# Testing independence in high-dimensional data: $\rho V$-coefficient based approach

Masashi Hyodo[†], Takahiro Nishiyama[††], Tatjana Pavlenko[†††]

[†] *Department of Mathematical Sciences, Graduate School of Engineering, Osaka Prefecture University,*
*1-1, Gakuen-cho, Naka-ku, Sakai-shi, Osaka 599-8531, Japan.*
*E-Mail:* `hyodo@ms.osakafu-u.ac.jp`
[††] *Department of Business Administration, Senshu University, 2–1–1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214–8580, Japan.*
*E-Mail:* `nishiyama@isc.senshu-u.ac.jp`
[†††] *Department of Mathematics, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden.*

**Abstract**

We treat the problem of testing mutual independence of $k$ high-dimensional random vectors when the data are multivariate normal and $k \geq 2$ is a fixed integer. For this purpose, we focus on the the vector correlation coefficient, $\rho V$ and propose an extension of its classical estimator which is constructed to correct potential sources of inconsistency related to the high dimensionality. Building on the proposed estimator of $\rho V$, we derive the new test statistic and study its limiting behavior in a general high-dimensional asymptotic framework which allows the vector's dimensionality arbitrarily exceed the sample size. Specifically, we show that the asymptotic distribution of the test statistic under the main hypothesis of independence is standard normal and that the proposed test is size and power consistent. Using our statistics, we further construct the step-down multiple comparison procedure based the closed testing strategy for the simultaneous test for independence. Accuracy of the proposed tests in finite samples is shown through simulations for a variety of high-dimensional scenarios in combination with a number of alternative dependence structures. Real data analysis is performed to illustrate the utility of the test procedures.

*Key words:*
RV-coefficient, Testing hypotheses, Multiple comparison procedure

## 1. Introduction

Testing independence of random variables is a standard task of statistical inference which naturally arises whenever it is needed to handle the dependence structures in multivariate data. Test of independence based on the product-moment correlation was initially explored in the classical seminar paper by Karl Pearson [13], followed by a substantial amount of research literature regarding this topic and its variants. One specific problem which emerges in contemporary applications is the test of independence of $k$, $p$-dimensional random vectors, where $k \geq 2$ is an integer representing the number of underlying populations. In this study, we address this issue and propose the test of significance based on the high-dimensional extension of the $\rho V$ vector correlation, initially introduced by Escoufier in [3] for characterizing the relationship of random vectors with a scalar measure of multivariate dependence. Based on the extended estimator of $\rho V$ and its asymptotic theory, we further develop two types of tests of independence of $k$ random vectors in arbitrarily high dimensions, and show that both tests apply whether $p \geq n$ or $p < n$ settings, where $n$ denotes the sample size.

### *1.1. Background and motivation*

Extensive overview of the classical, large $n$ and fixed $p$ independence testing techniques is provided in the textbooks on multivariate statistical analysis, see e.g., Mardia et al [12], Anderson [12], Fang and Zhang [12], and references there in. But, due to ever growing need of analyzing high- and ultra-high dimensional data, examples of applied areas include signal processing, astronomy, functional genomics and proteomics, just a few to name, the development of high-dimensional extensions of the classical testing procedures is of crucial importance. For instance, in functional genomics, multiple and high-dimensional data sets are frequently generated on the same samples of the biological system. This naturally calls for data fusion techniques which make it possible to extract the mutual information from all datasets simultaneously. The first step of the fusion strategy is to accurately identify whether certain similarities of the configuration of the samples (i.e., dependencies) occur between the datasets. Thus, it is necessary to develop novel testing methodologies suitable for testing the independence between such pairs of high-dimensional data sets. Another example motivating the research of this paper is discussed by Efron [4], who analyzed effects of the independence assumption for Cardio microarrays data comprising $n = 63$ arrays and $p = 20426$ genes. Starting with the presumption of independence across microarrays, which underlies most of conventional statistical inferential procedures, Efron demonstrated that the presence of dependence can invalidate the usual choice of a null hypothesis, leading to flawed assessments of significance. Hence, before conducting further high-dimensional statistical analyses such as classification, testing hypothesis of equality of mean vectors and covariance matrices, it is important to know when independence fails. For this purpose, testing procedures that are able to cope with nowadays $p \gg n$ data must be designed.

Our focus in this paper is on testing mutual independence of multivariate components building on the high-dimensional extension of $\rho V$. As for the review of the existing literature on the subject of our study, we refer to Josse et al. [10] who considered $\rho V$-based independence testing and argued for the permutation test strategy to approximate the distribution of the test statistic.

Further relevant approaches include Jiang et al. [9] who employed a high-dimensional correction of LRT to construct the test of independence of two vectors. However, the asymptotic theory of these corrected LRT statistic, such as its distribution under $\mathcal{H}$, is restricted to a bounded limiting ratio for both sub-vectors, i.e., to the high-dimensional case where $p_k/n \to c_k \in (0,1]$, $k = 1, 2$. Testing the independence of two normal sub-vectors based on the structure of the covariance matrix was considered by Srivastava and Reid [15], and further generalized by Hyodo et al [8] to testing the independence of $k$ sub-vectors. Yang and Pan [18] presented the independence test based on the sum of regularized sample canonical correlation coefficients. Testing of independence that does not require normality and is based on the distance correlation are presented by Székely and Rizzo [15] and [16]. Non-parametric approaches to the problem of testing independence can be found in e.g., a Han and Liu [6] who treated the maxima of rank correlations measure, such as Kendall's tau, and Leung and Drton [11] who used the framework of $U$-statistics and propose a family of test statistics which is based on sum of squares of sample rank correlations such as Kendall's tau, Hoeffding's $D$ statistics and a dominating term of Spearman's $\rho$.

### *1.2. Preliminaries and notations*

In what follows, we focus on the more precise problem statement, after some prefatory notations are in place. Henceforth, for an integer $k \geq 2$, we will denote by $[k]$ the set $\{1, \ldots, k\}$. Let $\mathbf{x} = (\mathbf{x}_1^\top, \ldots, \mathbf{x}_k^\top)^\top$ denote a $(p \times k)$-dimensional random vector, in which $\mathbf{x}_g$ possesses a dimension $p$ for each $g \in [k]$. Denote further by $\boldsymbol{\mu}_g$, by $\boldsymbol{\Sigma}_{gg}$, and by $\boldsymbol{\Sigma}_{gh}$, the mean vector of the $g$th sub-vector of $\mathbf{x}$, the covariance matrix of the $g$th sub-vector of $\mathbf{x}$, and the cross-covariance matrix of $\mathbf{x}_g$ and $\mathbf{x}_h$, respectively, for $g \neq h \in [k]$. Then $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \ldots, \boldsymbol{\mu}_k^\top)^\top$ and $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{gh})$, $g, h \in [k]$ are the mean vector and covariance matrix of $\mathbf{x}$, respectively. We are interested in testing the following hypothesis

$$\mathcal{H}: \quad \forall g, h \in [k] \quad \mathbf{x}_g \text{ and } \mathbf{x}_h \text{ are independent} \quad \text{vs.} \quad \mathcal{A}: \neg\mathcal{H}. \tag{1}$$

To this end, we draw a sample of independent observations of $\mathbf{x}$ using the following sampling scheme. Without loss of generality, we first assume that $1 \leq n_1 \leq \cdots \leq n_k$ and set $n_0 = 0$. Further, $\forall h \in [k]$ and $\forall j \in \{n_{h-1}+1, \ldots, n_h\}$, we denote $p(k-h+1)$ dimensional vectors by $\mathbf{x}_{\langle h \rangle j} = (\mathbf{x}_{hj}^\top, \ldots, \mathbf{x}_{kj}^\top)^\top$ and $\boldsymbol{\mu}_{\langle h \rangle} = (\boldsymbol{\mu}_h^\top, \ldots, \boldsymbol{\mu}_k^\top)^\top$, respectively. By considering a partition of $\boldsymbol{\Sigma}$ which is compatible with $\mathbf{x}_{\langle h \rangle}$ and $\boldsymbol{\mu}_{\langle h \rangle}$, we introduce a (positive definite) matrix

$$\boldsymbol{\Sigma}_{\langle h \rangle} = \begin{pmatrix} \boldsymbol{\Sigma}_{hh} & \cdots & \boldsymbol{\Sigma}_{hk} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{hk} & \cdots & \boldsymbol{\Sigma}_{kk} \end{pmatrix},$$

and assume that $\mathbf{x}_{\langle h \rangle j} \overset{i.i.d.}{\sim} \mathcal{N}_{p(k-h+1)}(\boldsymbol{\mu}_{\langle h \rangle}, \boldsymbol{\Sigma}_{\langle h \rangle})$. In addition,

$$\mathbf{x}_{\langle 1 \rangle 1}, \ldots, \mathbf{x}_{\langle 1 \rangle n_1}, \ldots, \mathbf{x}_{\langle k \rangle n_{k-1}+1}, \ldots, \mathbf{x}_{\langle k \rangle n_k}$$

are assumed to be mutually independent across $k$ populations, constituting thereby a sample of independent observations of $\mathbf{x}$ to be used for constructing the test procedure.

