Low dimensional factor model-based tests for assessing vector correlation in high-dimensional settings

Masashi Hyodo[†], Takahiro Nishiyama[††], Shoichi Narita[†††]

[†] *Department of Economics, Kanagawa university, 3-27-1 Rokkakubashi, Kanagawa-ku, Yokohama-shi, Kanagawa, Japan. E-Mail: caicmhy@gmail.com*[††] *Department of Business Administration, Senshu University, 2-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa, Japan. E-Mail: nishiyama@isc.senshu-u.ac.jp*

[†††] *Graduate School of Economics, Kanagawa university, 3-27-1 Rokkakubashi, Kanagawa-ku, Yokohama-shi, Kanagawa, Japan. E-Mail: r202470016kq@jindai.jp*

Abstract

This study proposes a new test for vector correlation in a high-dimensional framework, while accommodating a low-dimensional latent factor model. Our test, built under low-dimensional factor models, distinguishes from previous normal approximation- based tests, which are valid under a weak spike structure. We propose a modified RV coefficient for high-dimensional data, and show that its null-limiting distributions follow a weighted mixture of chi-square distributions under a high-dimensional asymptotic regime integrated with weak technical conditions. By applying this asymptotic result and estimation theory of the number of factors in a low-dimensional factor model, we propose a new approximation test for vector correlations. We also derive the asymptotic power function for the proposed test. Lastly, we examine the finite sample and dimensional performance of this test using Monte Carlo simulations.

AMS 2000 subject classification: Primary 62H15; secondary 62F05.

Key words:
Tests of covariance structure, high dimension, statistical hypothesis testing, non-normality.

---

1. Introduction

We let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be p-dimensional random sample with a population mean vector $\boldsymbol{\mu}$ and population covariance matrix $\boldsymbol{\Sigma}$. We further partition $\mathbf{x}_i$, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$ into 2 components:

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{x}_{1i} \\ \mathbf{x}_{2i} \end{pmatrix},\ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix},\ \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

where $\mathbf{x}_{gi}$ and $\boldsymbol{\mu}_g$ are $p_g \times 1$ vectors, and $\boldsymbol{\Sigma}_{gh}$ is a $p_g \times p_h$ matrix, $g, h \in \{1, 2\}$. Note that $p = p_1 + p_2$. The test for assessing the vector correlation can be fomulated as

$$\mathcal{H} : \mathbf{\Sigma}_{12} = \mathbf{O} \quad \text{vs.} \quad \mathcal{A} : \mathbf{\Sigma}_{12} \neq \mathbf{O}. \tag{1.1}$$

To construct test (1.1), we introduce the ρV coefficient introduced in [4]. The ρV coefficient is defined as

$$\rho V_{12} = \frac{\|\mathbf{\Sigma}_{12}\|_F^2}{\|\mathbf{\Sigma}_{11}\|_F \|\mathbf{\Sigma}_{22}\|_F},$$

where $\| \cdot \|_F$ denotes the Frobenius norm. The ρV-coefficient measures the correlation between two probability vectors. Particularly , if $p_1 = p_2 = 1$, it corresponds to the square of Pearson's correlation coefficient. Because $\mathbf{\Sigma}_{12} = \mathbf{O}$ and $\rho V_{12} = 0$ are equivalent, the estimator of ρV_{12} can be used to hypothesize testing (1.1). The RV coefficient introduced by [9] can be interpreted as a naive estimator of ρV-coefficient. [10] verifies the accuracy of several permutation tests for (1.1). [10] introduced an RV coefficient based test, assumes a situation where the dimensions are not very large, and compares several permutation tests by simulation. [7] states that the RV coefficient takes high values when the sample size n is small, and when both p_1 and p_2 are large. Further, they corrected the RV coefficient so that it is consistent even in high-dimensional settings, and showed the asymptotic normality of the corrected RV under a high-dimensional framework with a multivariate normal population and the following covariance structure: (hereafter referred to as weak-spike structure).

$$\frac{\|\mathbf{\Sigma}_{gg}^2\|_F^2}{\|\mathbf{\Sigma}_{gg}\|_F^4} = o(1) \quad (p \to \infty). \tag{1.2}$$

A sufficient condition for condition (1.2) is that the largest eigenvalue of $\mathbf{\Sigma}_{gg}$ grows at a rate $o(p_g^{1/4})$ or $O(1)$. [13] proposed an approximate test for deducing the block-diagonal covariance structure of a covariance matrix under a population distribution with relaxed normality assumptions. When the number of blocks is 2, it matches test (1.1). Their test is based on the asymptotic normality of the unbiased estimator of the L_2 squared norm of the off-diagonal block matrix, $\mathbf{\Sigma}_{12}$. This asymptotic normality is justified under the weak spiking condition (1.2). Recentry, [1] also proposed a modified RV coefficient under a population distribution with relaxed normality assumptions, and construct a normal approximation-based test for (1.1). Surprisingly, their study shows the asymptotic normality of the modified RV coefficient under the assumption that $\mathrm{tr}(\mathbf{\Sigma}_{gg})$ grows at rate $O(p_g)$. Note that this assumption holds, even if the largest eigenvalue of $\mathbf{\Sigma}_{gg}$ grows at rate $O(p_g)$. Therefore, they indicate that the asymptotic normality holds under fairly relaxed conditions. However, this study indicates that their results are incorrect, and derive the correct asymptotic distribution of the modified RV coefficient under the condition that some eigenvalues of $\mathbf{\Sigma}_{gg}$ grow at rate $O(p_g)$ (See Theorems 1 and Remark 1). In these studies, normal approximation-based tests for high-dimesnional data are justified under the weak spiking condition (1.2), for example [11], [12], [5], [6].

In other words, there is a concern that the normal approximation-based test for L_2 statistic will not work properly if the condition (1.2) does not hold. This was also noted by [8]. This study provides ρV-based test for (1.1) without the normality assumption and weak spike structure (1.2), while allowing the dimension p to be much larger than the sample size n.

This paper proceeds as follows. Section 2 lays out the high-dimensional asymptotic framework, presents the modified RV coefficients and their asymptotic properties, and provides data-driven test procedures. Section 3 evaluates the finite-sample performance of the proposed test. Lastly, Section 4 presents the conclusions. The Appendix further presents all proofs and auxiliary technical results.

2. Main results

2.1. Data generation model and asymptotic framework

The data generation model is assumed to be a latent factor model expressed as

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{B}\mathbf{f} + \boldsymbol{\epsilon}. \tag{2.1}$$

Here, $\boldsymbol{\mu} \in \mathbb{R}^p$ is the population mean vector $\mathbf{B}$ is the $p \times d$ non-random matrix $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_p)^\top$ that satisfies $\mathrm{rank}(\mathbf{B}) = d$, and elements $\mathbf{b}_1, \ldots, \mathbf{b}_p$ are referred to as factor loadings. $\mathbf{f} \in \mathbb{R}^d$ and $\boldsymbol{\epsilon} \in \mathbb{R}^p$ are random vectors for common and specific factors, respectively. We assume that $\mathbf{f}$ and $\boldsymbol{\epsilon}$ are independent. We let $\mathbf{f} = (\mathrm{f}_1, \ldots, \mathrm{f}_d)$ and $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_p)^\top$. Furthermore, we assume that f_i is iid with $\mathrm{E}(\mathrm{f}_i) = 0$, $\mathrm{E}(\mathrm{f}_i^2) = 1$, and $\mathrm{E}(\mathrm{f}_i^4) = \kappa + 3 < \infty$. and ϵ_j are iid with $\mathrm{E}(\epsilon_j) = 0$, $0 < \mathrm{E}(\epsilon_j^2) = \psi_j < \infty$, $\mathrm{E}(\epsilon_j^4) = \psi_j^2(\kappa + 3) < \infty$ for $i \in \{1, \ldots, d\}$, and $j \in \{1, \ldots, p\}$. Under these assumptions, $\mathrm{E}(\mathbf{f}) = \mathbf{0}$, $\mathrm{E}(\boldsymbol{\epsilon}) = \mathbf{0}$, $\mathrm{cov}(\mathbf{f}) = \mathbf{I}_d$ and $\mathrm{cov}(\boldsymbol{\epsilon}) = \boldsymbol{\Psi} = \mathrm{diag}(\psi_1, \ldots, \psi_p)$.

