.

Further, a series of simulations is provided to demonstrate the advantage of the correction procedure suggested in (3.4) for the critical point of the test. We simulate the ASL of our test using $\hat{z}(\alpha)$ from the approximation (3.4), and ASL of AY which uses $\Delta_L z_\alpha/(z_\alpha + z_\beta)$ as the critical point (see for details Aoshima and Yata (2011)). Both S and SD use z_α. Tables 9-12 give ASL for the following cases:

$$\begin{aligned}
\text{Table } 9 &: \Sigma_1 = I,\ \Sigma_2 = (0.5^{|i-j|})\ N_1 = 10, \text{ and } N_2 = 20,\\
\text{Table } 10 &: \Sigma_1 = I,\ \Sigma_2 = (0.5^{|i-j|})\ N_1 = 20, \text{ and } N_2 = 30,\\
\text{Table } 11 &: \Sigma_1 = (0.2^{|i-j|}),\ \Sigma_2 = (0.5^{|i-j|})\ N_1 = 10, \text{ and } N_2 = 20,\\
\text{Table } 12 &: \Sigma_1 = (0.2^{|i-j|}),\ \Sigma_2 = (0.5^{|i-j|})\ N_1 = 20, \text{ and } N_2 = 30.
\end{aligned}$$

Also, AY(2.5) and AY(5) denote AY with $\Delta_L = 2.5$ and $\Delta_L = 5$, respectively.

Tables 9 to 12 show that the our test procedure based on $\widehat{z}(\alpha)$ outperforms all other procedures for both types of alternatives, $\Delta = 5$ and $\Delta = 10$, and for all the settings of p and N_i, yielding the value of ASL closest to the nominal level α.

Lastly, we study the effect of p on the power of the test. From tables 5 to 12 one can see that for both $\Delta = 5$ and $\Delta = 10$ our test statistic appear to have most stable power when p grows, and this result remain valid for very small sizes.

5. Concluding remarks

We have proposed a new test statistic for testing linear hypothesis of mean vectors assuming that covariance matrices are unequal. We also suggested a method for correcting the critical point of the test that leads to better performance in large p and small N_i case. Simulations indicate that the newly derived test statistic, $\widetilde{T}_D$, in (2.4) appear to perform well for a range of settings of H_0 specified by β_i's and $k > 2$, when $\widehat{z}(\alpha)$ from Corollary 2.2 is used as a critical point.

For the multivariate Behrens-Fisher problem, our test procedure has a comparable power performance to that of Aoshima and Yata (2011), and outperforms both procedures derived in Srivastava (2007) and Srivastava and Du (2008) for all the high-dimensional settings of p and N_i and a number of settings of Σ_i's. It is especially important to point out that our procedure performs well for small deviations from H_0, i.e. when $\Delta = 5$.

In conclusion, our test can be recommended for testing the mean vectors for both $k > 2$ case and multivariate Behrens-Fisher problem, when p is much larger than N_i, and when a small deviation from H_0 is suspected.

Appendix A.

A.1. Proof of Theorem 2.1.

To derive the asymptotic expansion of the distribution of $\widetilde{T}_D/\widehat{\sigma}_D$ under null hypothesis we need estimators of a_1 and a_2. So we consider unbiased and (N,p)-consistent estimators of a_1 and a_2 derived in Srivastava (2005) (see, section 2). By $(N,p)-$consistency, we obtain

$$\frac{\widetilde{T}_D}{\widehat{\sigma}_D} = \frac{N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}S}{\sqrt{2p\widehat{a}_2}}.$$

Let $w = n_1(\widehat{a}_2 - a_2)$ which implies that $w = O_p(1)$ (see Srivastava (2005)) and $\widetilde{T}_D/\widehat{\sigma}_D$ can be expanded as

$$\begin{aligned}\frac{\widetilde{T}_D}{\widehat{\sigma}_D} &= \frac{1}{\sqrt{2p\widehat{a}_2}}\left(N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}S\right)\left(1 + \frac{1}{n_1 a_2}w\right)^{-1/2} \\ &= \frac{1}{\sqrt{2p\widehat{a}_2}}\left(N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}S\right)\left(1 - \frac{1}{2n_1 a_2}w + O_p(n_1^{-2})\right).\end{aligned}$$

To further explore the distribution of $\widetilde{T}_D/\widehat{\sigma}_D$, we derive the asymptotic expansion of the characteristic function. The following lemma is proved in Appendix B.

Lemma A.1. *Under the asymptotic framework* (A.1)*, the characteristic function of* $\widetilde{T}_D/\widehat{\sigma}_D$ *can be expressed as*

$$C(t) = \left|I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2}\left|I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2} E_{\boldsymbol{z}_1^*, Z_2^*}[g(S,\overline{\boldsymbol{y}})],$$

where

$$g(S,\overline{\boldsymbol{y}}) = 1 - \frac{(it)w}{2n_1 a_2\sqrt{2pa_2}}\left(N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}S\right) + O_p(n_1^{-2}).$$