Observe that under the null hypothesis $\mathcal{H}$ of (1) stated in the multivariate normal setting, the population covariance matrix $\mathbf{\Sigma}_{\langle h \rangle}$ has all the cross-covariance components $\mathbf{\Sigma}_{gh} = \mathbf{O}$ which explicitly represents the classical inferential assumption of independence among $k$ populations. To enhance the presentation, we use the notation

$$\mathbf{D}_{\langle h \rangle} = \operatorname{diag}(\mathbf{\Sigma}_{hh}, \ldots, \mathbf{\Sigma}_{kk}) = \begin{pmatrix} \mathbf{\Sigma}_{hh} & \cdots & \mathbf{O} \\ \vdots & \ddots & \vdots \\ \mathbf{O} & \cdots & \mathbf{\Sigma}_{kk} \end{pmatrix},$$

to denote the diagonal block matrix with blocks $\mathbf{\Sigma}_{hh}, \ldots, \mathbf{\Sigma}_{kk}$, i.e., $^{\forall} h \in [k]$ all off-diagonal blocks of $\mathbf{D}_{\langle h \rangle}$ are $\mathbf{O}$. Here, $\mathbf{O}_{p \times p}$ will be used to denote the $p \times p$ null matrix and will be abbreviated to $\mathbf{O}$ when the dimensionality will be clear from the context. With the aid of these notations, the test of independence (1) can equivalently stated as

$$\mathcal{H} : \mathbf{\Sigma}_{\langle 1 \rangle} = \mathbf{D}_{\langle 1 \rangle} \quad \text{vs.} \quad \mathcal{A} : \neg \mathcal{H}. \tag{2}$$

The natural approach is to design a test statistic that measures the dependence among the components of $\mathbf{x}$ based on the sample, and reject $\mathcal{H}$ when its value is too large, where the critical value of rejection is set according to the asymptotic distribution of the test statistic under the null. Our focus in this paper is on the use of $\rho V$ vector correlations in settings where the dimension $p$ may exceed the sample size $n_i$. The new statistic we propose for testing $\mathcal{H}$ is constructed as a function of consistent estimators of the pairwise vector correlation coefficients and the corresponding asymptotic theory is developed to obtain the limit null distribution of this statistic with $p, n_i \to \infty$. The test statistic is presented in the next section, beginning with the high-dimensional adjustment of the estimator of the vector correlation coefficient, followed by the characterization of the test's asymptotic behavior. A simultaneous test of independence is further constructed in Section 3, where the proposed statistic is incorporated into the step-down multiple comparison algorithm. A finite sample performance of the proposed tests is shown in Section 4 through a number of simulation scenarios and application. Section 5 summarizes the main results. More technical details and proofs are gathered in the appendix.

Throughout the paper, $\operatorname{tr}(\mathbf{M})$ and $\|\mathbf{M}\|_F^2 = \operatorname{tr}(\mathbf{M}\mathbf{M}^\top)$ represent the trace of a square matrix $\mathbf{M}$ and its squared Frobenious norm, respectively. The symbol $\rightsquigarrow$ denotes convergence in distribution. The symbol $\otimes$ denotes Kronecker product.

## 2. Methodology and theory

Our proposed testing procedures will be studied under the high-dimensional or, as is frequently known as, $(n, p)$-asymptotic regime which is kept general implying that both $n_i \to \infty$ and $p \to \infty$, but without requiring the two indices to satisfy any specific relationship of mutual growth order, i.e. $p$ may arbitrarily exceed $n_i$. With the sampling scheme derived in Section 1.2, without loss of generality, we focus on $n_1$ and denote the high-dimensional asymptotic regime by $n_1, p \to \infty$ throughout the paper.

*2.1. Vector correlation coefficient in high-dimensional setting*

For any indices $g \neq h \in [k]$, let $\rho V_{gh}$ denote the vector correlation coefficient between the two components of $\mathbf{x}$, $\mathbf{x}_g$ and $\mathbf{x}_h$, defined as (see [3])

$$\rho V_{gh} = \frac{\|\mathbf{\Sigma}_{gh}\|_F^2}{\|\mathbf{\Sigma}_{gg}\|_F \|\mathbf{\Sigma}_{hh}\|_F}.$$

It is immediately clear that Pearson's product-moment correlation coefficient is a special case $\rho V_{gh}$ when $p = 1$. Furthermore, $\rho V_{gh} = \rho V_{hg}$, and $\rho V_{gh} = 0$ if and only if $\mathbf{\Sigma}_{gh} = \mathbf{O}$. In a view of this, if the joint distribution of $\mathbf{x}_g$ and $\mathbf{x}_h$ is normal, independence between $\mathbf{x}_g$ and $\mathbf{x}_h$ is equivalent to asserting that the population vector correlations all vanish, i.e., $\forall g \neq h \in [k]$, $\rho V_{gh} = 0$. Thus, the summation of these measurements over all $(g, h)$ pairs, subject to $g < h$, serves as an effective population measure of the overall dependency among $k$ parts of $\mathbf{x}$ and the natural criteria for testing $\mathcal{H}$ should be based on a suitable statistic for $\sum_{g<h}^{k} \rho V_{gh}$.

The sample counterpart of $\rho V_{hg}$ can be obtained as

$$RV_{gh} = \frac{\|\mathbf{S}_{gh}\|_F^2}{\|\mathbf{S}_{gg}\|_F \|\mathbf{S}_{hh}\|_F},$$

where the sample covariance matrix of $\mathbf{x}_\ell$ and the cross sample covariance matrix of $\mathbf{x}_g$ and $\mathbf{x}_h$ are constructed as

$$\forall \ell \in \{g, h\} \quad \mathbf{S}_{\ell\ell} = \frac{1}{n_\ell - 1} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \widetilde{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \widetilde{\mathbf{x}}_\ell)^\top,$$

$$\mathbf{S}_{gh} = \frac{1}{n_g - 1} \sum_{j=1}^{n_g} (\mathbf{x}_{gj} - \overline{\mathbf{x}}_g)(\mathbf{x}_{hj} - \overline{\mathbf{x}}_h)^\top, \ \mathbf{S}_{hg} = \mathbf{S}_{gh}^\top.$$

Here, $\overline{\mathbf{x}}_\ell = n_g^{-1} \sum_{j=1}^{n_g} \mathbf{x}_{\ell j}$, $\widetilde{\mathbf{x}}_\ell = n_\ell^{-1} \sum_{j=1}^{n_\ell} \mathbf{x}_{\ell j}$ for $\ell \in \{g, h\}$.

Note that the empirical measure of the vector correlation, $RV_{gh}$, is invariant by location, rotation, and overall scaling, and consistent for the classical case of the sample size $n$ tending to infinity and the dimension $p$ remaining fixed. The invariance property of $RV_{gh}$ is of special advantage because it allows to discuss the asymptotic behavior of the test statistic constructed from $RV_{gh}$ without knowing explicit information of the population mean vector and covariance matrix. However, as the "naive plug-in" estimator of the dependency measure, $RV_{gh}$ is flawed when $p > n_1$ and needs to be modified. Specifically $RV_{gh}$ lacks consistency under $\mathcal{H}$ when $p$ tends to infinity along with $n_1$ which is justified by the following lemma proved in Appendix section A.

**Lemma 1.** *Let $RV_{gh}$ be as already defined. Then, for any indices $g \neq h \in [k]$, the following representation holds*

$$RV_{gh} = \left( \rho V_{gh} + \frac{\mathrm{tr}(\mathbf{\Sigma}_{gg})\mathrm{tr}(\mathbf{\Sigma}_{hh})}{n_g \|\mathbf{\Sigma}_{gg}\|_F \|\mathbf{\Sigma}_{hh}\|_F} \right) \prod_{\ell \in \{g,h\}} \left( 1 + \frac{\{\mathrm{tr}(\mathbf{\Sigma}_{\ell\ell})\}^2}{n_\ell \|\mathbf{\Sigma}_{\ell\ell}\|_F^2} \right)^{-1/2} + o_p(1) \tag{3}$$

*as $p \to \infty$.*

To realize the essence of Lemma 1, observe that for $\mathbf{\Sigma}_{\ell\ell} = \mathbf{I}_p$ and $n_\ell = o(p)$, $R_{gh} = 1 + o_p(1)$ as $p \to \infty$, indicating that the $RV_{gh}$ coefficient is not able to capture the correlation.

By these arguments, the crucial step in our construction of test statistic for testing (1)-(2) is to obtain an estimator of $\rho V_{gh}$ suitable for high-dimensional settings. We first consider the following unbiased estimators of $\|\mathbf{\Sigma}_{gh}\|_F^2$ and $\|\mathbf{\Sigma}_{\ell\ell}\|_F^2$, (see Srivastava and Reid [15])

$$\forall g < h \in [k], \quad \widehat{\|\mathbf{\Sigma}_{gh}\|_F^2} = \frac{(n_g - 1)^2}{(n_g - 2)(n_g + 1)} \left\{ \|\mathbf{S}_{gh}\|_F^2 - \frac{\mathrm{tr}(\mathbf{S}_{gg})\mathrm{tr}(\mathbf{S}_{hh})}{n_g - 1} \right\},$$

$$\forall \ell \in [k], \quad \widehat{\|\mathbf{\Sigma}_{\ell\ell}\|_F^2} = \frac{(n_\ell - 1)^2}{(n_\ell - 2)(n_\ell + 1)} \left[ \|\mathbf{S}_{\ell\ell}\|_F^2 - \frac{\{\mathrm{tr}(\mathbf{S}_{\ell\ell})\}^2}{n_\ell - 1} \right].$$

and then define the estimator of $\rho V_{gh}$ with the high dimensionality adjustment as

$$HRV_{gh} = \frac{\widehat{\|\mathbf{\Sigma}_{gh}\|_F^2}}{\widehat{\|\mathbf{\Sigma}_{gg}\|_F} \widehat{\|\mathbf{\Sigma}_{hh}\|_F}}. \tag{4}$$

Apparently, the adjustment proposed in (4) preserves the invariance of $HRV_{gh}$ by location, rotation, and overall scaling. Now, to proceed further with the test statistic construction, we need one more result. The following theorem, proved in Appendix section B, shows consistency of $HRV_{gh}$ in high-dimensional regime.