We further partition $\mathbf{B}$, $\boldsymbol{\Psi}$, and $\boldsymbol{\epsilon}$ into 2 components:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}, \ \boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_1 & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Psi}_2 \end{pmatrix}, \ \boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \end{pmatrix},$$

where $\mathbf{B}_g$ is $p_g \times d$ nonrandom matrix that satisfies $\mathrm{rank}(\mathbf{B}_g) = d_g > 0$, $\boldsymbol{\Psi}_g$ is $p_g \times p_g$ diagonal matrix, and $\boldsymbol{\epsilon}_g$ is p_g-dimensional random vector. These assumptions, along with Equation (3.1), imply that

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \ \boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}^\top + \boldsymbol{\Psi} = \begin{pmatrix} \mathbf{B}_1\mathbf{B}_1^\top + \boldsymbol{\Psi}_1 & \mathbf{B}_1\mathbf{B}_2^\top \\ \mathbf{B}_2\mathbf{B}_1^\top & \mathbf{B}_2\mathbf{B}_2^\top + \boldsymbol{\Psi}_2 \end{pmatrix}.$$

For the asymptotic evaluation, we impose the following regularity conditions:

(A1) $p_g = p_g(n)$ $(g \in \{1, 2\})$ is a function of n such that p_g tends to infinity along with $n \to \infty$, $n/p_g \to \theta_g \in (0, \infty)$, and positive integer d is fixed.

(A2) $\psi_{\max} = \max\{\psi_1, \psi_2, \ldots, \psi_p\}$ is bounded.

(A3) There are two positive semidefinite matrices $\mathbf{B}_{11}^*$ and $\mathbf{B}_{22}^*$ such that $\mathrm{rank}(\mathbf{B}_{11}^*) = d_1 > 0$, $\mathrm{rank}(\mathbf{B}_{22}^*) = d_2 > 0$, and $\|(1/p_g)\mathbf{B}_g^\top \mathbf{B}_g - \mathbf{B}_{gg}^*\|_F \to 0$ $(p_g \to \infty)$ for $g \in \{1,2\}$.

(A4) $\mathbf{f} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$.

Remark 1. *Under (A1)–(A3),* $\mathrm{tr}(\mathbf{\Sigma}_{gg}^2) \asymp p_g^2 \mathrm{tr}(\mathbf{B}_{gg}^{*2}) \asymp p_g^2$ *and* $\mathrm{tr}(\mathbf{\Sigma}_{gg}^4) \asymp p_g^4 \mathrm{tr}(\mathbf{B}_{gg}^{*4}) \asymp p_g^4$. *Therefore, the weak-spike structure (1.2) does not hold. This unique feature distinguishes the weak spike structure in our study from that in other studies (1.2) considered in the test for (1.1), for example [13], [7].*

2.2. Consistent estimator of ρV and its sampling distribution

We let ρV_{12} denote the vector correlation coefficients between the two components $\mathbf{x}$, $\mathbf{x}_1$ and $\mathbf{x}_2$, defined as (see [4])

$$\rho V_{12} = \frac{\|\mathbf{\Sigma}_{12}\|_F^2}{\|\mathbf{\Sigma}_{11}\|_F \|\mathbf{\Sigma}_{22}\|_F}.$$

It is clear that the Pearson's product-moment correlation coefficient is a special case, ρV_{12} when $p = 1$. Furthermore, $\rho V_{12} = \rho V_{21}$, and $\rho V_{12} = 0$ if and only if $\mathbf{\Sigma}_{12} = \mathbf{O}$. Therefore, the natural criterion for testing $\mathcal{H}$ must be based on a suitable statistic for ρV_{12}.

The sample counterpart of ρV_{12} is obtained as

$$RV_{12} = \frac{\|\mathbf{S}_{12}\|_F^2}{\|\mathbf{S}_{11}\|_F \|\mathbf{S}_{22}\|_F},$$

where the sample covariance matrix of $\mathbf{x}_g$ and the cross-sample covariance matrix of $\mathbf{x}_1$ and $\mathbf{x}_2$ are constructed as

$$\forall g \in \{1,2\},\ \mathbf{S}_{gg} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_{gi} - \overline{\mathbf{x}}_g)(\mathbf{x}_{gi} - \overline{\mathbf{x}}_g)^\top,$$

$$\mathbf{S}_{12} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_{1i} - \overline{\mathbf{x}}_1)(\mathbf{x}_{2i} - \overline{\mathbf{x}}_2)^\top,\ \mathbf{S}_{21} = \mathbf{S}_{12}^\top$$

with $\overline{\mathbf{x}}_g = n^{-1}\sum_{i=1}^n \mathbf{x}_{gi}$ for $g \in \{1,2\}$. The invariance of RV_{12} was confirmed by [9]. RV_{12} is a consistent estimator of ρV_{12} when $n \to \infty$ and p are fixed; however, it is not a consistent estimator of ρV_{12} when $n \to \infty$ and $p \to \infty$.

Based on these arguments, the crucial step while constructing the test statistic for testing (1.1) is obtaining an estimator of ρV_{gh} suitable for high-dimensional settings. Note that $\mathbf{S}_{gh}$ is an unbiased estimator of $\mathbf{\Sigma}_{gh}$, but $\|\mathbf{S}_{gh}\|_F^2$ is not an unbiased estimator of $\|\mathbf{\Sigma}_{gh}\|_F^2$. In particular, in high dimensions, $\|\mathbf{S}_{gh}\|_F^2$ has a large bias. Therefore, it is better to use an unbiased estimator of $\|\mathbf{\Sigma}_{gh}\|_F^2$. Therefore, we first consider the following unbiased estimators of $\|\mathbf{\Sigma}_{gh}\|_F^2$.

$$\widehat{\|\mathbf{\Sigma}_{gh}\|_F^2} = \frac{n-1}{n(n-2)(n-3)}[(n-1)(n-2)\mathrm{tr}(\mathbf{S}_{gh}\mathbf{S}_{hg}) + \mathrm{tr}(\mathbf{S}_{gg})\mathrm{tr}(\mathbf{S}_{hh}) - nK_{gh}]$$

for $g, h \in \{1, 2\}$. Here,

$$K_{gh} = \frac{1}{n-1} \sum_{i=1}^{n} \|\mathbf{x}_{gi} - \overline{\mathbf{x}}_g\|^2 \|\mathbf{x}_{hi} - \overline{\mathbf{x}}_h\|^2.$$