We now proceed to show Theorem 2.1 by analyzing

$$\text{(i)}\ \left|I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2}\left|I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2}, \quad \text{(ii)}\ \mathrm{E}_{\boldsymbol{z}_1^*, Z_2^*}[g(S,\overline{\boldsymbol{y}})],$$

respectively. At first, for (i), the following equality are hold

$$
\begin{aligned}
&\log\left|I_p-\frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2}\log\left|I_p+\frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2}\\
&=\frac{1}{2}\left\{\frac{\sqrt{2}(it)}{\sqrt{pa_2}}\mathrm{tr}\Sigma+\frac{1}{2}\left(\frac{\sqrt{2}(it)}{\sqrt{pa_2}}\right)^2\mathrm{tr}\Sigma^2+\frac{1}{3}\left(\frac{\sqrt{2}(it)}{\sqrt{pa_2}}\right)^3\mathrm{tr}\Sigma^3\right.\\
&\quad\left.+\frac{1}{4}\left(\frac{\sqrt{2}(it)}{\sqrt{pa_2}}\right)^4\mathrm{tr}\Sigma^4+O(p^{-3/2})\right\}-\frac{n_1}{2}\left\{\frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\mathrm{tr}\Sigma\right.\\
&\quad\left.-\frac{(it)^2}{n_1^2pa_2}\mathrm{tr}\Sigma^2+O(p^{-7/2})\right\}\\
&=\frac{(it)^2}{2}+\frac{\sqrt{2}(it)^3a_3}{3\sqrt{pa_2^3}}+\frac{(it)^4a_4}{2pa_2^2}+\frac{(it)^2}{2n_1}+O(n_1^{-3/2}).
\end{aligned}
$$

Therefore (i) can be rewritten as

$$
\begin{aligned}
&\left|I_p-\frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2}\left|I_p+\frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2}\\
&=\exp\left\{\frac{(it)^2}{2}+\frac{\sqrt{2}(it)^3a_3}{3\sqrt{pa_2^3}}+\frac{(it)^4a_4}{2pa_2^2}+\frac{(it)^2}{2n_1}+O(n_1^{-3/2})\right\}\\
&=\exp\left\{\frac{(it)^2}{2}\right\}\left\{1+\frac{\sqrt{2}(it)^3a_3}{3\sqrt{pa_2^3}}+\frac{(it)^4a_4}{2pa_2^2}+\frac{(it)^6a_3^2}{9pa_2^3}+\frac{(it)^2}{2n_1}\right\}+O(n_1^{-3/2}).
\end{aligned}
\tag{1}
$$

Secondly, we expand $\mathrm{E}_{\boldsymbol{z}_1^*,Z_2^*}[g(S,\overline{\boldsymbol{y}})]$ in (ii) as follows:

$$
\begin{aligned}
&\mathrm{E}_{(\boldsymbol{z}_1^*,Z_2^*)}[g(S,\overline{\boldsymbol{y}})]\\
&=1-\frac{(it)}{2n_1a_2\sqrt{2pa_2}}\mathrm{E}[w(N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}}-\mathrm{tr}S)+O_p(n_1^{-2})]\\
&=1-\frac{(it)}{2n_1a_2\sqrt{2pa_2}}\mathrm{E}\Big[\Big\{\frac{n_1^2}{p(n_1+2)(n_1-1)}\Big(\frac{\mathrm{tr}(\Sigma Z_2Z_2'\Sigma Z_2Z_2')}{n_1^2}\\
&\quad-\frac{(\mathrm{tr}(\Sigma Z_2Z_2'))^2}{n_1^3}\Big)-\frac{\mathrm{tr}\Sigma^2}{p}\Big\}\Big(\mathrm{tr}(\Sigma\boldsymbol{z}_1\boldsymbol{z}_1')-\frac{\mathrm{tr}(\Sigma Z_2Z_2')}{n_1}\Big)+O_p(n_1^{-2})\Big]
\end{aligned}
$$

$$= 1 - \frac{(it)}{2n_1a_2\sqrt{2pa_2}}\mathrm{E}\Big[\Big\{\frac{1}{(n_1+2)(n_1-1)}\mathrm{tr}\Big(BZ_2^*Z_2^{*'}BZ_2^*Z_2^{*'}\Big)$$
$$-\frac{1}{n_1(n_1+2)(n_1-1)}\Big(\mathrm{tr}\Big(BZ_2^*Z_2^{*'}\Big)\Big)^2 - \mathrm{tr}\Sigma^2\Big\}\Big\{\mathrm{tr}\Big(A\boldsymbol{z}_1^*\boldsymbol{z}_1^{*'}\Big)$$
$$-\frac{1}{n_1}\mathrm{tr}\Big(BZ_2^*Z_2^{*'}\Big)\Big\}\Big] + O(n_1^{-2}), \tag{2}$$

where

$$A = \Sigma\left(I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right)^{-1}, \quad B = \Sigma\left(I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right)^{-1},$$

respectively. To calculate the expectation in (2), we need the following lemma.

Lemma A.2. *Let $\boldsymbol{z}_1^*$ and Z_2^* be mutually independently and distributed as $N_p(\mathbf{0}, I_p)$ and $N_{pn_1}(\mathbf{0}, I_p \otimes I_{n_1})$, respectively. Then the following expectations are calculated as*

$$\begin{aligned}
\mathrm{E}[\mathrm{tr}(BZ_2^*Z_2^{*'}BZ_2^*Z_2^{*'})]\mathrm{E}[\mathrm{tr}(A\boldsymbol{z}_1^*\boldsymbol{z}_1^{*'})] &= \mathrm{tr}A\{n_1(n_1+1)\mathrm{tr}B^2 + n_1(\mathrm{tr}A)^2\},\\
\mathrm{E}[\mathrm{tr}(BZ_2^*Z_2^{*'}BZ_2^*Z_2^{*'})\mathrm{tr}(BZ_2^*Z_2^{*'})] &= 4n_1(n_1+1)\mathrm{tr}B^3 + n_1(n_1^2+n_1+4)\\
&\qquad \times\mathrm{tr}B\mathrm{tr}B^2 + n_1^2(\mathrm{tr}B)^3,\\
\mathrm{E}[\{\mathrm{tr}(BZ_2^*Z_2^{*'})\}^2\mathrm{tr}(A\boldsymbol{z}_1^*\boldsymbol{z}_1^{*'})] &= \mathrm{tr}A\{2n_1\mathrm{tr}B^2 + n_1^2(\mathrm{tr}B)^2\},\\
\mathrm{E}[\{\mathrm{tr}(BZ_2^*Z_2^{*'})\}^3] &= 8n_1\mathrm{tr}B^3 + 6n_1^2\mathrm{tr}B\mathrm{tr}B^2 + n_1^3(\mathrm{tr}B)^3,\\
\mathrm{E}[\mathrm{tr}\Sigma^2\mathrm{tr}(A\boldsymbol{z}_1^*\boldsymbol{z}_1^{*'})] &= \mathrm{tr}\Sigma^2\mathrm{tr}A,\\
\mathrm{E}[\mathrm{tr}\Sigma^2\mathrm{tr}(BZ_2^*Z_2^{*'})] &= \mathrm{tr}\Sigma^2\mathrm{tr}B.
\end{aligned}$$

Proof. See, Himeno (2007).