**Theorem 1.** *Let $HRV_{gh}$ be as already defined. Then, as $n_1, p \to \infty$, for any indices $g \neq h \in [k]$, it holds that $HRV_{gh} = \rho V_{gh} + o_p(1)$.*

**Remark 1.** *Theorem 1 remains valid with $p$ fixed and $n_1 \to \infty$.*

### *2.2. The proposed test statistic and its asymptotic properties*

In order to construct the test statistic, we first observe that the tests (1) and (2) can be restated in terms of $\rho V$ coefficient as

$$\mathcal{H} : \forall g \neq h \in [k] \quad \rho V_{gh} = 0 \quad \text{vs.} \quad \mathcal{A} : \rho V_{gh} > 0. \tag{5}$$

Further, with the high-dimensional adjustment $HRV_{gh}$ at hand, we propose our test statistic for (1), (2) and (5), namely the vector correlation type statistics,

$$T = \sum_{1 \leq g < h \leq k} HRV_{gh},$$

which consistently estimates the population measure of the overall dependency, $\sum_{1\le g<h\le k}\rho V_{gh}$ in the joint distribution of $\mathbf{x}_1,\ldots,\mathbf{x}_k$, and sums all pairwise sample correlations for a "one-sided" test. Note that under the null hypothesis, all of the population's $\rho V_{gh}$ should be zero corresponding to zero off-diagonal blocks $\mathbf{\Sigma}_{gh}$. Hence, as an immediate consequence of Theorem 1, the asymptotic behavior of proposed test statistic under the null $\mathcal{H}$ and alternative $\mathcal{A}$ when $n_1, p\to\infty$ is as follows

$$T=\begin{cases} o_p(1) & \text{under } \ \mathcal{H},\\ \sum_{1\le g<h\le k}\rho V_{gh}+o_p(1) & \text{under } \ \mathcal{A},\end{cases}$$

i.e. the large values of $T$ indicate departures from $\mathcal{H}$.

To state the size-$\alpha$ test of significance using $T$, we need to characterize its null asymptotic distribution. One of the main results of our study is the central limit theorem for $T$ under $\mathcal{H}$, provided below. Let $\forall g\in[k]$, $\|\mathbf{\Sigma}_{gg}\|_F$ and $\|\mathbf{\Sigma}_{gg}^2\|_F$ be functions of $p$, and, letting $p$ be the asymptotic driving index assume that

$$\text{(A1) } \|\mathbf{\Sigma}_{gg}^2\|_F^2/\|\mathbf{\Sigma}_{gg}\|_F^4=o(1) \ \textit{as } p\to\infty.$$

**Theorem 2.** *Suppose that the null hypothesis $\mathcal{H}$ from (5) is true. Suppose further that (A1) is satisfied for all $g\in[k]$ and consider the asymptotic regime $n_1,p\to\infty$. Then, after suitable rescaling, $T$ is asymptotically normal, namely, $\sigma^{-1}T\rightsquigarrow\mathcal{N}(0,1)$ with $\sigma^2=2\sum_{1\le g<h\le k}n_g^{-2}$.*

*Proof.* See, Appendix C. □

**Remark 2.** *Under $\mathcal{H}$ and (A1),* $\mathrm{var}(T)/\sigma^2=1+o(1)$ *as $n_1,p\to\infty$.*

By Theorem 2, a critical value for the approximate size-$\alpha$ test can be calibrated based on the normal quantiles.

An alternative idea of how to express the test and show that we can control size is as follows. Using Theorem 2, our proposed approximate size-$\alpha$ test of the null hypothesis $\mathcal{H}$ can be based on the statistic

$$Q_\alpha=\mathbf{1}\left[\sum_{1\le g<h\le k}HRV_{gh}/\sigma\ge z_{1-\alpha}\right],$$

where $\mathbf{1}(\cdot)$ represents the indicator function and $z_\alpha=\Phi^{-1}(\alpha)$ denotes the upper $\alpha$ quantile of $\mathcal{N}(0,1)$. The following corollary states that the test $Q_\alpha$ can efficiently control the size.

**Collorary 1.** *Suppose that the condition (A1) holds, then, as $n_1,p\to\infty$,*

$$\Pr(Q_\alpha=1\mid\mathcal{H})=\alpha+o(1).$$

We further evaluate the power of $T$ under a kind of local alternative. Consider the alternative hypothesis

$$\mathcal{A}:\quad \mathbf{x}_g \text{ and } \mathbf{x}_h \text{ are dependent for some } g,h\in[k]$$

satisfying condition (6) below. Draw $n_i$ samples from such alternatives $\mathbf{x}$ following sampling scheme of Section 1 to form the respective analogues of $HRV_{gh}$ and $Q_\alpha$ and denote them by $HRV_{gh}^{\mathcal{A}}$ and $Q_\alpha^{\mathcal{A}}$, respectively.

**Theorem 3.** *In addition to assumptions in Theorems 2 let*

$$\Theta_{n_1} = \{\mathbf{\Sigma}_{\langle 1\rangle} : \max_{g<h} \rho V_{gh} \geq n_1^{-\delta}\} \tag{6}$$

*be a set of alternatives* $\mathbf{\Sigma}_{\langle 1\rangle}$ *such that* $\max_{g<h} \rho V_{gh} \geq n_1^{-\delta}$*, where* $0 < \delta < 1$*. Then, as* $n_1, p \to \infty$*,*

$$\inf_{\Theta_{n_1}} \Pr(Q_\alpha^{\mathcal{A}} = 1 \,|\, \mathcal{A}) = 1 + o(1).$$

*Proof.* See, Appendix D. □

## 3. Stepwise multiple significance test

In this section, we proceed to explore the proposed statistic $T$ and construct the new step-down multiple comparison significance test for simultaneous testing of independence.

Let $\mathcal{M}_q$ be the family of subsets with cardinal number $q \geq 2$ of the set $[k]$. Also, let these subsets be denoted by $m = \{\ell_1, \ldots, \ell_q\} \in \mathcal{M}_q$ where $\ell_1 < \cdots < \ell_q$ and let $\mathbf{\Sigma}^{(q,m)}$ be the following $pq \times pq$ matrix for these $m$:

$$\mathbf{\Sigma}^{(q,m)} = \begin{pmatrix} \mathbf{\Sigma}_{\ell_1\ell_1} & \cdots & \mathbf{\Sigma}_{\ell_1\ell_q} \\ \vdots & \ddots & \vdots \\ \mathbf{\Sigma}_{\ell_q\ell_1} & \cdots & \mathbf{\Sigma}_{\ell_q\ell_q} \end{pmatrix}.$$

We wish to test the following hypothesis:

$$\mathcal{H}^{\{q,m\}}: \quad \forall g \neq h \in \{\ell_1, \ldots, \ell_q\}, \ \ \mathbf{\Sigma}_{gh} = \mathbf{O} \quad \text{vs.} \quad \mathcal{A}^{\{q,m\}}: \quad \neg\mathcal{H}^{\{q,m\}},$$

and for this, we obtain the test statistic $T^{\{q,m\}}/\sigma^2_{\{q,m\}}$ based on results of Section 2, where $T^{\{q,m\}} = \sum_{g<h}^{\ell_q} HRV_{gh}$ and $g \neq h \in \{\ell_1, \ldots, \ell_q\}$. Here, we consider the problem of testing family of hypotheses $\mathcal{F} = \{\mathcal{H}^{\{2,m\}} : \mathbf{\Sigma}_{\ell_1\ell_2} = \mathbf{O}, m \in \mathcal{M}_2\}$.

Let $\mathcal{G}_q$ be the set consisting of all hypothesis $\mathcal{H}^{\{q,m\}}$ and let $\mathcal{G} = \cup_{q=2}^k \mathcal{G}_q$. Then the family $\mathcal{G}$ is closed. Hence, we can derive a step-down multiple comparison procedure based on closed testing procedure for $\mathcal{G}$. We define

$$\alpha_q = \begin{cases} 1 - (1-\alpha)^{q/k} & \text{for } q \in \{2, \ldots, k-2\} \\ \alpha & \text{for } q \in \{k-1, k\} \end{cases}$$

and let $t_{\{q,m\}}(\alpha)$ be the upper $\alpha$ percentiles of the statistic $T^{\{q,m\}}$ under $\mathcal{H}^{\{q,m\}}$, that is, $t_{\{q,m\}}(\alpha)$ satisfies $\Pr\{T^{\{q,m\}} \geq t_{\{q,m\}}(\alpha)\} = \alpha$. Then we carry out the following Tukey-Welsch type step-down multiple test for all hypotheses in $\mathcal{G}$ by using the $T^{\{q,m\}}$:

**Step 1.** We test hypothesis $\mathcal{H}^{\{k,m\}} = \mathcal{H}$.

(C1) If $T \geq t(\alpha_k)$, we reject $\mathcal{H}$ and go to Step 2.

(C2) If $T < t(\alpha_k)$, we retain all hypotheses in $\mathcal{G}$ and stop the test. Here, $t(\alpha)$ satisfies $\Pr\{T \geq t(\alpha)\} = \alpha$ under $\mathcal{H}$.

**Step 2.** We test all hypotheses $\mathcal{H}^{\{k-1,m\}}$ in $\mathcal{G}_{k-1}$.

(C1) If $T^{\{k-1,m\}} \geq t_{\{k-1,m\}}(\alpha_{k-1})$, we reject $\mathcal{H}^{\{k-1,m\}}$.

(C2) If $T^{\{k-1,m\}} < t_{\{k-1,m\}}(\alpha_{k-1})$, we retain $\mathcal{H}^{\{k-1,m\}}$ and all hypotheses in $\cup_{q=2}^{k-2}\mathcal{G}_q$ implied by $\mathcal{H}^{\{k-1,m\}}$.

If all hypotheses in $\cup_{q=2}^{k-2}\mathcal{G}_q$ are retained, we finish the test. Otherwise, we go to Step 3.

**Step 3.** We test all hypotheses $\mathcal{H}^{\{k-2,m\}}$ in $\mathcal{G}_{k-2}$ which are not retained in Step 2.

(C1) If $T^{\{k-2,m\}} \geq t_{\{k-2,m\}}(\alpha_{k-2})$, we reject $\mathcal{H}^{\{k-2,m\}}$.

(C2) If $T^{\{k-2,m\}} < t_{\{k-2,m\}}(\alpha_{k-2})$, we retain $\mathcal{H}^{\{k-2,m\}}$ and all hypotheses in $\cup_{q=2}^{k-3}\mathcal{G}_q$ implied by $\mathcal{H}^{\{k-2,m\}}$.