$\widehat{\|\boldsymbol{\Sigma}_{12}\|_F^2}$ is an unbiased estimator of $\|\boldsymbol{\Sigma}_{12}\|_F^2$ obtained from [13]. Additionally, $\widehat{\|\boldsymbol{\Sigma}_{gg}\|_F^2}$ ($g \in \{1, 2\}$) is an unbiased estimator of $\|\boldsymbol{\Sigma}_{gg}\|_F^2$ derived by [13]. Using $\widehat{\|\boldsymbol{\Sigma}_{11}\|_F^2}$, $\widehat{\|\boldsymbol{\Sigma}_{22}\|_F^2}$, and $\widehat{\|\boldsymbol{\Sigma}_{12}\|_F^2}$, we define the estimator of ρV_{12} with a high-dimensionality adjustment as

$$MRV_{12} = \frac{\widehat{\|\boldsymbol{\Sigma}_{12}\|_F^2}}{\widehat{\|\boldsymbol{\Sigma}_{11}\|_F}\widehat{\|\boldsymbol{\Sigma}_{22}\|_F}}. \tag{2.2}$$

Estimator (2.2) is the same as that proposed by [1]. Our essential contribution is not to propose an estimator (2.2), rather to clarify the asymptotic properties of this estimator in situations where the weak spike condition (1.2) does not hold. The following theorem lists the asymptotic properties of the proposed modified estimator (2.2).

Theorem 1. *Under (A1)–(A3),* $MRV_{12} = \rho V_{12} + o_p(1)$ *as* $n, p_1, p_2 \to \infty$.

Proof. See Appendix A. □

To construct a hypothesis test (1.1), we consider the null distribution of MRV_{12}. The following theorem provides an asymptotic null distribution of the adjusted MRV_{12} under (A1)–(A4):

Theorem 2. *Suppose the null hypothesis $\mathcal{H}$ in (1.1) is true. Under (A1)–(A4),*

$$nMRV_{12} + \frac{\mathrm{tr}(\boldsymbol{\Lambda}_1)\mathrm{tr}(\boldsymbol{\Lambda}_2)}{\sqrt{\mathrm{tr}(\boldsymbol{\Lambda}_1^2)\mathrm{tr}(\boldsymbol{\Lambda}_2^2)}} \rightsquigarrow \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} \frac{\lambda_{1i}\lambda_{2j}}{\sqrt{\mathrm{tr}(\boldsymbol{\Lambda}_1^2)\mathrm{tr}(\boldsymbol{\Lambda}_2^2)}}\chi_{ij}^2 \quad (n, p_1, p_2 \to \infty), \tag{2.3}$$

where $\chi_{11}^2, \ldots, \chi_{1d_1}^2, \chi_{21}^2, \ldots, \chi_{2d_2}^2$ are mutually independent chi-squared distributed random variables with one degree of freedom, $\boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_{11}, \ldots, \lambda_{1d_1})$ is $d_1 \times d_1$ diagonal matrix whose diagonal components are the nonzero eigenvalues of $\mathbf{B}_{11}^*$, and $\boldsymbol{\Lambda}_2 = \mathrm{diag}(\lambda_{21}, \ldots, \lambda_{2d_2})$ is $d_2 \times d_2$-diagonal matrix whose diagonal components are the nonzero eigenvalues of $\mathbf{B}_{22}^*$.

Proof. See, Appendix B. □

The following remark argues that the sufficient condition $\mathrm{tr}(\boldsymbol{\Sigma}_{gg})/p_g = O(1)$ ($g \in \{1, 2\}$) for the asymptotic normality of MRV_{12} in [1] is incorrect.

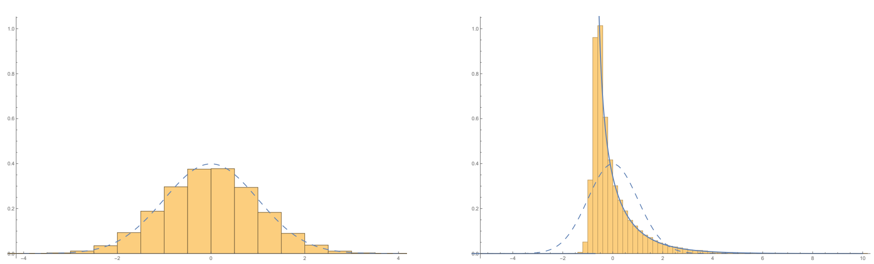
Remark 2. *[1] show the asymptotic normality of $nMRV_{12}/\sqrt{2}$ under $\mathrm{tr}(\boldsymbol{\Sigma}_{gg})/p_g = O(1)$ ($g \in \{1, 2\}$) and other regularity conditions. Note that under Assumptions (A2) and (A3), $\mathrm{tr}(\boldsymbol{\Sigma}_{gg})/p_g = O(1)$ ($g \in \{1, 2\}$) holds. In Theorem 2, if*

$d_1 = d_2 = 1$, *$nMRV_{12} + 1$ converges to a chi-square distribution with one degree of freedom. Therefore, this example is a counterexample in which the asymptotic normality of $nMRV_{12}/\sqrt{2}$ does not hold, even though* $\mathrm{tr}(\mathbf{\Sigma}_{gg})/p_g = O(1)$ *($g \in \{1, 2\}$) holds.*

We further examine the behavior of $nMRV_{12}/\sqrt{2}$ when condition (1.2) holds and when it does not, using a toy example. We set $p_1 = p_2 = 500$, $n = 100$, and $d = 2$. The data generation model is as follows.

$$\mathbf{x} = \begin{pmatrix} \sigma_1 \mathbf{1}_{p_1} & \mathbf{0} \\ \mathbf{0} & \sigma_1 \mathbf{1}_{p_2} \end{pmatrix} \mathbf{f} + \boldsymbol{\epsilon}, \tag{2.4}$$

where $\mathbf{f} \sim \mathcal{N}_2(\mathbf{0}, \mathbf{I}_2)$, $\boldsymbol{\epsilon} \sim \mathcal{N}_p(\mathbf{0}, \sigma^2 \mathbf{I}_p)$, and $\mathbf{f}$ and $\boldsymbol{\epsilon}$ are independet. Then, $\mathbf{\Sigma}_{gg} = \sigma_1^2 \mathbf{1}_{p_g} \mathbf{1}_{p_g}^\top + \sigma^2 \mathbf{I}_p$ and $\mathbf{\Sigma}_{12}$. We consider two cases: (a) $\sigma_1 = 0$ and (b) $\sigma_1 = 1$. Note that (a) is the case when condition (1.2) holds, and (b) is the case when condition (1.2) does not hold. We generated independent pseudorandom observations from model (2.4) and calculated $nMRV_{12}/\sqrt{2}$ 100,000 times. Fig. 1 shows histograms for (a) and (b). $nMRV_{12}/\sqrt{2}$ converges to $\mathcal{N}(0, 1)$ in (a). However, it does not converge in (b). In case (b), as mentioned in the remark, it is almost the same as the density function of $(\chi_1^2 - 1)/\sqrt{2}$. Remark 1 and the numerical results show that the sufficient condition for the asymptotic normality of $nMRV_{12}/\sqrt{2}$ is (1.2) and not $\mathrm{tr}(\mathbf{\Sigma}_{gg})/p_g = O(1)$ ($g \in \{1, 2\}$).