Now, by applying Lemma A.2 to (2), we obtain:

$$\mathrm{E}_{(\boldsymbol{z}_1^*, Z_2^*)}[g(S, \overline{\boldsymbol{y}})] = 1 + O(n_1^{-3/2}).$$

Summarizing (1) and (2), we obtain the expansion of the characteristic function

$$C(t) = \exp\left\{\frac{(it)^2}{2}\right\} \times \Big[1 + \frac{\sqrt{2}(it)^3a_3}{3\sqrt{pa_2^3}} + \frac{(it)^4a_4}{2pa_2^2} + \frac{(it)^6a_3^2}{9pa_2^3} + \frac{(it)^2}{2n_1}\Big] + O(n_1^{-3/2}).$$

By inverting this characteristic function, we get the following density function of $\widetilde{T}_D/\widehat{\sigma}_D$;

$$\begin{aligned} f(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-itz\}C(t)dt \\ &= \phi(z)\Big[1+\frac{1}{\sqrt{p}}c_3h_3(z)+\frac{1}{p}\{c_4h_4(z)+c_6h_6(z)\}+\frac{1}{n_1}c_2h_2(z)\Big]+O(n_1^{-3/2}), \end{aligned}$$

where $\phi(z)$ is the density function of the standard normal distribution,

$$c_2 = \frac{1}{2}, \quad c_3 = \frac{\sqrt{2}a_3}{3\sqrt{a_2^3}}, \quad c_4 = \frac{a_4}{2a_2^2}, \quad c_6 = \frac{a_3^2}{9a_2^3},$$

and $h_i(z)$'s $(i = 1, \dots, 6)$ are the Hermite polynomials given by

$$\begin{aligned} &h_1(z) = z, \quad h_2(z) = z^2 - 1, \quad h_3(z) = z^3 - 3z, \quad h_4(z) = z^4 - 6z^2 + 3, \\ &h_5(z) = z^5 - 10z^3 + 15z, \quad h_6(z) = z^6 - 15z^4 + 45z^2 - 15. \end{aligned}$$

Therefore, we obtain Theorem 2.1.

A.2. Proof of Corollary 2.2.

We prepare following Lemma to derive the Cornish-Fisher expansion of the upper 100α percentiles of $\widetilde{T}_D/\widehat{\sigma}_D$.

Lemma A.3. *Let $z(\alpha)$ be*

$$z(\alpha) = z_\alpha + \frac{1}{\sqrt{p}}q_1(z_\alpha) + \frac{1}{p}q_2(z_\alpha) + \frac{1}{n_1}q_3(z_\alpha),$$

where z_α is the upper $100\alpha\%$ point of the standard normal distribution and

$$\begin{aligned} q_1(z_\alpha) &= \frac{\sqrt{2}a_3}{3\sqrt{a_2^3}}(z_\alpha^2 - 1), \\ q_2(z_\alpha) &= \frac{a_4}{2a_2^2}z_\alpha(z_\alpha^2 - 3) - \frac{2a_3^2}{9a_2^3}z_\alpha(2z_\alpha^2 - 5), \\ q_3(z_\alpha) &= \frac{z_\alpha}{2}. \end{aligned}$$

Then under the framework (A.1) *and assumption* (A.3),

$$P\left(\frac{\widetilde{T}_D}{\widehat{\sigma}_D} \le z(\alpha)\right) = 1 - \alpha + O(p^{-3/2}).$$

Proof. See, Appendix B.2.

Now, by replacing a_i's in Lemma A.3 with their unbiased and consistent estimators $\widehat{a}_i$'s for $i = 1, \ldots, 4$, we obtain Corollary 2.2.

A.3. Proof of Theorem 2.3.

We derive the limiting distribution of the statistic T_D^*/σ_D^* under (A.1), (A.2) and $H_1^{L(\delta)}$. T_D^*/σ_D^* is expanded as

$$\begin{aligned}\frac{T_D^*}{\sigma_D^*} &= \frac{(\mathrm{tr}\Sigma/\mathrm{tr}S)N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}\Sigma - N_1\boldsymbol{\mu}'\boldsymbol{\mu}}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}} \\ &= \frac{1}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}(\boldsymbol{z}_1'\Sigma\boldsymbol{z}_1 + 2\sqrt{N_1}\boldsymbol{\mu}'\Sigma^{1/2}\boldsymbol{z}_1 - \mathrm{tr}\Sigma) + o_p(1).\end{aligned}$$