If all hypotheses in $\cup_{q=2}^{k-3}\mathcal{G}_q$ are retained, we finish the test. Otherwise, we repeat similar judgments till Step $k-1$ at the maximum.

**Remark 3.** *From a principle of closed testing procedure, we note that the maximum type-I FWE (family-wise error rate) of our proposed step-down multiple comparison procedure is not greater than $\alpha$.*

By the results of Theorem 2, the critical values $t_{\{q,m\}}(\alpha)$ for an approximate $\alpha$-size test can be set as $\sigma_{\{q,m\}}z_{1-\alpha}$, where $\sigma_{\{q,m\}}z_{1-\alpha}$ satisfies $\Pr\{T^{\{q,m\}} \geq \sigma_{\{q,m\}}z_{1-\alpha}\} = \alpha + o(1)$ under $\mathcal{H}^{\{q,m\}}$ and assuming that (A1) holds.

## 4. Numerical study

We present results from numerical studies which are designed to evaluate the performance of the proposed statistic $T$ for testing independence hypothesis $\mathcal{H}$ and for the multiple comparison procedure based on the simultaneous test of independence. Our simulations explore the size of the tests when critical values are selected using asymptotic normality of $T$ and compare their power for a number of alternative scenarios. We also employ the proposed test to analyze data from Electroencephalograph (EEG) experiment to illustrate the application of our results.

*4.1. Simulation experiments*

We first assess the accuracy of the proposed test statistic for its size control. The test compares a rescaled statistic $T$ to the limiting standard normal distribution from the Theorem 2. Targeting the size of $\alpha = 0.05$, the null hypothesis (2) is rejected when the value of the rescaled statistic exceeds the 0.95th percentile of the standard normal distribution. With $\ell$ replications of the data set generated under the null hypothesis $\mathcal{H}$, we calculate the empirical size as

$$\hat{\alpha}_T = \frac{\#\{T^{\mathcal{H}}/\sigma \geq z_\alpha\}}{\ell},$$

where $T^{\mathcal{H}}$ represents the values of of the test statistic $T$ based on the data generated under the null hypothesis.

Table 1: The empirical power of proposed test $\alpha = 0.05$.

| $\boldsymbol{n}^\top$ | $p$ | Size | Power ($\mathcal{A}_1$) | | Power ($\mathcal{A}_2$) | |
|---|---|---|---|---|---|---|
| | | $\rho = 0.0$ | $\rho = 0.4$ | $\rho = 0.6$ | $\rho = 0.4$ | $\rho = 0.6$ |
| (20,20,20,20,20) | 50 | 0.060 | 0.159 | 0.361 | 0.333 | 0.783 |
| | 100 | 0.059 | 0.160 | 0.365 | 0.337 | 0.796 |
| | 200 | 0.060 | 0.156 | 0.367 | 0.336 | 0.801 |
| | 300 | 0.060 | 0.157 | 0.366 | 0.335 | 0.803 |
| (10,15,20,25,30) | 50 | 0.064 | 0.122 | 0.228 | 0.215 | 0.507 |
| | 100 | 0.064 | 0.124 | 0.231 | 0.215 | 0.518 |
| | 200 | 0.063 | 0.121 | 0.229 | 0.211 | 0.520 |
| | 300 | 0.063 | 0.122 | 0.229 | 0.215 | 0.519 |
| (40,40,40,40,40) | 50 | 0.055 | 0.313 | 0.788 | 0.728 | 0.998 |
| | 100 | 0.055 | 0.317 | 0.796 | 0.742 | 0.999 |
| | 200 | 0.055 | 0.312 | 0.798 | 0.745 | 0.999 |
| | 300 | 0.055 | 0.313 | 0.799 | 0.759 | 1.000 |
| (20,30,40,50,60) | 50 | 0.058 | 0.198 | 0.495 | 0.450 | 0.928 |
| | 100 | 0.056 | 0.196 | 0.496 | 0.455 | 0.937 |
| | 200 | 0.057 | 0.195 | 0.500 | 0.459 | 0.943 |
| | 300 | 0.055 | 0.194 | 0.498 | 0.456 | 0.943 |
| (60,60,60,60,60) | 50 | 0.053 | 0.522 | 0.976 | 0.947 | 1.000 |
| | 100 | 0.053 | 0.521 | 0.978 | 0.956 | 1.000 |
| | 200 | 0.054 | 0.524 | 0.979 | 0.961 | 1.000 |
| | 300 | 0.052 | 0.522 | 0.979 | 0.966 | 1.000 |
| (40,50,60,70,80) | 50 | 0.056 | 0.388 | 0.891 | 0.840 | 1.000 |
| | 100 | 0.053 | 0.387 | 0.895 | 0.853 | 1.000 |
| | 200 | 0.052 | 0.388 | 0.900 | 0.860 | 1.000 |
| | 300 | 0.053 | 0.390 | 0.901 | 0.862 | 1.000 |

Table 1 reports Monte-Carlo estimates of the finite-sample sizes for a variety of combinations of $\boldsymbol{n} = (n_1, \dots, n_k)$ and $p$ to reflect both large-sample and high-dimensional scenarios. The data underlying the table are i.i.d. $p \times k$-variate normal with $k = 5$ and the covariance matrix having the following within-block structures $\boldsymbol{\Sigma}_{\langle 1 \rangle} = \text{diag}(\boldsymbol{\Sigma}_{11}, \cdots, \boldsymbol{\Sigma}_{55})$, where each $\boldsymbol{\Sigma}_{gg}$ has an $AR(1)$ structure, i.e., $\boldsymbol{\Sigma}_{gg} = g(0.5^{|i-j|})$. For each combination of $p$ and $\boldsymbol{n}$, empirical sizes of the tests are calculated from $\ell = 100,000$ independently generated data sets. As expected, in general, the tests have their sizes converging to the the nominal level 0.05 as both $p$ and $n_k$ increase together. For certain combinations of $p$ and $n_k$, the test sizes are not very satisfactory when $n_k$ are very small, but they all become close to the nominal 0.05 level when $n_k$ get above 40-50, indicating that the asymptotic properties of $T$ described by Theorem 2 pitch in.

Next, we consider the power of the tests, as studied in Section 2. The empirical power is calculated as

$$\hat{\beta}_T = \frac{\#\left\{T^{\mathcal{A}}/\sigma \geq z_\alpha\right\}}{\ell},$$

where $T^{\mathcal{A}}$ represents the values of the test statistic $T$ based on the data generated under the alternative hypothesis. For different combinations of $\boldsymbol{n} = (n_1, \ldots, n_k)$ and $p$, we generate data as a set of independent draws from two $p \times k$-variate normal distributions with different forms of alternatives of the covariance structure of $\boldsymbol{\Sigma}_{[1]}$. These are

$$\text{i) } \mathcal{A}_1 : \boldsymbol{\Sigma}_{[1]} = \operatorname{diag}(\boldsymbol{\Sigma}_{11}, \cdots, \boldsymbol{\Sigma}_{55}) + \begin{pmatrix} 0 & \eta & 0 & 0 & 0 \\ \eta & 0 & \eta & 0 & 0 \\ 0 & \eta & 0 & \eta & 0 \\ 0 & 0 & \eta & 0 & \eta \\ 0 & 0 & 0 & \eta & 0 \end{pmatrix} \otimes \boldsymbol{\Sigma}_{11}.$$

$$\text{ii) } \mathcal{A}_2 : \boldsymbol{\Sigma}_{[1]} = \operatorname{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{55}) + \begin{pmatrix} 0 & \eta & \eta & \eta & 0 \\ \eta & 0 & \eta & \eta & \eta \\ \eta & \eta & 0 & \eta & \eta \\ \eta & \eta & \eta & 0 & \eta \\ 0 & \eta & \eta & \eta & 0 \end{pmatrix} \otimes \boldsymbol{\Sigma}_{11}.$$

Further, for each distribution, two levels of $\eta$'s are considered: 0.4 and 0.6. The power of the proposed test is largely dependent on (i) the sample size $n_g$, and (ii) the variation in $\eta$ as it determines $\sum_{g<h}^{k} \rho V_{gh}$, the quantity which in turn determines the asymptotic power of the test as shown in Theorem 3. Specifically, if $\mathcal{A}_1$ holds, then $\sum_{g<h}^{k} \rho V_{gh} = 8\eta^2/5$, and if $\mathcal{A}_2$ holds, then $\sum_{g<h}^{k} \rho V_{gh} = 67\eta^2/20$. Observe that the value $\sum_{g<h}^{k} \rho V_{gh}$ does not depend on the dimension $p$, hence the power is expected to be mainly related to the value of $\eta$ and number of $\eta$ entries in the covariance structures under alternatives. In particular, the first alternative, $\mathcal{A}_1$, is designed to challenge the test procedure for some near block-diagonal structures with sparsely distributed non-zero off-blocks entries, whereas the second alternative, $\mathcal{A}_2$ represents a dense alternative. Table 1 reports empirical power $\hat{\beta}$ of the test for a range of configurations of $p$ and $\boldsymbol{n} = (n_1, n_2, n_3, n_4, n_5)^\top$, computed based on $\ell = 100{,}000$ replications of the experiments for the test based on $T$. We find the powers for the alternative $\mathcal{A}_1$ are less affected by the increased dimensionality as compared to $\mathcal{A}_2$. Overall, the power of $T$ under the second alternative increases systematically much faster than that under the first alternative, as the sample sizes and the dimension are increased. Tacking into account that the value of $\sum_{g<h}^{k} \rho V_{gh}$ is systematically

larger under $\mathcal{A}_2$ as compared to $\mathcal{A}_1$, this is a natural trend. And when $\eta$ increases from 0.4 to 0.6 the power gets larger under both alternatives since the increase of $\eta$ contributes to the increase of each $\rho V_{gh}$, which measures the departure from the null hypothesis. With $\eta$ increased under the second alternative, many entries of empirical powers of the test approach 1, which could be viewed as an empirical indication of the proposed test being consistent.