(a) When $\sigma_1 = 0$ **(b)** When $\sigma_2 = 1$

Figure 1: The histograms of $nMRV_{12}/\sqrt{2}$ for the case when condition (1.2) holds (in the left panel) and for the case when condition (1.2) does not hold (in the right panel). The dashed line denotes the density function of $\mathcal{N}(0, 1)$. The solid line denotes the density function of $(\chi_1^2 - 1)/\sqrt{2}$.

2.3. Approximation test and its asymptotic properties

By estimating the unknown parameters in the random variable on the left-hand side of (2.3), we construct a test statistic for (1.1).

First, by applying the idea in [2], we estimate d_g. We focus on the criterion function originally proposed by [2]:

$$ER_g(i) = \frac{\lambda_i(\mathbf{S}_{gg})}{\lambda_{i+1}(\mathbf{S}_{gg})},$$

where $\lambda_i(\cdot)$ is the i-th largest eigenvalue and ER_g is the eigenvalue ratio. The estimator of d_g is given by the number i that minimizes $ER_g(i)$, that is,

$$\widehat{d}_g = \underset{1 \le i \le i_{g,\max}}{\arg\max}\, ER_g(i), \tag{2.5}$$

where $i_{g,\max}$ denotes the prespecified upper bound of i.

We further estimate the unknown parameters $\mathrm{tr}(\mathbf{\Lambda}_g)$ and $\mathrm{tr}(\mathbf{\Lambda}_g^2)$ in (2.3) using

$$\widehat{\mathrm{tr}(\mathbf{\Lambda}_g)} = \sum_{i=1}^{\widehat{d}_g} \widehat{\lambda}_{gi} \quad \text{and} \quad \widehat{\mathrm{tr}(\mathbf{\Lambda}_g^2)} = \sum_{i=1}^{\widehat{d}_g} \widehat{\lambda}_{gi}^2,$$

respectively. Here, $\widehat{\lambda}_{gi} = \lambda_i(\mathbf{S}_{gg})/p_g$ for $i \in \{1, 2, \ldots, \widehat{d}_g\}$ and $g \in \{1, 2\}$.

Using these estimators, we propose a test statistic, defined as

$$T = nMRV_{12} + \frac{\widehat{\mathrm{tr}(\mathbf{\Lambda}_1)}\widehat{\mathrm{tr}(\mathbf{\Lambda}_2)}}{\sqrt{\widehat{\mathrm{tr}(\mathbf{\Lambda}_1^2)}\widehat{\mathrm{tr}(\mathbf{\Lambda}_2^2)}}}.$$

The following theorem further shows that T has the same limiting null distribution as the random variable on the left side of (2.3).

Theorem 3. *Suppose the null hypothesis $\mathcal{H}$ in (1.1) is true. In (A1)–(A4),*

$$T \rightsquigarrow \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \frac{\lambda_{1i}\lambda_{2j}}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}_1^2)\mathrm{tr}(\mathbf{\Lambda}_2^2)}} \chi_{ij}^2 \quad (n, p_1, p_2 \to \infty).$$

Proof. See, Appendix C. □

Based on the results of Theorem 3, we provide an approximate test for (1.1). The test involves four steps.

1. We draw n observations from the population and calculate $\widehat{d}_g$, $\widehat{\lambda}_{gi}$ for $i \in \{1, \ldots, \widehat{d}_g\}$, $\widehat{\mathrm{tr}(\mathbf{\Lambda}_g)}$, and $\widehat{\mathrm{tr}(\mathbf{\Lambda}_g^2)}$ for $g \in \{1, 2\}$. Using these estimators, we construct T.
2. We further draw a sample of $\widehat{d}_1 \times \widehat{d}_2$ independently and χ_{ij}^2-distributed random variables to obtain

$$\widetilde{T} = \sum_{i=1}^{\widehat{d}_1} \sum_{j=1}^{\widehat{d}_2} \frac{\widehat{\lambda}_{1i}\widehat{\lambda}_{2j}}{\sqrt{\widehat{\mathrm{tr}(\mathbf{\Lambda}_1^2)}\widehat{\mathrm{tr}(\mathbf{\Lambda}_2^2)}}} \chi_{ij}^2.$$

3. We then repeat step 2 until we obtain a Monte Carlo estimate of the distribution for the random variable $\widetilde{T}$ and its $(1-\alpha)$-quantile $\hat{t}_\alpha$.
4. We further realized an approximate test with the nominal size α as follows:

$$\text{Reject } \mathcal{H} \overset{\text{def}}{\iff} T > \widehat{t}_\alpha. \tag{2.6}$$

To examine the power of test (2.6), we consider the following fixed alternative:

$\mathcal{A}_F$: $\|\mathbf{\Sigma}_{12}\|_F^2/(p_1p_2)$ convergence to some $\delta \in (0, \|\mathbf{B}_{11}^*\|_F\|\mathbf{B}_{22}^*\|_F)$ as $p_1, p_2 \to \infty$.

Under the fixed alternatives $\mathcal{A}_F$, (A2), and (A3), ρV_{12} converges to $\rho V_{12}^* = \delta/(\|\mathbf{B}_{11}^*\|_F\|\mathbf{B}_{11}^*\|_F) \in (0,1)$ as $p_1, p_2 \to \infty$. The following theorem provides the asymptotic distribution of the MRV coefficient under the fixed alternative hypothesis:

Theorem 4. *Under fixed alternatives $\mathcal{A}_F$ and (A1)–(A4),*

$$\frac{\sqrt{n}(MRV_{12} - \rho V_{12})}{\sigma} \rightsquigarrow \mathcal{N}(0,1) \ (n, p_1, p_2 \to \infty),$$

where

$$\begin{aligned}\sigma^2 =& 2\rho V_{12}^2 \left[\frac{p_1^4\|\mathbf{B}_{11}^{*2}\|_F^2}{\|\mathbf{\Sigma}_{11}\|_F^4} + \frac{p_2^4\|\mathbf{B}_{22}^{*2}\|_F^2}{\|\mathbf{\Sigma}_{22}\|_F^4} + 2p_1^2p_2^2 \frac{\mathrm{tr}\{(\mathbf{B}_{11}^*\mathbf{B}_{22}^*)^2\} + \mathrm{tr}(\mathbf{B}_{11}^{2*}\mathbf{B}_{22}^{2*})}{\|\mathbf{\Sigma}_{12}\|_F^4} \right. \\ &\left. + \frac{2p_1^2p_2^2\mathrm{tr}(\mathbf{B}_{11}^{*2}\mathbf{B}_{22}^{*2})}{\|\mathbf{\Sigma}_{11}\|_F^2\|\mathbf{\Sigma}_{22}\|_F^2} - \frac{4p_1^3p_2\mathrm{tr}(\mathbf{B}_{11}^{*3}\mathbf{B}_{22}^*)}{\|\mathbf{\Sigma}_{11}\|_F^2\|\mathbf{\Sigma}_{12}\|_F^2} - \frac{4p_1p_2^3\mathrm{tr}(\mathbf{B}_{22}^{*3}\mathbf{B}_{11}^*)}{\|\mathbf{\Sigma}_{22}\|_F^2\|\mathbf{\Sigma}_{21}\|_F^2} \right]. \end{aligned} \tag{2.7}$$

Proof. See, Appendix D. □

Applying this theorem, we obtain the following corollary of the asymptotic power under the fixed alternative $\mathcal{A}_F$.