Also, the characteristic function of T_D^*/σ_D^* is calculated as

$$\begin{aligned}C(t) &= \mathrm{E}_{\boldsymbol{z}_1}\left[\exp\left\{\frac{it(\boldsymbol{z}_1'\Sigma\boldsymbol{z}_1 + 2\sqrt{N_1}\boldsymbol{\mu}\Sigma^{1/2}\boldsymbol{z}_1 - \mathrm{tr}\Sigma)}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right\}\right] + o(1) \\ &= \mathrm{etr}\left\{\frac{-(it)\mathrm{tr}\Sigma}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right\}\int_{\boldsymbol{z}_1}(2\pi)^{-p/2} \times \mathrm{etr}\left\{-\frac{1}{2}\left(I_p\right.\right. \\ &\quad \left.\left.-\frac{2(it)\Sigma}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right)\boldsymbol{z}_1\boldsymbol{z}_1' + \frac{2(it)\sqrt{N_1}\boldsymbol{\mu}'\Sigma^{1/2}\boldsymbol{z}_1}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right\}d\boldsymbol{z}_1 + o(1).\end{aligned}$$

Further, we consider the following transformation

$$\begin{aligned}\boldsymbol{z}_1^* = &\left(I_p - \frac{2(it)\Sigma}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right)^{1/2}\boldsymbol{z}_1 \\ &-\left(I_p - \frac{2(it)\Sigma}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right)^{-1/2}\frac{2(it)}{\sqrt{2\mathrm{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\sqrt{N_1}\Sigma^{1/2}\boldsymbol{\mu},\end{aligned}$$

whose Jacobian is given by $\left|I_p - 2(it)\Sigma/\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}\right|^{-1/2}$. Also, under (A.2),

$$\log\left|I_p - \frac{2(it)\Sigma}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right|^{-1/2}$$
$$= \frac{(it)\text{tr}\Sigma}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}} + \frac{(it)^2\text{tr}\Sigma^2}{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}} + o(1).$$

Therefore,

$$\left|I_p - \frac{2(it)\Sigma}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right|^{-1/2}$$
$$= \exp\left(\frac{(it)\text{tr}\Sigma}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}} + \frac{(it)^2\text{tr}\Sigma^2}{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}\right) + o(1),$$

and then we obtain the expansion of the characteristic function

$$C(t) = \exp\left(-\frac{(it)\text{tr}\Sigma}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right)\exp\left(\frac{(it)\text{tr}\Sigma}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right.$$
$$\left.+\frac{(it)^2\text{tr}\Sigma^2}{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}\right)\exp\left(\frac{(it)^2 2N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}{\sqrt{2\text{tr}\Sigma^2 + 4N_1\boldsymbol{\mu}'\Sigma\boldsymbol{\mu}}}\right) + o(1)$$
$$= \exp\left\{\frac{(it)^2}{2}\right\} + o(1).$$

Therefore, $T_D^*/\sigma_D^* \xrightarrow{d} N(0,1)$.

A.4. Proof of Theorem 2.4.

Let

$$\Delta_T = \widetilde{T}_D - T_D^* = \sqrt{p}\frac{N_1\boldsymbol{\mu}'\boldsymbol{\mu}}{\text{tr}\Sigma},$$

then the power of $\widetilde{T}_D$ with significance level α can be expressed as:

$$Power_\alpha(\widetilde{T}_D, \delta) \ = \ P(T_D^* > \sigma_D z_\alpha - \Delta_T \mid H_1^{L(\delta)}).$$

Then, by the results of Theorem 2.3,

$$\lim_{p\to\infty} Power_\alpha(\widetilde{T}_D,\delta) = \lim_{p\to\infty} \Phi\left(\frac{\Delta_T - \sigma_D z_\alpha}{\sigma_D^*}\right).$$

By the assumption (A.2), $Power_\alpha(\widetilde{T}_D,\delta) \to 1$ when $1/2 < \delta < 1$, since $\Delta_T \to \infty$, and $Power_\alpha(\widetilde{T}_D,\delta) \to \alpha$ when $0 < \delta < 1/2$, since $\Delta_T \to 0$ and $\sigma_D^* \to \sigma_D$. When $\delta = 1/2$, we obtain

$$\lim_{p\to\infty} Power_\alpha(\widetilde{T}_D,\delta) = \Phi\left(\frac{\Delta_T}{\sigma_D} - z_\alpha\right) = \Phi\left(\frac{N_1\boldsymbol{\mu}'\boldsymbol{\mu}}{\sqrt{2\mathrm{tr}\Sigma^2}} - z_\alpha\right).$$

Appendix B.

B.1. Proof of Lemma A.1.

The characteristic function of $\widetilde{T}_D/\widehat{\sigma}_D$ is calculated as

$$C(t) \;=\; \mathrm{E}\left[\exp\left(\frac{(it)\widetilde{T}_D}{\widehat{\sigma}_D}\right)\right] = \mathrm{E}\left[\exp\left\{\frac{(it)}{\sqrt{2pa_2}}\left(N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}S\right)\right\} g(S,\overline{\boldsymbol{y}})\right],$$

where

$$g(S,\overline{\boldsymbol{y}}) = 1 - \frac{(it)w}{2n_1a_2\sqrt{2pa_2}}\left(N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} - \mathrm{tr}S\right) + O_p(n_1^{-2}).$$