Next, we investigate the behavior of test power when the alternative hypothesis depends on $n_1$. According to Theorem 3, the power converge to 1 when $\delta < 1$. To reflect the conditions of Theorem 3, we further $\delta$ and $c$:

$$\mathcal{A}_3 : \mathbf{\Sigma}_{\langle 1 \rangle} = \mathrm{diag}(\mathbf{\Sigma}_{11}, \cdots, \mathbf{\Sigma}_{55}) + c n_1^{-\delta/2} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \otimes \mathbf{\Sigma}_{11}.$$

We set $(\delta, c) \in \{(0.3, 1.0), (0.5, 1.5), (0.7, 2.0)\}$, $p = 5 \times n_1$, $n_1 = \cdots = n_5$, and $n_1 = 10 \times i$, where $i \in \{1, \ldots, 20\}$. The empirical powers of our proposed test are listed in Figure 1 for the case $(\delta, c) = (0.3, 1.0)$ ($\bullet$), for the case $(\delta, c) = (0.5, 1.5)$ ($\times$), and for the case $(\delta, c) = (0.7, 2.0)$ ($\triangle$).

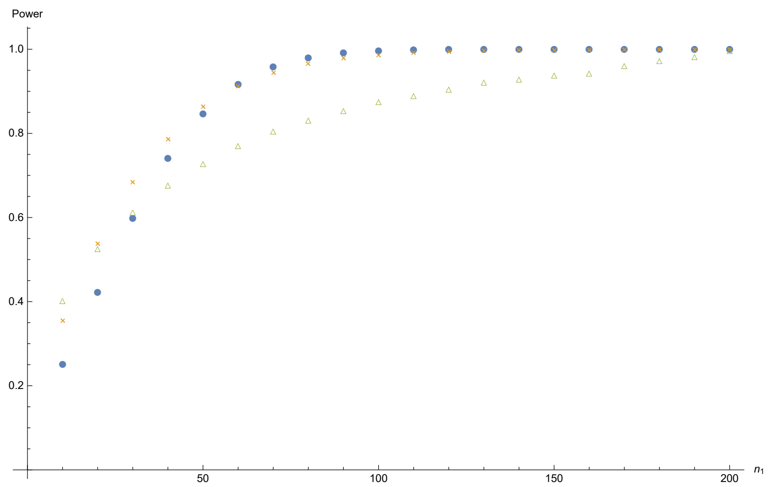


Figure 1: The empirical powers

As the results in Figure 1 show, $T$ tends to have rather similar power converging to 1 across the set of alternatives generated by $(\delta, c)$. This convergence becomes slower with larger values of $\delta$.

Finally, we investigate the probability of selecting the correct model with the proposed multiple comparison procedure. Let $\boldsymbol{\Sigma}_{\langle 1\rangle}$ has the following structure

$$\boldsymbol{\Sigma}_{\langle 1\rangle} = \mathrm{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{44}) + \begin{pmatrix} 0 & \sqrt{2}\eta & 0 & 0 \\ \sqrt{2}\eta & 0 & \sqrt{6}\eta & 0 \\ 0 & \sqrt{6}\eta & 0 & 2\sqrt{3}\eta \\ 0 & 0 & 2\sqrt{3}\eta & 0 \end{pmatrix} \otimes \boldsymbol{\Sigma}_{11}.$$

We check whether the proposed procedure can correctly capture this covariance structure. That is, we count the number of times that $\mathcal{H}^{\{2,\{1,3\}\}}$, $\mathcal{H}^{\{2,\{1,4\}\}}$ and $\mathcal{H}^{\{2,\{2,4\}\}}$ are retained by the procedure. As $\eta$ is larger, the selection probability tends to be larger. Also, as the dimension $p$ increases, the selection probability slightly increases. When the total sample size is the same, the balance type has a higher selection probability than the unbalanced type.

Table 2: The empirical power of multiple comparison procedure.

| $\rho$ | 0.4 | | | 0.5 | | | 0.6 | | |
|---|---|---|---|---|---|---|---|---|---|
| $\boldsymbol{n}^\top \setminus p$ | 100 | 200 | 300 | 100 | 200 | 300 | 100 | 200 | 300 |
| (40,40,40,40) | 0.321 | 0.336 | 0.342 | 0.889 | 0.906 | 0.909 | 0.999 | 0.999 | 0.999 |
| (30,35,45,50) | 0.173 | 0.182 | 0.183 | 0.735 | 0.747 | 0.750 | 0.980 | 0.980 | 0.981 |
| (50,50,50,50) | 0.654 | 0.674 | 0.678 | 0.987 | 0.989 | 0.991 | 1.000 | 1.000 | 1.000 |
| (40,45,55,60) | 0.490 | 0.510 | 0.514 | 0.948 | 0.954 | 0.955 | 1.000 | 1.000 | 1.000 |
| (60,60,60,60) | 0.865 | 0.879 | 0.881 | 0.999 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 |
| (50,55,65,70) | 0.762 | 0.779 | 0.783 | 0.995 | 0.996 | 0.997 | 1.000 | 1.000 | 1.000 |
| (70,70,70,70) | 0.960 | 0.965 | 0.967 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| (60,65,75,80) | 0.915 | 0.923 | 0.926 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

*4.2. Applications : An example*

For illustration, we employ the step-down multiple comparison significance testing to analyze the Electroencephalography (EEG) data publicly available at the University of California-Irvine Machine Learning Repository, web address

`https://archive.ics.uci.edu/ml/datasets/EEG+Database`

The data arose from a large study to examine Electroencephalograph (EEG) correlates of genetic predisposition to alcoholism. Monitoring of the brain electric activity is performed with 64 electrodes evenly distributed over subjects scalps and recording 256 measurements for 1 second. The initial study involved two groups of subjects: alcoholic and control. Each subject was exposed to either a single stimulus (S1) or to two stimuli (S1 and S2) which were pictures of objects chosen from a picture set. 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 data set has been analyzed by several statisticians for various purposes, see e.g Harrar and Kong [7] whose main hypotheses of interest are whether ERP profiles are similar between the alcoholic and control groups, and if different, to identify for which electrode (which part of the brain) dissimilarity occurs.

In this paper, we conduct the analysis the for the single stimulus (S1) exposure in the alcoholic group. We are interested in testing the independence of the level of electrical activity withing the frontal regions of the brain. Specifically, the data set we focus on, consists of four channels (electrodes) FC1, FCz, FC2 and Cz where each channel has names identifying the location of the electrode on the scalp; F stands for frontal lobe, letter z (zero) is used for the mid-line and C identifies the central location between the frontal and parietal lobes. Combinations of two letters indicates intermediate locations, for example FC is in between frontal and central electrode locations (see Figure 5 of Harrar and Kong [7] for illustration). In the notations of the paper, this data set comprises $k = 4$ sub-vectors (FC1 (1), FCz (2), FC2 (3), Cz (4)), each of dimensionality $p = 256$ with equal sample sizes, that is $n_i = 77$ for $i = 1, \dots, k$. The multiple comparison procedure proposed in Section 3 is applied to clarify whether the levels of the brain activity at FC1, FCz, FC2 and Cz channels are mutually independent. By setting the significance level as $\alpha = 0.05$, the testing model is established in a stepwise fashion as follows.

**Step 1.** We test hypothesis $\mathcal{H}^{\{4,\{1,2,3,4\}\}}$.
We calculate test statistic $T^{\{4,\{1,2,3,4\}\}}/\sigma_{\{4,\{1,2,3,4\}\}} \approx 32.44$, and we also obtain $z_{0.05} \approx 1.645$ as an approximate critical value. Therefore, we reject $\mathcal{H}^{\{4,\{1,2,3,4\}\}}$ and move on to Step 2.

**Step 2.** We test following hypotheses:

$$\mathcal{H}^{\{3,\{1,2,3\}\}},\ \mathcal{H}^{\{3,\{1,2,4\}\}},\ \mathcal{H}^{\{3,\{1,3,4\}\}},\ \mathcal{H}^{\{3,\{2,3,4\}\}}.$$

The test statistic corresponding to each hypothesis is calculated as follows:

$$T^{\{3,\{1,2,3\}\}}/\sigma_{\{3,\{1,2,3\}\}} \approx 44.56,\ T^{\{3,\{1,2,4\}\}}/\sigma_{\{3,\{1,2,4\}\}} \approx 18.43,$$
$$T^{\{3,\{1,3,4\}\}}/\sigma_{\{3,\{1,3,4\}\}} \approx 15.94,\ T^{\{3,\{2,3,4\}\}}/\sigma_{\{3,\{2,3,4\}\}} \approx 12.84.$$

We also obtain $z_{0.05} \approx 1.645$ as an approximate critical value. Therefore, we reject all hypotheses and move on to Step 3.

**Step 3.** We test following hypotheses:

$$\mathcal{H}^{\{2,\{1,2\}\}},\ \mathcal{H}^{\{2,\{1,3\}\}},\ \mathcal{H}^{\{2,\{1,4\}\}},\ \mathcal{H}^{\{2,\{2,3\}\}},\ \mathcal{H}^{\{2,\{2,4\}\}},\ \mathcal{H}^{\{2,\{3,4\}\}}$$

The test statistic corresponding to each hypothesis is calculated as follows:

$$T^{\{2,\{1,2\}\}}/\sigma_{\{2,\{1,2\}\}} \approx 29.41,\ T^{\{2,\{1,3\}\}}/\sigma_{\{2,\{1,3\}\}} \approx 26.38,$$
$$T^{\{2,\{1,4\}\}}/\sigma_{\{2,\{1,4\}\}} \approx 1.45,\ T^{\{2,\{2,3\}\}}/\sigma_{\{2,\{2,3\}\}} \approx 21.39,$$
$$T^{\{2,\{2,4\}\}}/\sigma_{\{2,\{2,4\}\}} \approx 1.06,\ T^{\{2,\{3,4\}\}}/\sigma_{\{2,\{3,4\}\}} \approx -0.22.$$

We also obtain $z_{1-\sqrt{0.95}} \approx 1.955$ as an approximate critical value.