Corollary 1. *Under fixed alternatives $\mathcal{A}_F$ and (A1)–(A4), the asymptotic power function converges to* 1 *as $n, p_1, p_2 \to \infty$.*

Proof. See, Appendix E. □

Note that the asymptotic power function is equal to one for all values under a fixed alternative $\mathcal{A}_F$. This is the notion of consistency for a test: it has asymptotic power 1 under every fixed alternative $\mathcal{A}_F$. Accordingly, the asymptotic power vs. fixed alternative $\mathcal{A}_F$ is not a sufficiently discerning asymptotic criterion for distinguishing between the tests. This problem can be addressed by considering the following local alternatives:

$\mathcal{A}_L$: Let η be a constant greater than or equal to 1/2. There exists a $d \times d$ matrix $\mathbf{\Xi}$ such that all diagonal elements are 0 and at least one off-diagonal element is not 0 such that the following condition is met:

$$\left\| \frac{n^\eta}{p_1p_2} \mathbf{B}_1^\top \mathbf{B}_1 \mathbf{B}_2^\top \mathbf{B}_2 - \mathbf{\Xi} \right\|_F \to 0 \ \ (n, p_1, p_2 \to \infty).$$

Furthermore, there exists a positive real number Δ such that the following condition is met:

$$\frac{n^{2\eta}}{p_1p_2}\|\mathbf{\Sigma}_{12}\|_F^2 = \frac{n^{2\eta}}{p_1p_2}\mathrm{tr}(\mathbf{B}_1^\top \mathbf{B}_1 \mathbf{B}_2^\top \mathbf{B}_2) \to \Delta \ \ (n, p_1, p_2 \to \infty).$$

Note that $\rho V_{12} \asymp n^{-2\eta}$ under (A1)–(A3) and $\mathcal{A}_L$.

Theorem 5. *Under the local alternatives $\mathcal{A}_L$ and (A1)–(A4),*

$$nMRV_{12} + \frac{\mathrm{tr}(\boldsymbol{\Lambda}_1)\mathrm{tr}(\boldsymbol{\Lambda}_2)}{\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F} \rightsquigarrow \begin{cases} \Delta/(\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F) + \mathbf{z}^\top\mathbf{C}^*\mathbf{z} + \mathbf{c}^{*\top}\mathbf{z} & \eta = 1/2, \\ \mathbf{z}^\top\mathbf{C}^*\mathbf{z} & \eta > 1/2. \end{cases}$$

where $\mathbf{z}$ has a d^2-variate normal distribution with a mean vector $\mathbf{0}$ and covariance matrix $\mathbf{I}_{d^2} + \mathbf{K}_{d^2}$ and

$$\mathbf{C}^* = \frac{1}{\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F}(\mathbf{B}^*_{11} \otimes \mathbf{B}^*_{22}),\ \mathbf{c}^* = \frac{1}{\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F}\mathrm{vec}(\boldsymbol{\Xi} + \boldsymbol{\Xi}^\top).$$

Here, $\mathbf{K}_{d^2}$ denotes the commutation matrix.

Proof. See Appendix F. □

Applying the theorem, we obtain the following corollary of the asymptotic power under local alternative (L).

Corollary 2. *Under (A1)–(A4), the asymptotic power function is*

$$\Pr(T > \hat{t}_\alpha | \mathcal{A}_L) = \begin{cases} G\{t_\alpha - \Delta/(\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F)\} + o(1) & \eta = 1/2, \\ \alpha + o(1) & \eta > 1/2, \end{cases}$$

where $G(\cdot)$ denotes the cumulative distribution function of $\mathbf{z}^\top\mathbf{C}^*\mathbf{z} + \mathbf{c}^{*\top}\mathbf{z}$.

Proof. See Appendix G. □

3. A simulation study

In this section, we examine the size and power of test (2.6) in a finite sample. Throughout this section, the sample size n is $n \in \{p/2, p, 2p\}$, the dimension p_1 and p_2 are $(p_1, p_2) = (100\ell, 150\ell)$ for $\ell \in \{1, 2, 4, 8\}$, the number of factors d is $d \in \{2, 4\}$, and the nominal size α is $\alpha = 0.05$. We assume that the data $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$ follow the model

$$x_{ij} = \mathbf{b}_j^\top \mathbf{f}_i + \epsilon_{ij} \quad i \in \{1, \dots, n\},\ j \in \{1, \dots, p\}. \tag{3.1}$$

The specific factors were drawn from the following three distribution scenarios:.

(D1): $\epsilon_{ij} \overset{iid}{\sim} \mathcal{N}(0, 1)$.

(D2): Let $e_{ij} \overset{iid}{\sim} \chi^2_1$, then $\epsilon_{ij} = (e_{ij} - 1)/\sqrt{2}$.

(D3): Let $e_{ij} \overset{iid}{\sim} \mathcal{LN}(0, 1)$, then $\epsilon_{ij} = (e_{ij} - e^{1/2})/\sqrt{e(e-1)}$.

The common factors $\mathbf{f}_i \overset{iid}{\sim} \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$.

For each fixed p_1, p_2, and n, we repeat the following steps 10,000 times, and calculate the empirical size or power of the proposed test (2.6) by recording the number of times the null hypothesis $\mathcal{H}$ is rejected.

1. Generate independently $\{\mathbf{b}_j\}_{j=1}^p$ and set $\mathbf{B}$.
2. Generate independently $\{\boldsymbol{\epsilon}_i\}_{i=1}^n$.
3. Generate independently $\{\mathbf{f}_i\}_{i=1}^n$.
4. Calculate $\{\mathbf{x}_i\}_{i=1}^n$ from $\mathbf{x}_i = \mathbf{B}\mathbf{f}_i + \boldsymbol{\epsilon}_i$.
5. Set prespecified upper bound in (2.5) to 10, estimate $\widehat{d}_1$ and $\widehat{d}_2$. Execute test steps 1 to 4 in Section 2.3 and record 1 if the null hypothesis $\mathcal{H}$ is rejected.

3.1. Size comparison with previous research

To ascertain the size of the test (2.6), we set the factor loadings $\mathbf{b}_j = (b_{j1}, \ldots, b_{jd})^\top$ as follows for each $d = 2$ and $d = 4$ so that the null hypothesis $\mathcal{H}$ holds.

- $d = 2$:For $j \in \{1, \ldots, p_1\}$, $b_{j1} \overset{iid}{\sim} \mathcal{N}(1, 1)$, $b_{j2} = 0$. For $j \in \{p_1 + 1, \ldots, p\}$, $b_{j1} = 0$, $b_{j2} \overset{iid}{\sim} \mathcal{N}(1/2, 1)$.
- $d = 4$:For $j \in \{1, \ldots, p_1\}$, $b_{j1}, b_{j2} \overset{iid}{\sim} \mathcal{N}(1, 1)$, $b_{j3} = b_{j4} = 0$. For $j \in \{p_1 + 1, \ldots, p\}$, $b_{j1} = b_{j2} = 0$, $b_{j3}, b_{j4} \overset{iid}{\sim} \mathcal{N}(1/2, 1)$.

We also compared the proposed test (2.6) with the test procedures in [3] and [1] in terms of size control. These tests are denoted by Cor, and Ahm, respectively, and our proposed test is denoted by HNN throughout this section.