Let $\boldsymbol{z}_1$ be a p-dimensional random vector distributed as $N_p(\mathbf{0}, I_p)$, and Z_2 be a $p\times n_1$ random matrix such that $\mathrm{vec}(Z_2)$ is distributed as $N_{pn_1}(\mathbf{0}, I_p\otimes I_{n_1})$. Then we note that

$$N_1\overline{\boldsymbol{y}}'\overline{\boldsymbol{y}} = \mathrm{tr}(\Sigma^{1/2}\boldsymbol{z}_1\boldsymbol{z}_1'\Sigma^{1/2}), \quad n_1S = \Sigma^{1/2}Z_2Z_2'\Sigma^{1/2},$$

and we can rewrite the characteristic function as

$$\begin{aligned} C(t) \;&=\; \mathrm{E}\left[\exp\left\{\frac{(it)}{\sqrt{2pa_2}}\left(\mathrm{tr}(\Sigma^{1/2}\boldsymbol{z}_1\boldsymbol{z}_1'\Sigma^{1/2}) - \frac{\mathrm{tr}(\Sigma^{1/2}Z_2Z_2'\Sigma^{1/2})}{n_1}\right)\right\} g(S,\overline{\boldsymbol{y}})\right] \\ &=\; \int\!\!\int (2\pi)^{-(n_1+1)p/2}\mathrm{etr}\left\{-\frac{1}{2}\left(I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right)\boldsymbol{z}_1\boldsymbol{z}_1'\right\} \\ &\qquad \times\mathrm{etr}\left\{-\frac{1}{2}\left(I_p + \frac{1}{n_1\sqrt{pa_2}}\Sigma\right)Z_2Z_2'\right\} g(S,\overline{\boldsymbol{y}})d\boldsymbol{z}_1dZ_2. \end{aligned}$$

Here, we consider following transformations

$$\boldsymbol{z}_1 = \left(I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right)^{-1/2} \boldsymbol{z}_1^*, \ Z_2 = \left(I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right)^{-n_1/2} Z_2^*,$$

respectively, and the Jacobians for these transformations are

$$\left|I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2}, \quad \left|I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2}.$$

Therefore the characteristic function can be written as

$$\begin{aligned} C(t) &= \left|I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2} \left|I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2} \\ &\quad \times \int\int (2\pi)^{-(n_1+1)p/2} \mathrm{etr}\left\{-\frac{1}{2}\boldsymbol{z}_1^*\boldsymbol{z}_1^{*'} - \frac{1}{2}Z_2^*Z_2^{*'}\right\} g(S,\overline{y}) d\boldsymbol{z}_1^* dZ_2^* \\ &= \left|I_p - \frac{\sqrt{2}(it)}{\sqrt{pa_2}}\Sigma\right|^{-1/2} \left|I_p + \frac{\sqrt{2}(it)}{n_1\sqrt{pa_2}}\Sigma\right|^{-n_1/2} \mathrm{E}_{\boldsymbol{z}_1^*, Z_2^*}[g(S, \overline{\boldsymbol{y}})]. \end{aligned}$$

B.2. Proof of Lemma A.3

Let $z(\alpha)$ be the upper 100α percentile of $\widetilde{T}_D/\widehat{\sigma}_D$. We further expand $z(\alpha)$ as

$$z(\alpha) = u + \frac{1}{\sqrt{p}}q_1(u) + \frac{1}{p}q_2(u) + \frac{1}{n_1}q_3(u).$$

Now, from the result of Theorem 2.1, we derive the following expansion

$$\begin{aligned} 1-\alpha &= P\left(\frac{\widetilde{T}_D}{\widehat{\sigma}_D} \leq z(\alpha)\right) \\ &= \Phi(z(\alpha)) - \phi(z(\alpha))\left[\frac{1}{\sqrt{p}}c_3h_2(z(\alpha)) \right. \\ &\quad \left. + \frac{1}{p}\left\{c_4h_3(z(\alpha)) + c_6h_5(z(\alpha))\right\} + \frac{1}{n_1}c_2h_1(z(\alpha))\right] + O(p^{-3/2}). \end{aligned}$$

Then, by Taylor expansion of Φ, ϕ and h_i's around u, we obtain

$$P\left(\frac{\widetilde{T}_D}{\widehat{\sigma}_D} \leq z(\alpha)\right) = \Phi(u) - \phi(u)\left[\frac{1}{\sqrt{p}}\{q_1(u) - c_3 h_2(u)\}\right.$$
$$\left. + \frac{1}{p}\{q_2(u) - c_4 h_3(u) - c_6 h_5(u)\} + \frac{1}{n_1}\{q_3(u) - c_2 h_1(u)\}\right] + O(p^{-3/2}).$$

Therefore, we have

$$\begin{aligned} q_1(u) &= \frac{\sqrt{2}a_3}{3\sqrt{a_2^3}}(u^2 - 1), \\ q_2(u) &= \frac{a_4}{2a_2^2}u(u^2 - 3) - \frac{2a_3^2}{9a_2^3}u(2u^2 - 5), \\ q_3(u) &= \frac{u}{2}. \end{aligned}$$

Acknowledgements

T. Nishiyama's research was in part supported by the travel grant from The Royal Swedish Academy of Sciences, G.S. Magnuson's fond. M. Hyodo's research was in part supported by Japan Society for the Promotion of Science (JSPS). T. Seo's research was in part supported by Grant-in-Aid for Scientific Research (C)(23500360). T. Pavlenko's research was in part supported by grant from Swedish Research Council (421-2008-1966).

References

[1] Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis Third Edition*. Wiley, New york.

[2] Aoshima, M. and Yata, K. (2011). Two stage procedures for high-dimensional data. *Sequential Analysis*, **30**, 356–399.

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

[4] Bennett, B. M. (1951). Note on a solution of the generalized Behrens-Fisher problem. *Annals of the Institute of Statistical Mathematics*, **2**, 87–90.

[5] 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.

[6] Dempster, A. P. (1958). A high dimensional two sample significance test. *Annals of Mathematical Statistics*, **29**, 995–1010.

[7] Dempster, A. P. (1960). A significant test for the separation of two highly multivariate small samples. *Biometrics*, **16**, 41–50.

[8] Himeno, T. (2007). Asymptotic expansions of the null distributions for the Dempster trace criterion. *Hiroshima Mathematical Journal*, **37**, 431–454.

[9] Hyodo, M., Takahashi, S. and Nishiyama, T. (2012). Multiple comparisons among mean vectors when the dimension is larger than the total sample size. *Technical Report No.12-01*, Statistical Research Group, Hiroshima University.