To summarize, the hypotheses $\mathcal{H}^{\{2,\{1,2\}\}}$, $\mathcal{H}^{\{2,\{1,3\}\}}$, $\mathcal{H}^{\{2,\{2,3\}\}}$ are rejected, whereas $\mathcal{H}^{\{2,\{1,4\}\}}$, $\mathcal{H}^{\{2,\{2,4\}\}}$, $\mathcal{H}^{\{2,\{3,4\}\}}$ are retained. Hence, with the results above we have strong evidence to believe that the three channels, FC1, FCz, and FC2 correlate with each other, but there is no correlation between (FC1, FCz, FC2) and Cz. This suggests that the assumption on the cross-channel independence in such empirical studies may not be appropriate. The four steps of the testing model for this example are illustrated on the Figure below.

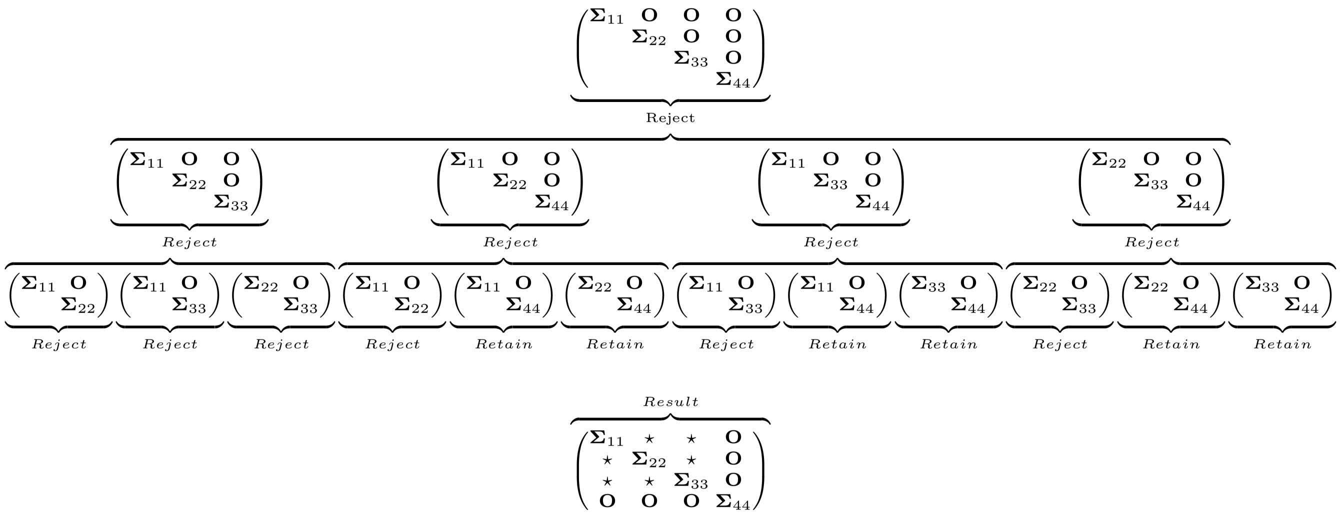


Figure 2:Three steps of testing independence of the brain activity in four brain channels with EGG data. The symbol in a block of the resulting structure indicates the presence of dependence between the corresponding channels.

## 5. Summary

A test statistic for mutual independence of $k$ random vectors coming for multivariate normal populations is developed when the dimensionality is large, possibly much larger than the sample size. Such a test is usually carried out as a preliminary test in large scale multivariate inference; examples include discriminant analysis or model-based clustering, testing equality of mean vectors and covariance matrices where the ability to detect departures from independence is of crucial importance.

With a view towards alternatives in which dependence is spread out over vector components, the test statistic is formed as a sum of consistent estimators of the pairwise vector correlation coefficients. The corresponding asymptotic theory is then developed to derive the asymptotic normal limit of the proposed test when both sample size and dimensionality go to infinity. A step-down multiple comparison procedure that allows to controls the family-wise error rate under independence is presented as a direct by-product.

Simulation results are used to demonstrate the finite sample performance of the test with respect to its size control and power, for large samples, arbitrary dimensions and a variety of dependence structure models often used in multivariate analysis. Our methodology is illustrated with the Electroencephalography (EEG) data where we applied the proposed step-down procedure for assessment of independence of electrical activity over certain regions of the human brain.

## 6. Acknowledgements

The research of the first and second authors were supported by JSPS KAKENHI Grant Numbers 17K14238, 17K00056. The third author is in part supported by Grant 2013-45266 of the National Research Council of Sweden (VR).

## References

[1] T. W. Anderson, An Introduction to Multivariate Statistical Analysis,thirded., John Wiley and Sons, NewYork, 2003.

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

[3] Y. Escoufier, Le Traitement des Variables Vectorielles, Biometrics 29 (1973) 751–760.

[4] B. Efron, Are a set of microarrays independent of each other?, The Annals of Applied Statistics 3 (3) (2009), 922–942.

[5] K. T. Fang, Y. T. Zhang, Generalized Multivariate Analysis. Springer, 1990.

[6] Han, F. and Liu, H. (2014). Distribution-free tests of independence with applications to testing more structures. ArXiv:1410.4179.

[7] S. W. Harrar, X. Kong, High-dimensional multivariate repeated measures analysis with unequal covariance matrices, Journal of Multivariate Analysis 145 (2016) 1-21.

[8] M. Hyodo, N. Shutoh, T. Nishiyama, T. Pavlenko, Testing block-diagonal covariance structure for high-dimensional data, Statistica Neerlandica 69 (2015) 460–482.

[9] D. Jiang, Z. Bai, S. Zheng, Testing the independence of sets of large-dimensional variables, Science China Mathematics 56 (2013) 135–147.

[10] J. Josse, J. Pagés, F. Husson, Testing the significance of the $RV$ coefficient, Computational Statistics and Data Analysis 53 (2008) 82–91.

[11] D. Leung, M. Drton, Testing independence in high dimensions with sums of rank correlations https://arxiv.org/abs/1501.01732

[12] K. V. Mardia, J. T. Kent, J. M. Bibby, Multivariate analysis. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, 1979.

[13] K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Philosophical Magazine 50 (1900) 157–175.

[14] A. N. Shiryaev, Probability, 2nd Ed., Springer-Verlag, New-York, 1984.

[15] M. S. Srivastava, N. M. Reid, Testing the structure of the covariance matrix with fewer observations than the dimension, Journal of Multivariate Analysis 112 (2012) 156–171.

[16] G. J. Szkely, M. L. Rizzo, The distance correlation $t$-test of independence in high dimension, Journal of Multivariate Analysis 112 (2013) 193–213.

[17] S. Taskinen, H. Oja, R. H. Randles, Multivariate nonparametric tests of independence, Journal of the American Statistical Association 100 (2005) 916–925.

[18] Y. Yang, G. Pan, Independence test for high dimensional data based on regularized canonical correlation coefficients, The Annals of Statistics 43 (2015) 467–500.

## A. Proof of Lemma 1

First, we evaluate expectation and variance of $\|\mathbf{S}_{\ell\ell}\|_F^2$:

$$\begin{aligned}
\mathrm{E}(\|\mathbf{S}_{\ell\ell}\|_F^2) &= \frac{n_\ell}{n_\ell - 1}\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^2 + \frac{\{\mathrm{tr}(\boldsymbol{\Sigma}_{\ell\ell})\}^2}{n_\ell - 1}, \\
\mathrm{var}\left(\|\mathbf{S}_{\ell\ell}\|_F^2\right) &= \frac{8}{(n_\ell - 1)^3}\{\mathrm{tr}(\boldsymbol{\Sigma}_{\ell\ell})\}^2\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^2 + \frac{4n_\ell}{(n_\ell - 1)^3}\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^4 \\
&\quad + \frac{16n_\ell}{(n_\ell - 1)^3}\mathrm{tr}(\boldsymbol{\Sigma}_{\ell\ell})\mathrm{tr}(\boldsymbol{\Sigma}_{\ell\ell}^3) + \frac{4(2n_\ell^2 + n_\ell + 2)}{(n_\ell - 1)^3}\mathrm{tr}(\boldsymbol{\Sigma}_{\ell\ell}^4).
\end{aligned}$$

Thus, we obtain

$$\|\mathbf{S}_{\ell\ell}\|_F^2 = 1 + \frac{\{\mathrm{tr}(\boldsymbol{\Sigma}_{\ell\ell})\}^2}{n_\ell\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^2} + O_p(n_\ell^{-1/2}). \tag{7}$$

Next, we evaluate expectation and variance of $\|\mathbf{S}_{gh}\|_F^2$:

$$\begin{aligned}
\mathrm{E}(\|\mathbf{S}_{gh}\|_F^2) &= \frac{n_g}{n_g - 1}\|\boldsymbol{\Sigma}_{gh}\|_F^2 + \frac{\mathrm{tr}(\boldsymbol{\Sigma}_{gg})\mathrm{tr}(\boldsymbol{\Sigma}_{hh})}{n_g - 1}, \\
\mathrm{var}\left(\|\mathbf{S}_{gh}\|_F^2\right) &= \frac{2}{(n_g - 1)^3}[\{\mathrm{tr}(\boldsymbol{\Sigma}_{gg})\}^2\|\boldsymbol{\Sigma}_{hh}\|_F^2 + \{\mathrm{tr}(\boldsymbol{\Sigma}_{hh})\}^2\|\boldsymbol{\Sigma}_{gg}\|_F^2 \\
&\quad + 2\mathrm{tr}(\boldsymbol{\Sigma}_{gg})\mathrm{tr}(\boldsymbol{\Sigma}_{hh})\|\boldsymbol{\Sigma}_{gh}\|_F^2] \\
&\quad + \frac{2n_g}{(n_g - 1)^3}(\|\boldsymbol{\Sigma}_{gh}\|_F^4 + \|\boldsymbol{\Sigma}_{gg}\|_F^2\|\boldsymbol{\Sigma}_{hh}\|_F^2) \\
&\quad + \frac{8n_g}{(n_g - 1)^3}\mathrm{tr}(\boldsymbol{\Sigma}_{gg})\mathrm{tr}(\boldsymbol{\Sigma}_{hh}\boldsymbol{\Sigma}_{hg}\boldsymbol{\Sigma}_{gh}) \\
&\quad + \frac{8n_g}{(n_g - 1)^3}\mathrm{tr}(\boldsymbol{\Sigma}_{hh})\mathrm{tr}(\boldsymbol{\Sigma}_{gg}\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hg}) \\
&\quad + \frac{4(n_g^2 + n_g + 2)}{(n_g - 1)^3}\mathrm{tr}(\boldsymbol{\Sigma}_{gg}\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hh}\boldsymbol{\Sigma}_{hg}) \\
&\quad + \frac{4n_g^2}{(n_g - 1)^3}\|\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hg}\|_F^2.
\end{aligned}$$

Thus, we obtain

$$\frac{\|\mathbf{S}_{gh}\|_F^2}{\|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F} = \rho V_{gh} + \frac{\mathrm{tr}(\boldsymbol{\Sigma}_{gg})\mathrm{tr}(\boldsymbol{\Sigma}_{hh})}{n_g\|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F} + O_p(n_g^{-1/2}). \tag{8}$$

Combining (7) and (8), the result is established.