The covariance structure in this simulation setting satisfies assumptions (A2) and (A3). Therefore, Theorem 3 shows that the size of our test converges to the nominal size for large p and n. Tables 1 and 2 indicate that our proposed test HNN provides a valid asymptotic test with an accurate size control for most simulation settings. The test sizes for HNN are close to the nominal levels even for $(p_1, p_2, n) = (100, 150, 125)$ in Table 1, indicating that the asymptotic condition described by Theorem 3 occurs even with relatively small dimensions. Furthermore, HNN systematically outperforms both Cor and Ahm in terms of size control across different distributions. However , the sizes of Cor and Ahm are larger than the nominal size in all cases considered in this simulation. Cor is theoretically valid in large-sample settings; therefore, it is likely that it does not work in high dimensions. Additionally, as mentioned in Remark 2, Ahm has no theoretical validity under the covariance structures in this simulation settings. Therefore, the approximation accuracy may deteriorate. As expected, Ahm exhibits an inflated size in almost all simulation settings. Additionally, the size does not converge to the nominal level because both p and n_g increase. This is likely because the asymptotic normality of the test statistic does not hold under the covariance structures in the simulation settings.

Table 1: This table presents the case where the number of factors d is $d = 2$. The sizes of the proposed tests (HNN), approximate permutation test by [3] (Cor), and tests based on normal approximation by [1] (Ahm) are evaluated under three different distribution scenarios for specific factors at a significance level of 5%. The average size in each column is listed in the "mean" row, and the standard deviation of size in each column is listed in the "sd" row.

		$n = p/2$			$n = p$			$n = 2p$		
(p_1, p_2)	ϵ_{ij}	HNN	Ahm	Cor	HNN	Ahm	Cor	HNN	Ahm	Cor
$(100, 150)$	(D1)	4.88	6.81	6.67	5.44	7.43	7.25	5.10	6.68	6.24
	(D2)	5.36	7.11	7.08	5.17	6.86	6.70	5.02	6.72	6.21
	(D3)	5.17	6.94	6.79	5.04	6.82	6.69	4.88	6.77	6.24
$(200, 300)$	(D1)	5.02	6.83	6.78	4.90	6.80	6.79	5.25	6.91	6.37
	(D2)	5.18	6.79	6.87	4.78	6.41	6.26	5.33	7.30	6.75
	(D3)	5.35	7.37	7.34	4.76	6.80	6.70	4.93	7.10	6.49
$(400, 600)$	(D1)	4.98	6.85	6.71	5.40	7.15	7.19	4.91	6.87	6.66
	(D2)	5.14	6.88	6.81	5.17	6.99	6.91	5.09	6.88	6.69
	(D3)	4.87	6.96	6.83	5.28	6.83	6.78	5.27	7.12	6.88
$(800, 1200)$	(D1)	4.87	6.61	6.58	4.97	6.84	6.81	4.74	6.37	6.23
	(D2)	5.19	6.87	6.89	4.77	6.60	6.60	5.26	7.07	6.91
	(D3)	4.93	7.04	7.08	4.93	6.52	6.56	5.01	6.77	6.62
mean		5.08	6.92	6.87	5.05	6.84	6.77	5.07	6.88	6.52
sd		0.17	0.18	0.20	0.23	0.26	0.25	0.18	0.24	0.25

Table 2: This table presents the case where the number of factors d is $d = 4$. The sizes of the proposed tests (HNN), the approximate permutation test by [3] (Cor), and tests based on normal approximation by [1] (Ahm) are evaluated under three different distribution scenarios for specific factors at a significance level of 5%. The average size in each column is listed in the "mean" row, and the standard deviation of size in each column is listed in the "sd" row.

		$n = p/2$			$n = p$			$n = 2p$		
(p_1, p_2)	ϵ_{ij}	HNN	Ahm	Cor	HNN	Ahm	Cor	HNN	Ahm	Cor
$(100, 150)$	(D1)	4.36	6.06	6.15	5.10	7.16	6.94	5.18	7.07	6.20
	(D2)	5.14	6.98	6.99	5.48	7.23	7.07	4.81	6.60	5.79
	(D3)	5.46	7.15	7.04	5.07	6.48	6.31	4.93	6.60	5.77
$(200, 300)$	(D1)	5.20	7.29	7.16	4.66	6.51	6.37	4.92	7.04	6.37
	(D2)	5.06	7.09	7.03	5.06	6.79	6.67	5.34	7.24	6.68
	(D3)	4.69	6.51	6.44	5.19	7.16	7.03	5.18	6.96	6.29
$(400, 600)$	(D1)	5.08	7.10	7.10	4.72	6.68	6.62	4.88	6.60	6.35
	(D2)	5.21	7.43	7.34	5.22	7.08	7.06	4.67	6.52	6.33
	(D3)	5.02	6.89	6.82	5.07	7.19	7.16	4.87	6.88	6.51
$(800, 1200)$	(D1)	5.35	7.33	7.27	4.74	6.86	6.86	4.87	6.67	6.53
	(D2)	4.74	6.70	6.63	5.13	7.05	6.97	4.83	6.77	6.65
	(D3)	4.95	6.75	6.83	5.47	7.39	7.32	5.31	6.98	6.98
mean		5.02	6.94	6.90	5.08	6.97	6.87	4.98	6.83	6.37
sd		0.29	0.37	0.33	0.25	0.28	0.30	0.21	0.22	0.33

3.2. Accuracy of approximate power function

Because the discrepancies in the size and nominal size for Cor and Ahm are large, we consider only the power of the proposed test in this section. We verify the finite-sample accuracy of the asymptotic power function when $\eta = 1/2$ in Corollary 2. To investigate the power of test (2.6), we set the factor loadings $\mathbf{b}_j = (b_{j1}, \ldots, b_{jd})^\top$ as follows for each $d = 2$ and $d = 4$ so that the local alternative $\mathcal{A}_L$ $(\eta = 1/2)$ holds.

- $d = 2$:For $j \in \{1, \ldots, p_1\}$, $b_{j1} \overset{iid}{\sim} \mathcal{N}(1, 1)$, $b_{j2} = 3n^{-1/2}$. For $j \in \{p_1 +$

$1,\ldots,p\}$, $b_{j1}=0$, $b_{j2} \overset{iid}{\sim} \mathcal{N}(1/2,1)$.

- $d=4$:For $j \in \{1,\ldots,p_1\}$, $b_{j1}, b_{j2} \overset{iid}{\sim} \mathcal{N}(1,1)$, $b_{j3}=b_{j4}=3n^{-1/2}$. For $j \in \{p_1+1,\ldots,p\}$, $b_{j1}=b_{j2}=0$, $b_{j3}, b_{j4} \overset{iid}{\sim} \mathcal{N}(1/2,1)$.