[10] Johnson, R. A. and Weerahendi, S. (1988). A Baysian solution to the multivariate Behrens-Fisher problem. *Journal of the American Statistical Association*, **83**, 145–149.

[11] Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high-dimensional data. *Journal of Japan Statistical Society*, **35**, 251–272.

[12] Srivastava, M. S. (2007). Multivariate theory for analyzing high dimensional data. *Journal of Japan Statistical Society*, **37**, 53–86.

[13] 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.

[14] Srivastava, M. S. and Yanagihara, H. (2010). Testing the equality of several covariance matrices with fewer observations than the dimension. *Journal of Multivariate Analysis*, **101**, 1319–1329.

[15] Yanagihara, H. and Yuan, K. H. (2005). Three approximate solutions to the multivariate Behrens-Fisher problem. *Communications in Statistics - Simulation and Computation*, **34**, 975–988.

Table 1: ASL in the case of $k = 3$ and $\beta_1 = \beta_2 = \beta_3 = 1$ $(i = 1, 2, 3)$.

			$N_i = 10(1+i)$			$N_i = 20(1+i)$		
p	α	z_α	t	ASL_α^1	ASL_α^2	t	ASL_α^1	ASL_α^2
	0.01	2.326	2.756	0.022	0.012	2.682	0.020	0.010
60	0.05	1.645	1.835	0.067	0.052	1.788	0.063	0.051
	0.1	1.282	1.374	0.114	0.102	1.341	0.109	0.101
	0.01	2.326	2.691	0.020	0.011	2.623	0.018	0.010
90	0.05	1.645	1.807	0.065	0.052	1.768	0.061	0.051
	0.1	1.282	1.366	0.113	0.102	1.331	0.108	0.100
	0.01	2.326	2.661	0.019	0.011	2.581	0.017	0.010
120	0.05	1.645	1.796	0.064	0.052	1.755	0.061	0.051
	0.1	1.282	1.362	0.113	0.102	1.334	0.108	0.101
	0.01	2.326	2.625	0.018	0.011	2.565	0.017	0.010
150	0.05	1.645	1.783	0.063	0.052	1.746	0.059	0.051
	0.1	1.282	1.355	0.111	0.102	1.326	0.107	0.100
	0.01	2.326	2.591	0.017	0.011	2.541	0.016	0.010
200	0.05	1.645	1.767	0.062	0.051	1.737	0.059	0.051
	0.1	1.282	1.350	0.111	0.101	1.326	0.107	0.101

Table 2: ASL in the case of $k = 3$, $\beta_1 = 1$ and $\beta_2 = \beta_3 = -1/2$ $(i = 1, 2, 3)$.

			$N_i = 10(1+i)$			$N_i = 20(1+i)$		
p	α	z_α	t	ASL_α^1	ASL_α^2	t	ASL_α^1	ASL_α^2
	0.01	2.326	2.728	0.021	0.011	2.661	0.019	0.011
60	0.05	1.645	1.824	0.066	0.052	1.782	0.062	0.051
	0.1	1.282	1.371	0.114	0.102	1.343	0.110	0.101
	0.01	2.326	2.658	0.019	0.011	2.612	0.018	0.011
90	0.05	1.645	1.790	0.063	0.051	1.762	0.061	0.051
	0.1	1.282	1.349	0.111	0.100	1.336	0.108	0.101
	0.01	2.326	2.630	0.018	0.011	2.583	0.017	0.011
120	0.05	1.645	1.787	0.064	0.052	1.743	0.059	0.050
	0.1	1.282	1.363	0.113	0.103	1.323	0.107	0.100
	0.01	2.326	2.602	0.017	0.011	2.542	0.016	0.010
150	0.05	1.645	1.775	0.063	0.052	1.740	0.059	0.051
	0.1	1.282	1.354	0.111	0.102	1.332	0.108	0.102
	0.01	2.326	2.583	0.017	0.011	2.505	0.015	0.010
200	0.05	1.645	1.757	0.061	0.051	1.725	0.058	0.050
	0.1	1.282	1.340	0.109	0.100	1.323	0.107	0.100

Table 3: ASL in the case of $k = 4$ and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$ $(i = 1, \ldots, 4)$.

			$N_i = 10(1+i)$			$N_i = 20(1+i)$		
p	α	z_α	t	ASL^1_α	ASL^2_α	t	ASL^1_α	ASL^2_α
	0.01	2.326	2.731	0.021	0.012	2.659	0.019	0.010
60	0.05	1.645	1.821	0.066	0.052	1.780	0.063	0.051
	0.1	1.282	1.371	0.113	0.102	1.339	0.109	0.101
	0.01	2.326	2.668	0.020	0.011	2.611	0.018	0.011
90	0.05	1.645	1.799	0.065	0.052	1.758	0.061	0.050
	0.1	1.282	1.363	0.113	0.102	1.331	0.108	0.100
	0.01	2.326	2.632	0.019	0.011	2.562	0.016	0.010
120	0.05	1.645	1.785	0.063	0.052	1.747	0.060	0.050
	0.1	1.282	1.355	0.112	0.101	1.330	0.108	0.101
	0.01	2.326	2.617	0.018	0.011	2.546	0.016	0.010
150	0.05	1.645	1.775	0.062	0.051	1.741	0.059	0.051
	0.1	1.282	1.352	0.111	0.102	1.324	0.107	0.100
	0.01	2.326	2.584	0.017	0.011	2.519	0.015	0.010
200	0.05	1.645	1.766	0.062	0.051	1.730	0.058	0.050
	0.1	1.282	1.346	0.110	0.101	1.323	0.107	0.100

Table 4: ASL in the case of $k = 4$, $\beta_1 = \beta_3 = 1$ and $\beta_2 = \beta_4 = -1$ $(i = 1, \ldots, 4)$.