## B. Proof of Theorem 1

First, we evaluate the variance of $\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2$, for which we obtain

$$\begin{aligned}\operatorname{var}\left(\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2\right) &= \frac{2}{(n_g-2)(n_g+1)}(\|\boldsymbol{\Sigma}_{gh}\|_F^4 + \|\boldsymbol{\Sigma}_{gg}\|_F^2\|\boldsymbol{\Sigma}_{hh}\|_F^2) \\ &+ \frac{4(n_g^2-5)}{(n_g-2)(n_g-1)(n_g+1)}\operatorname{tr}(\boldsymbol{\Sigma}_{gg}\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hh}\boldsymbol{\Sigma}_{hg}) \\ &+ \frac{4}{n_g-1}\operatorname{tr}\{(\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hg})^2\}.\end{aligned}$$

From $\operatorname{tr}(\boldsymbol{\Sigma}_{gg}\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hh}\boldsymbol{\Sigma}_{hg}) \le \|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F\|\boldsymbol{\Sigma}_{gh}\|_F^2$, $\operatorname{tr}\{(\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hg})^2\} \le \|\boldsymbol{\Sigma}_{gh}\|_F^4$ and $\rho V_{gh} < 1$, we get $\operatorname{var}(\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2)/(\|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F)^2 = O\left(n_g^{-1}\right)$. Using Chebyshev's inequality, we obtain

$$\frac{\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2}{\|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F} = \rho V_{gh} + o_p\left(1\right). \tag{9}$$

Next, we evaluate the variance of $\|\widehat{\boldsymbol{\Sigma}_{\ell\ell}}\|_F^2$ for $\ell \in \{g,h\}$. It is obtained by

$$\operatorname{var}\left(\|\widehat{\boldsymbol{\Sigma}_{\ell\ell}}\|_F^2\right) = \frac{4\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^4}{(n_\ell-2)(n_\ell+1)} + \frac{4(2n_\ell^2-n_\ell-7)\|\boldsymbol{\Sigma}_{\ell\ell}^2\|_F^2}{(n_\ell-2)(n_\ell-1)(n_\ell+1)}.$$

From $\|\boldsymbol{\Sigma}_{\ell\ell}^2\|_F^2 \le \|\boldsymbol{\Sigma}_{\ell\ell}\|_F^4$, we get $\operatorname{var}(\|\widehat{\boldsymbol{\Sigma}_{\ell\ell}}\|_F^2)/\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^4 = O\left(n_g^{-1}\right)$. Using Chebyshev's inequality, we obtain

$$\frac{\|\widehat{\boldsymbol{\Sigma}_{\ell\ell}}\|_F^2}{\|\boldsymbol{\Sigma}_{\ell\ell}\|_F^2} = 1 + o_p\left(1\right). \tag{10}$$

From (9) and (10), we obtain $HRV_{gh} = \rho V_{gh} + o_p(1)$.

## C. Proof of Theorem 2

From (10), under (A1), $\|\widehat{\boldsymbol{\Sigma}_{\ell\ell}}\|_F = \|\boldsymbol{\Sigma}_{\ell\ell}\|_F(1+o_p(1))$ for $\ell \in \{g,h\}$. Thus $T = \widetilde{T} + o_p(1)$, where

$$\widetilde{T} = \sum_{g<h}^{k} \frac{\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2}{\|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F}.$$

Therefore, it sufficient to show the asymptotic normality of $\widetilde{T}$.

Under $\mathcal{H}$, we have

$$\forall g \in [k],\ j \in [n_k]\ \mathbf{x}_{gj} = \boldsymbol{\Sigma}_{gg}^{1/2}\mathbf{z}_{gj} + \boldsymbol{\mu}_g,$$

where $\mathbf{z}_{gj} \sim \mathcal{N}_p(\mathbf{0}, I_p)$ and $\mathbf{z}_{gj}$ are mutually independent. We define $p \times \ell$ matrix, $\mathbf{Z}_{g(\ell)} = (\mathbf{z}_{g1}, \ldots, \mathbf{z}_{g\ell})$ for $\ell \le n_g$. Note that each component of $\mathbf{Z}_{g(n_g)}$ independently $\mathcal{N}(0,1)$ distributed. For $g < h$, i.e. $n_g \le n_h$, under $\mathcal{H}$,

$$\begin{aligned}\|\widehat{\mathbf{\Sigma}_{gh}}\|_F^2 &= \frac{\mathrm{tr}(\mathbf{\Sigma}_{gg}\mathbf{Z}_{g(n_g)}\mathbf{Z}_{h(n_g)}^\top\mathbf{\Sigma}_{hh}\mathbf{Z}_{h(n_g)}\mathbf{Z}_{g(n_g)}^\top)}{n_g^2} \\ &\quad -\frac{\mathrm{tr}(\mathbf{\Sigma}_{gg}\mathbf{Z}_{g(n_g)}\mathbf{Z}_{g(n_g)}^\top)\mathrm{tr}(\mathbf{\Sigma}_{hh}\mathbf{Z}_{h(n_g)}\mathbf{Z}_{h(n_g)}^\top)}{n_g^3} + o_p\left(\frac{\|\mathbf{\Sigma}_{gg}\|_F\|\mathbf{\Sigma}_{hh}\|_F}{n_g}\right).\end{aligned}$$

Let $\mathbf{\Gamma}_g$ be orthogonal matrix s.t. $\mathbf{\Gamma}_g^\top\mathbf{\Sigma}_{gg}\mathbf{\Gamma}_g = \mathbf{\Lambda}_g = \mathrm{diag}(\lambda_{g1}, \ldots, \lambda_{gp})$. For $i \in [p]$, we define $\mathbf{u}_{gi} = (\mathbf{e}_i^\top\mathbf{\Gamma}_g^\top\mathbf{z}_{g1}, \ldots, \mathbf{e}_i^\top\mathbf{\Gamma}_g^\top\mathbf{z}_{gn_g})^\top$. Then $\mathbf{u}_{gi} \sim \mathcal{N}_{n_g}(\mathbf{0}, I_{n_g})$ and $\boldsymbol{e}_j^\top\mathbf{u}_{gi} = \mathbf{z}_{gj}^\top\mathbf{\Gamma}_g\boldsymbol{e}_i$ are mutually independent whenever $(g, i, j)$ are distinct indices.

Let $\mathbf{u}_{gi(\ell)} = (\mathbf{e}_i^\top\mathbf{\Gamma}_g^\top\mathbf{z}_{g1}, \ldots, \mathbf{e}_i^\top\mathbf{\Gamma}_g^\top\mathbf{z}_{g\ell})^\top$. Then

$$\mathbf{\Gamma}_{gg}^\top\mathbf{Z}_{g(n_g)} = (\mathbf{u}_{g1(n_g)}, \ldots, \mathbf{u}_{gp(n_g)})^\top,\ \mathbf{\Gamma}_{hh}^\top\mathbf{Z}_{h(n_g)} = (\mathbf{u}_{h1(n_g)}, \ldots, \mathbf{u}_{hp(n_g)})^\top.$$

Using these variables, we rewrite

$$\begin{aligned}\|\widehat{\mathbf{\Sigma}_{gh}}\|_F^2 &= n_g^{-2}\sum_{i=1}^p\sum_{j=1}^p\lambda_{gi}\lambda_{hj}(\mathbf{u}_{gi(n_g)}^\top\mathbf{u}_{hj(n_g)})^2 \\ &\quad -n_g^{-3}\sum_{i=1}^p\sum_{j=1}^p\lambda_{gi}\lambda_{hj}\mathbf{u}_{gi(n_g)}^\top\mathbf{u}_{gj(n_g)}\mathbf{u}_{hi(n_g)}^\top\mathbf{u}_{hj(n_g)} \\ &\quad +o_p\left(\frac{\|\mathbf{\Sigma}_{gg}\|_F\|\mathbf{\Sigma}_{hh}\|_F}{n_g}\right).\end{aligned}$$