The following steps are followed to calculate the approximate power based on the results in Corollary 2:

1. For $\eta=1/2$, we caluclate $\widetilde{\mathbf{B}}_{gg}^{*}=(1/p_g)\mathbf{B}_g^{\top}\mathbf{B}_g$, $\widetilde{\lambda}_{gi}=\lambda_i(\widetilde{\mathbf{B}}_{gg}^{*})$, $\widetilde{\mathbf{\Lambda}}_g=\mathrm{diag}(\widetilde{\lambda}_{g1},\ldots,\widetilde{\lambda}_{gd_g})$, $\widetilde{\Delta}=n^{2\eta}\mathrm{tr}(\widetilde{\mathbf{B}}_{11}^{*}\widetilde{\mathbf{B}}_{22}^{*})$, and $\widetilde{\mathbf{\Xi}}=n^{\eta}\widetilde{\mathbf{B}}_{11}^{*}\widetilde{\mathbf{B}}_{22}^{*}$.
2. We draw a sample of $d_1 \times d_2$ independently and χ^2_{ij}-distributed random variables to obtain
$$T_{\mathcal{H}}=\sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\frac{\widetilde{\lambda}_{1i}\widetilde{\lambda}_{2j}}{\sqrt{\mathrm{tr}(\widetilde{\mathbf{\Lambda}}_1^2)\mathrm{tr}(\widetilde{\mathbf{\Lambda}}_2^2)}}\chi^2_{ij}.$$
We repeat this operation until we obtain a Monte Carlo estimate of the distribution for the random variable $T_{\mathcal{H}}$ and its $(1-\alpha)$-quantile $\widetilde{t}_\alpha$.
3. We draw a sample of $\mathbf{z}$ from $\mathcal{N}_{d^2}(\mathbf{0},\mathbf{K}_{d^2}+\mathbf{I}_{d^2})$ and caluclate
$$T_{\mathcal{A}}=\mathbf{z}^{\top}\widetilde{\mathbf{C}}^{*}\mathbf{z}+\widetilde{\mathbf{c}}^{*\top}\mathbf{z},$$
where
$$\widetilde{\mathbf{C}}^{*}=\frac{1}{\|\widetilde{\mathbf{\Lambda}}_1\|_F\|\widetilde{\mathbf{\Lambda}}_2\|_F}(\widetilde{\mathbf{B}}_{11}^{*}\otimes\widetilde{\mathbf{B}}_{22}^{*}),\ \widetilde{\mathbf{c}}^{*}=\frac{1}{\|\widetilde{\mathbf{\Lambda}}_1\|_F\|\widetilde{\mathbf{\Lambda}}_2\|_F}\mathrm{vec}(\widetilde{\mathbf{\Xi}}+\widetilde{\mathbf{\Xi}}^{\top}).$$
We repeat this operation until we obtain a Monte Carlo estimate $\widetilde{G}(\cdot)$ of the distribution for the random variable $T_{\mathcal{A}}$. By using $\widetilde{G}(\cdot)$, we obtain
$$\Pr(T>\widehat{t}_\alpha|\mathcal{A}_L)\approx\widetilde{G}\left(\widetilde{t}_\alpha-\frac{\widetilde{\Delta}}{\|\widetilde{\mathbf{\Lambda}}_1\|_F\|\widetilde{\mathbf{\Lambda}}_2\|_F}\right).$$

In Tables 3 and 4, the approximate and empirical powers calculated by the simulation are denoted by APW and EPW, respectively. Tables 3 and 4 verify that the asymptotic power function given by Corollary 2 is a good approximation of power in high-dimensional settings. Tables 3 and 4 indicate that APW is very close to EPW for most simulation settings.

4. Conclusion

This study presents a test statistic for vector correlation for high-dimensional data using a low-dimensional factor model. We formulate the test statistic as a consistent estimator of ρV coefficient. We further employ the corresponding asymptotic theory to derive the null and non-null limiting distributions of the proposed test when both the sample size and dimensions approach infinity. The

Table 3: This table presents the case where the number of factors d is $d = 2$. The power of the proposed tests (EPW) and the approximate power (APW) are evaluated under three different distribution scenarios for specific factors at a significance level of 5%. The average power in each column is listed in the "mean" row, and the standard deviation of power in each column is listed in the "sd" row.

		$n = p/2$		$n = p$		$n = 2p$	
(p_1, p_2)	ϵ_{ij}	EPW	APW	EPW	APW	EPW	APW
(100, 150)	(D1)	56.88	58.75	55.74	55.68	56.57	55.41
	(D2)	57.59	56.40	56.53	56.39	56.67	57.08
	(D3)	57.29	55.82	57.09	55.31	56.64	56.15
(200, 300)	(D1)	67.94	67.21	67.66	67.96	68.06	67.54
	(D2)	66.67	66.84	67.40	67.44	67.27	66.27
	(D3)	68.03	67.69	67.95	66.04	67.03	66.61
(400, 600)	(D1)	57.73	57.34	57.41	57.49	58.05	58.99
	(D2)	57.09	57.33	58.27	57.34	57.65	58.30
	(D3)	57.19	57.80	57.41	57.20	57.25	58.06
(800, 1200)	(D1)	58.47	60.59	59.01	58.57	58.81	57.29
	(D2)	58.88	59.75	59.46	59.85	59.34	60.87
	(D3)	58.74	58.09	58.82	57.64	57.92	57.81
mean		60.21	60.30	60.23	59.74	60.11	60.03
sd		4.29	4.20	4.41	4.44	4.32	4.13

Table 4: This table presents the case where the number of factors d is $d = 4$. The power of the proposed tests (HNN) and the approximate power (Approx) are evaluated under three different distribution scenarios for specific factors at a significance level of 5%. The average power in each column is listed in the "mean" row, and the standard deviation of power in each column is listed in the "sd" row.

		$n = p/2$		$n = p$		$n = 2p$	
(p_1, p_2)	ϵ_{ij}	EPW	APW	EPW	APW	EPW	APW
(100, 150)	(D1)	62.87	63.98	64.36	63.13	65.35	65.01
	(D2)	63.82	62.89	64.16	65.40	64.00	65.50
	(D3)	63.38	62.91	64.18	63.23	65.44	64.59
(200, 300)	(D1)	62.65	62.65	63.48	63.36	62.55	65.56
	(D2)	62.93	62.25	63.37	64.00	63.09	62.13
	(D3)	62.79	62.26	63.96	63.66	62.56	64.57
(400, 600)	(D1)	59.23	58.45	58.11	59.72	59.53	58.21
	(D2)	58.48	58.48	59.04	58.42	59.22	59.86
	(D3)	58.64	60.45	58.90	59.75	59.02	59.60
(800, 1200)	(D1)	67.61	66.62	67.77	67.72	67.02	68.03
	(D2)	66.78	67.83	67.94	67.52	67.80	66.85
	(D3)	68.12	65.46	66.57	66.03	68.42	69.10
mean		63.11	62.85	63.49	63.50	63.67	64.08
sd		3.11	2.77	3.15	2.87	3.14	3.29

asymptotic theory of the test was developed under a few mild assumptions, and accommodates a wide class of highly spiked, high-dimensional covariance models for the population, which usually represent a factor structure.

We propose a chi-square mixture-type of asymptotic approximations of the test statistic, along with a Monte Carlo simulation scheme, to compute the critical values of the test. This approach is in contrast to that of [1], in which the asymptotic theory of the proposed tests centers around the normal limits of the null distribution. Both our theoretical findings and numerical studies justify that the proposed construction of the ρV coefficient–based test statistic, and its chi-squared mixture approximation allow for better test size control. Therefore, our proposed approach is more suitable for high-dimensional models with a latent factor structure than the test proposed by [1], which does not consider the structural aspects of the distribution of the underlying population.