			$N_i = 10(1+i)$			$N_i = 20(1+i)$		
p	α	z_α	t	ASL^1_α	ASL^2_α	t	ASL^1_α	ASL^2_α
	0.01	2.326	2.733	0.021	0.012	2.669	0.019	0.011
60	0.05	1.645	1.821	0.066	0.052	1.783	0.063	0.051
	0.1	1.282	1.368	0.113	0.102	1.342	0.110	0.101
	0.01	2.326	2.671	0.019	0.011	2.605	0.017	0.010
90	0.05	1.645	1.799	0.064	0.052	1.759	0.061	0.050
	0.1	1.282	1.362	0.113	0.102	1.333	0.108	0.101
	0.01	2.326	2.639	0.019	0.011	2.576	0.017	0.010
120	0.05	1.645	1.791	0.064	0.052	1.749	0.060	0.051
	0.1	1.282	1.360	0.112	0.102	1.329	0.108	0.100
	0.01	2.326	2.609	0.018	0.011	2.545	0.016	0.010
150	0.05	1.645	1.773	0.062	0.051	1.743	0.060	0.051
	0.1	1.282	1.350	0.111	0.101	1.331	0.108	0.101
	0.01	2.326	2.581	0.017	0.011	2.525	0.015	0.010
200	0.05	1.645	1.765	0.062	0.051	1.730	0.058	0.050
	0.1	1.282	1.348	0.111	0.101	1.322	0.107	0.100

Table 5: Empirical powers with $\Sigma_1 = I$, $\Sigma_2 = (0.5^{|i-j|})$ and $\Delta = 5$.

		$N_1 = 10,\ N_2 = 20$				$N_1 = 20,\ N_2 = 30$			
p	α	S	SD	AY	NHSP	S	SD	AY	NHSP
50	0.01	0.430	0.426	0.842	0.586	0.932	0.880	0.965	0.919
	0.05	0.551	0.623	0.863	0.769	0.965	0.952	0.975	0.973
	0.1	0.621	0.728	0.877	0.847	0.976	0.974	0.981	0.987
100	0.01	0.201	0.106	0.729	0.391	0.628	0.527	0.931	0.783
	0.05	0.307	0.280	0.751	0.606	0.751	0.761	0.946	0.910
	0.1	0.380	0.422	0.769	0.718	0.811	0.859	0.955	0.950
150	0.01	0.131	0.025	0.628	0.287	0.407	0.231	0.894	0.659
	0.05	0.217	0.116	0.652	0.497	0.559	0.517	0.912	0.839
	0.1	0.281	0.231	0.668	0.621	0.642	0.686	0.924	0.904
200	0.01	0.100	0.006	0.552	0.231	0.286	0.088	0.857	0.561
	0.05	0.175	0.048	0.573	0.433	0.429	0.314	0.877	0.772
	0.1	0.233	0.126	0.589	0.561	0.519	0.507	0.890	0.856

Table 6: Empirical powers with $\Sigma_1 = I$, $\Sigma_2 = (0.5^{|i-j|})$ and $\Delta = 10$.

		$N_1 = 10,\ N_2 = 20$				$N_1 = 20,\ N_2 = 30$			
p	α	S	SD	AY	NHSP	S	SD	AY	NHSP
50	0.01	0.700	0.883	0.971	0.953	0.996	0.999	0.998	1.000
	0.05	0.793	0.956	0.980	0.986	0.999	1.000	0.999	1.000
	0.1	0.837	0.977	0.985	0.994	0.999	1.000	1.000	1.000
100	0.01	0.384	0.485	0.927	0.848	0.905	0.977	0.996	0.997
	0.05	0.509	0.735	0.942	0.942	0.950	0.995	0.998	1.000
	0.1	0.582	0.845	0.953	0.969	0.967	0.999	0.999	1.000
150	0.01	0.251	0.186	0.873	0.743	0.733	0.860	0.991	0.988
	0.05	0.363	0.459	0.895	0.887	0.837	0.967	0.995	0.998
	0.1	0.438	0.641	0.909	0.936	0.884	0.989	0.996	0.999
200	0.01	0.182	0.058	0.811	0.650	0.578	0.639	0.984	0.971
	0.05	0.284	0.250	0.836	0.828	0.715	0.888	0.989	0.993
	0.1	0.355	0.444	0.854	0.896	0.783	0.956	0.993	0.997

Table 7: Empirical powers with $\Sigma_1 = (0.2^{|i-j|})$, $\Sigma_2 = (0.5^{|i-j|})$ and $\Delta = 5$.

		$N_1 = 10,\ N_2 = 20$				$N_1 = 20,\ N_2 = 30$			
p	α	S	SD	AY	NHSP	S	SD	AY	NHSP
	0.01	0.425	0.418	0.817	0.533	0.913	0.853	0.954	0.882
50	0.05	0.546	0.603	0.838	0.721	0.953	0.935	0.965	0.957
	0.1	0.618	0.704	0.854	0.810	0.968	0.963	0.972	0.978
	0.01	0.207	0.116	0.709	0.354	0.612	0.509	0.915	0.729
100	0.05	0.311	0.285	0.731	0.562	0.738	0.737	0.931	0.878
	0.1	0.382	0.420	0.746	0.677	0.801	0.837	0.941	0.928
	0.01	0.134	0.030	0.619	0.261	0.403	0.238	0.878	0.605
150	0.05	0.221	0.127	0.641	0.462	0.551	0.504	0.896	0.801
	0.1	0.284	0.239	0.656	0.587	0.635	0.667	0.909	0.876
	0.01	0.102	0.008	0.540	0.208	0.282	0.094	0.841	0.510
200	0.05	0.177	0.053	0.560	0.398	0.426	0.315	0.861	0.731
	0.1	0.236	0.132	0.575	0.522	0.516	0.500	0.875	0.824

Table 8: Empirical powers with $\Sigma_1 = (0.2^{|i-j|})$, $\Sigma_2 = (0.5^{|i-j|})$ and $\Delta = 10$.