Thus $\widetilde{T}$ can be expressed as $\widetilde{T}/\sigma = \sum_{i=1}^p \varepsilon_i + o_p(1)$, where

$$\begin{aligned}\varepsilon_i &= \sum_{g<h}\frac{1}{\sigma n_g^2}\left[\frac{\lambda_{gi}\lambda_{hi}}{\|\mathbf{\Sigma}_{gg}\|_F\|\mathbf{\Sigma}_{hh}\|_F}\left\{(\mathbf{u}_{gi(n_g)}^\top\mathbf{u}_{hi(n_g)})^2 - \frac{\|\mathbf{u}_{gi(n_g)}\|^2\|\mathbf{u}_{hi(n_g)}\|^2}{n_g}\right\}\right. \\ &\quad +\sum_{j=0}^{i-1}\frac{\lambda_{gi}\lambda_{hj}}{\|\mathbf{\Sigma}_{gg}\|_F\|\mathbf{\Sigma}_{hh}\|_F}\left\{(\mathbf{u}_{gi(n_g)}^\top\mathbf{u}_{hj(n_g)})^2 - \frac{\|\mathbf{u}_{gi(n_g)}\|^2\|\mathbf{u}_{hj(n_g)}\|^2}{n_g}\right\} \\ &\quad \left.+\sum_{j=0}^{i-1}\frac{\lambda_{gj}\lambda_{hi}}{\|\mathbf{\Sigma}_{gg}\|_F\|\mathbf{\Sigma}_{hh}\|_F}\left\{(\mathbf{u}_{gj(n_g)}^\top\mathbf{u}_{hi(n_g)})^2 - \frac{\|\mathbf{u}_{gj(n_g)}\|^2\|\mathbf{u}_{hi(n_g)}\|^2}{n_g}\right\}\right].\end{aligned}$$

Here, for $i = 1$, both the second and the third term ignore 0. Define $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and let $\mathcal{F}_i$ for $i \in \mathbb{N}$ be the $\sigma$-algebra generated by the random variables $U_{i-1}$, where

$$U_{i-1} = (\mathbf{u}_{11}, \ldots, \mathbf{u}_{1i-1}, \ldots, \mathbf{u}_{k1}, \ldots, \mathbf{u}_{ki-1}).$$

Then we find that $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_\infty$ and $\mathrm{E}(\varepsilon_i|\mathcal{F}_{i-1}) = 0$. Thus, $(\varepsilon_i)$ is a martingale difference sequence. We show the asymptotic normality of $\varepsilon_1 + \cdots + \varepsilon_p$ by adapting the martingale difference central limit theorem; see, e.g., [14]. Let $E_{i-1} = \mathrm{E}(\varepsilon_i^2|\mathcal{F}_{i-1})$. Then

$$\mathrm{E}\left(\sum_{i=1}^{p} E_{i-1}\right) = 1 + o(1), \ \mathrm{var}\left(\sum_{i=1}^{p} E_{i-1}\right) = O(n_1^{-1}).$$

Thus, (I) : $\sum_{i=1}^{p} E_{i-1} = 1 + o_p(1)$ as $n_1 \to \infty$. Also

$$\sum_{i=1}^{p} \mathrm{E}(\varepsilon_i^4) = O\left(\sum_{g=1}^{k} \frac{\|\mathbf{\Sigma}_{gg}^2\|^2}{\|\mathbf{\Sigma}_{gg}\|^4}\right).$$

Thus, under (A1), (II) : $\sum_{i=1}^{p} \mathrm{E}(\varepsilon_i^4) = o(1)$ as $p \to \infty$. The above results (I) and (II) complete the proof.

## D. Proof of Theorem 3

From (10), the power of our proposed test at $\mathbf{\Sigma}_{\langle 1\rangle}$ is $\Pr(Q_\alpha^{\mathcal{A}} = 1 \,|\, \mathcal{A}) = \Pr(\widetilde{T} \geq \sigma z_\alpha) + o(1)$ as $n_1 \to \infty$. Thus it is sufficient to show that $\Pr(\widetilde{T} \geq \sigma z_\alpha) = 1 + o(1)$ for any $\mathbf{\Sigma}_{\langle 1\rangle} \in \Theta_{n_1}$.

We note that $\mathrm{E}(\widetilde{T}) = \sum_{g<h} \rho V_{gh} > 0$ for any $\mathbf{\Sigma}_{\langle 1\rangle} \in \Theta_{n_1}$, and

$$\Pr\left(\widetilde{T} \geq \sigma z_\alpha\right) \geq 1 - \Pr\left(|\widetilde{T} - \mathrm{E}(\widetilde{T}) - \sigma z_\alpha| \geq \mathrm{E}(\widetilde{T})\right).$$

Using Markov's inequality and Cauchy-Schwarz inequality in the context of expectation, we obtain

$$\begin{aligned}\Pr\left(\left|\widetilde{T} - \mathrm{E}(\widetilde{T}) - \sigma z_\alpha\right| \geq \mathrm{E}(\widetilde{T})\right) &\leq \mathrm{E}\left(\left|\widetilde{T} - \mathrm{E}(\widetilde{T}) - \sigma z_\alpha\right| / \mathrm{E}(\widetilde{T})\right) \\ &\leq \mathrm{E}\left(\left|\widetilde{T} - \mathrm{E}(\widetilde{T}) - \sigma z_\alpha\right|^2\right) / \{\mathrm{E}(\widetilde{T})\}^2.\end{aligned}$$

Since $\mathrm{E}(|\widetilde{T} - \mathrm{E}(\widetilde{T}) - \sigma z_\alpha|^2) = \mathrm{var}(\widetilde{T}) + \sigma^2 z_\alpha^2$, we obtain

$$\Pr(\widetilde{T} \geq \sigma z_\alpha) \quad \geq \quad 1 - \frac{\mathrm{var}(\widetilde{T}) + \sigma^2 z_\alpha^2}{\{\mathrm{E}(\widetilde{T})\}^2}. \tag{11}$$

We further evaluate $\mathrm{var}(\widetilde{T})$. For any $g < h, g, h \in [k]$, we define $A_{gh} = \widetilde{HRV}_{gh} - \rho V_{gh}$, where $\widetilde{HRV}_{gh} = \|\widehat{\mathbf{\Sigma}_{gh}}\|_F^2 / (\|\mathbf{\Sigma}_{gg}\|_F \|\mathbf{\Sigma}_{hh}\|_F)$. Then

$$\frac{\mathrm{var}(\widetilde{T})}{\{\mathrm{E}(\widetilde{T})\}^2} = \frac{\mathrm{E}\{(\sum_{g<h} A_{gh})^2\}}{\{\mathrm{E}(\widetilde{T})\}^2} \leq \frac{k(k-1)\sum_{g<h} \mathrm{E}(A_{gh}^2)}{2\{\mathrm{E}(\widetilde{T})\}^2}.$$

$\mathrm{E}(A_{gh}^2)$ is obtained by

$$\mathrm{E}(A_{gh}^2) = \frac{2}{(n_g-2)(n_g+1)}(\rho V_{gh}^2+1) + \frac{4(n_g^2-5)\mathrm{tr}(\boldsymbol{\Sigma}_{gg}\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hh}\boldsymbol{\Sigma}_{hg})}{(n_g-2)(n_g-1)(n_g+1)\|\boldsymbol{\Sigma}_{gg}\|_F^2\|\boldsymbol{\Sigma}_{hh}\|_F^2} + \frac{4\mathrm{tr}\{(\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hg})^2\}}{(n_g-1)\|\boldsymbol{\Sigma}_{gg}\|_F^2\|\boldsymbol{\Sigma}_{hh}\|_F^2}.$$

From $\mathrm{tr}(\boldsymbol{\Sigma}_{gg}\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hh}\boldsymbol{\Sigma}_{hg}) \le \|\boldsymbol{\Sigma}_{gg}\|_F\|\boldsymbol{\Sigma}_{hh}\|_F\|\boldsymbol{\Sigma}_{gh}\|_F^2$ and $\mathrm{tr}\{(\boldsymbol{\Sigma}_{gh}\boldsymbol{\Sigma}_{hg})^2\} \le \|\boldsymbol{\Sigma}_{gh}\|_F^4$, we get

$$\frac{\mathrm{E}(A_{gh}^2)}{\{\mathrm{E}(\widetilde{T})\}^2} = O\left(\frac{\rho V_{gh}+\rho V_{gh}^2}{(\sum_{g<h}\rho V_{gh})^2 n_g} + \frac{\rho V_{gh}^2+1}{(\sum_{g<h}\rho V_{gh})^2 n_g^2}\right)$$

Note that $\mathrm{E}(\widetilde{T}) = \sum_{g<h}\rho V_{gh} \ge \max_{g<h}\rho V_{gh} \ge n_1^{-\delta}$. Thus, for any $\boldsymbol{\Sigma}_{\langle 1\rangle} \in \Theta_{n_1}$,

$$\frac{\mathrm{E}(A_{gh}^2)}{\{\mathrm{E}(\widetilde{T})\}^2} = O\left(\frac{1}{n_1^{1-\delta}} + \frac{1}{n_1} + \frac{1}{n_1^2}\right).$$

Since $k$ is fixed, we obtain

$$\frac{\mathrm{var}(\widetilde{T})}{\{\mathrm{E}(\widetilde{T})\}^2} = O\left(\frac{1}{n_1^{1-\delta}} + \frac{1}{n_1} + \frac{1}{n_1^2}\right). \tag{12}$$

Next, we evaluate $\sigma^2/\{\mathrm{E}(\widetilde{T})\}^2$. Since $\sigma^2 = O(n_1^{-2})$ and $\{\mathrm{E}(\widetilde{T})\}^2 \ge (\max_{g<h}\rho V_{gh})^2 \ge n_1^{-2\delta}$, we obtain

$$\frac{\sigma^2 z_\alpha^2}{\{\mathrm{E}(\widetilde{T})\}^2} = O\left(\frac{1}{n^{2(1-\delta)}}\right). \tag{13}$$

Substituting (12) and (13) to (11), for any $\boldsymbol{\Sigma}_{\langle 1\rangle} \in \Theta_{n_1}$,

$$\Pr(\widetilde{T} \ge \sigma z_\alpha) = 1 + O\left(\frac{1}{n_1^{1-\delta}} + \frac{1}{n_1^{2(1-\delta)}} + \frac{1}{n_1} + \frac{1}{n_1^2}\right).$$

Therefore, under $n_1 \to \infty$, $\inf_{\Theta_{n_1}} \Pr(\widetilde{T} \ge \sigma z_\alpha) = 1 + o(1)$.