		$N_1 = 10,\ N_2 = 20$				$N_1 = 20,\ N_2 = 30$			
p	α	S	SD	AY	NHSP	S	SD	AY	NHSP
	0.01	0.705	0.861	0.960	0.925	0.995	0.999	0.997	0.999
50	0.05	0.796	0.940	0.971	0.974	0.998	1.000	0.998	1.000
	0.1	0.841	0.966	0.977	0.988	0.999	1.000	0.999	1.000
	0.01	0.393	0.475	0.908	0.799	0.901	0.965	0.993	0.993
100	0.05	0.518	0.715	0.926	0.915	0.948	0.992	0.996	0.999
	0.1	0.592	0.825	0.938	0.953	0.966	0.997	0.997	0.999
	0.01	0.262	0.195	0.851	0.686	0.737	0.837	0.986	0.976
150	0.05	0.376	0.455	0.874	0.848	0.839	0.955	0.991	0.995
	0.1	0.452	0.628	0.889	0.910	0.884	0.982	0.994	0.998
	0.01	0.189	0.066	0.795	0.596	0.583	0.621	0.977	0.951
200	0.05	0.293	0.260	0.820	0.789	0.720	0.869	0.984	0.987
	0.1	0.364	0.444	0.838	0.867	0.786	0.943	0.988	0.994

Table 9: ASL in the case of $\Sigma_1 = I$, $\Sigma_2 = (0.5^{|i-j|})$.

		$N_1 = 10,\ N_2 = 20$				
p	α	S	SD	AY(2.5)	AY(5)	NHSP
	0.01	0.087	0.014	0.085	0.015	0.018
50	0.05	0.152	0.045	0.103	0.024	0.059
	0.1	0.203	0.078	0.119	0.034	0.108
	0.01	0.052	0.003	0.099	0.029	0.016
100	0.05	0.101	0.016	0.114	0.040	0.057
	0.1	0.142	0.037	0.126	0.052	0.105
	0.01	0.042	0.001	0.097	0.035	0.016
150	0.05	0.087	0.006	0.109	0.047	0.057
	0.1	0.126	0.019	0.118	0.057	0.105
	0.01	0.041	0.000	0.090	0.036	0.015
200	0.05	0.084	0.002	0.100	0.047	0.055
	0.1	0.121	0.010	0.108	0.056	0.103

Table 10: ASL in the case of $\Sigma_1 = I$, $\Sigma_2 = (0.5^{|i-j|})$.

		$N_1 = 20,\ N_2 = 30$				
p	α	S	SD	AY(2.5)	AY(5)	NHSP
	0.01	0.298	0.009	0.032	0.001	0.011
50	0.05	0.419	0.037	0.047	0.002	0.052
	0.1	0.497	0.073	0.062	0.004	0.102
	0.01	0.097	0.002	0.064	0.005	0.012
100	0.05	0.181	0.014	0.083	0.010	0.052
	0.1	0.246	0.037	0.102	0.017	0.101
	0.01	0.062	0.000	0.085	0.011	0.012
150	0.05	0.128	0.005	0.105	0.020	0.053
	0.1	0.185	0.020	0.122	0.030	0.103
	0.01	0.048	0.000	0.095	0.017	0.011
200	0.05	0.106	0.001	0.114	0.028	0.052
	0.1	0.158	0.011	0.130	0.039	0.102

Table 11: ASL in the case of $\Sigma_1 = (0.2^{|i-j|})$, $\Sigma_2 = (0.5^{|i-j|})$.

		$N_1 = 10,\ N_2 = 20$				
p	α	S	SD	AY(2.5)	AY(5)	NHSP
	0.01	0.083	0.016	0.093	0.020	0.017
50	0.05	0.144	0.047	0.111	0.030	0.058
	0.1	0.192	0.082	0.127	0.041	0.105
	0.01	0.050	0.004	0.111	0.036	0.016
100	0.05	0.099	0.019	0.126	0.049	0.056
	0.1	0.142	0.043	0.138	0.061	0.106
	0.01	0.043	0.001	0.110	0.044	0.017
150	0.05	0.086	0.007	0.123	0.057	0.058
	0.1	0.124	0.022	0.132	0.068	0.106
	0.01	0.039	0.000	0.101	0.045	0.015
200	0.05	0.081	0.003	0.112	0.056	0.055
	0.1	0.117	0.012	0.120	0.066	0.103

Table 12: ASL in the case of $\Sigma_1 = (0.2^{|i-j|})$, $\Sigma_2 = (0.5^{|i-j|})$.

		$N_1 = 20,\ N_2 = 30$				
p	α	S	SD	AY(2.5)	AY(5)	NHSP
	0.01	0.265	0.012	0.039	0.002	0.012
50	0.05	0.381	0.042	0.057	0.004	0.053
	0.1	0.456	0.079	0.073	0.007	0.103
	0.01	0.092	0.003	0.075	0.008	0.011
100	0.05	0.170	0.017	0.096	0.015	0.052
	0.1	0.233	0.043	0.114	0.023	0.102
	0.01	0.058	0.001	0.095	0.016	0.012
150	0.05	0.123	0.008	0.116	0.026	0.052
	0.1	0.178	0.024	0.133	0.037	0.101
	0.01	0.046	0.000	0.107	0.024	0.012
200	0.05	0.102	0.003	0.127	0.036	0.051
	0.1	0.152	0.014	0.143	0.048	0.102