Consequently, the proposed testing method may be appropriate for high-dimensional, vector correlation tests when the data have an unknown degree with an underlying low-dimensional latent factor structure. The test performs very well for many practical distributions of the factor vector and error term of (3.1) and spiked distributions, where the dimensions may greatly exceed the sample size, and even for a moderate number of independent samples.

Appendix

A. Proof of Theorem 1

Let $\mathbf{y}_{gi} = \mathbf{x}_{gi} - \boldsymbol{\mu}_g$ for $g \in \{1, 2\}$. Note that

$$\mathbf{y}_i = \begin{pmatrix} \mathbf{y}_{1i} \\ \mathbf{y}_{2i} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1 & \boldsymbol{\Psi}_1^{1/2} & \mathbf{O} \\ \mathbf{B}_2 & \mathbf{O} & \boldsymbol{\Psi}_2^{1/2} \end{pmatrix} \begin{pmatrix} \mathbf{f}_i \\ \mathbf{e}_{1i} \\ \mathbf{e}_{2i} \end{pmatrix},$$

where $\mathbf{e}_{gi} = \boldsymbol{\Psi}_g^{-1/2} \boldsymbol{\epsilon}_{gi}$ for $i \in \{1, \ldots, n\}$, $g \in \{1, 2\}$.

First, we evaluate $\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2$, denoted as

$$\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2 = J_1 - 2J_2 + J_3,$$

where

$$\begin{aligned} J_1 &= \frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hi}^\top \mathbf{y}_{hj}, \\ J_2 &= \frac{2}{n(n-1)(n-2)} \sum_{\substack{i,j,k=1 \\ i \neq j, j \neq k, k \neq i}}^{n} \mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hi}^\top \mathbf{y}_{hk}, \\ J_3 &= \frac{1}{n(n-1)(n-2)(n-3)} \sum_{\substack{i,j,k,\ell=1 \\ i \neq j \neq k \neq \ell, k \neq i \neq \ell \neq j}}^{n} \mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hk}^\top \mathbf{y}_{h\ell}. \end{aligned}$$

Because the expectation value of J_1 is $\|\boldsymbol{\Sigma}_{gh}\|_F^2$ and the expectation values of J_2 and J_3 are 0, the unbiasedness of $\|\widehat{\boldsymbol{\Sigma}_{gh}}\|_F^2$ is verified.

$\mathrm{var}(J_1)$ is calculated as

$$\begin{aligned}
\mathrm{var}(J_1) =& \frac{2}{n^2(n-1)^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \mathrm{E}\{(\mathbf{y}_{gi}^\top \mathbf{y}_{gj})^2 (\mathbf{y}_{hi}^\top \mathbf{y}_{hj})^2\} \\
&+ \frac{4}{n^2(n-1)^2} \sum_{\substack{i,j,k=1 \\ i \neq j, j \neq k, k \neq i}}^{n} \mathrm{E}(\mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hi}^\top \mathbf{y}_{hj} \mathbf{y}_{gi}^\top \mathbf{y}_{gk} \mathbf{y}_{hi}^\top \mathbf{y}_{hk}) \\
&+ \frac{1}{n^2(n-1)^2} \sum_{\substack{i,j,k,\ell=1 \\ i \neq j \neq k \neq \ell, k \neq i \neq \ell \neq j}}^{n} \mathrm{E}(\mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hi}^\top \mathbf{y}_{hj} \mathbf{y}_{gk}^\top \mathbf{y}_{g\ell} \mathbf{y}_{hk}^\top \mathbf{y}_{h\ell}) \\
=& \frac{2}{n(n-1)} \mathrm{tr}(\mathbf{A}_1^2) \mathrm{tr}(\mathbf{A}_2^2) + \frac{2}{n(n-1)} \{\mathrm{tr}(\mathbf{A}_1 \mathbf{A}_2)\}^2 \\
&+ \frac{4}{n} \mathrm{tr}\{(\mathbf{A}_1 \mathbf{A}_2)^2\} + \frac{4}{n} \mathrm{tr}(\mathbf{A}_1^2 \mathbf{A}_2^2) \\
&+ \frac{\kappa}{n} \mathrm{tr}\{(\mathbf{A}_1 \mathbf{A}_2) \odot (\mathbf{A}_1 \mathbf{A}_2)\} + \frac{\kappa}{n} \mathrm{tr}\{(\mathbf{A}_2 \mathbf{A}_1) \odot (\mathbf{A}_2 \mathbf{A}_1)\} \\
&+ \frac{2\kappa}{n} \mathrm{tr}\{(\mathbf{A}_1 \mathbf{A}_2) \odot (\mathbf{A}_2 \mathbf{A}_1)\} + \frac{2\kappa}{n(n-1)} \mathrm{tr}\{(\mathbf{A}_1 \odot \mathbf{A}_2)^2\} \\
&+ \frac{4\kappa}{n(n-1)} \mathrm{tr}(\mathbf{A}_1^2 \odot \mathbf{A}_2^2) + \frac{4\kappa}{n(n-1)} \mathrm{tr}\{(\mathbf{A}_1 \mathbf{A}_2) \odot (\mathbf{A}_2 \mathbf{A}_1)\} \\
&+ \frac{2}{n(n-1)} \mathrm{tr}\{(\mathbf{A}_1 \mathbf{A}_2)^2\} + \frac{2}{n(n-1)} \mathrm{tr}(\mathbf{A}_1^2 \mathbf{A}_2^2),
\end{aligned}$$

where

$$\mathbf{A}_1 = \begin{pmatrix} \mathbf{B}_1^\top \mathbf{B}_1 & \mathbf{B}_1^\top \mathbf{\Psi}_1^{1/2} & \mathbf{O} \\ \mathbf{\Psi}_1^{1/2} \mathbf{B}_1 & \mathbf{\Psi}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{pmatrix}, \ \mathbf{A}_2 = \begin{pmatrix} \mathbf{B}_2^\top \mathbf{B}_2 & \mathbf{O} & \mathbf{B}_2^\top \mathbf{\Psi}_2^{1/2} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{\Psi}_2^{1/2} \mathbf{B}_2 & \mathbf{O} & \mathbf{\Psi}_2 \end{pmatrix}.$$

Therefore, under (A1)–(A3),

$$\frac{1}{n^2(n-1)^2} \mathrm{var}\left(\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hi}^\top \mathbf{y}_{hj} \right) = O(p^4/n).$$

Under Assumptions (A1) and (A2), the variance of each term is:

$$\frac{4}{n^2(n-1)^2(n-2)^2} \mathrm{var}\left(\sum_{\substack{i,j,k=1 \\ i \neq j, j \neq k, k \neq i}}^{n} \mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hi}^\top \mathbf{y}_{hk} \right) = O(p^4/n^3),$$

$$\frac{1}{n^2(n-1)^2(n-2)^2(n-3)^2} \mathrm{var}\left(\sum_{\substack{i,j,k,\ell=1 \\ i \neq j \neq k \neq \ell, k \neq i \neq \ell \neq j}}^{n} \mathbf{y}_{gi}^\top \mathbf{y}_{gj} \mathbf{y}_{hk}^\top \mathbf{y}_{h\ell} \right) = O(p^4/n^4).$$

Therefore, for $g, h \in \{1,2\}$,

$$\frac{\widehat{\|\boldsymbol{\Sigma}_{gh}\|_F^2}}{p_g p_h} = \frac{\|\boldsymbol{\Sigma}_{gh}\|_F^2}{p_g p_h} + O_p(n^{-1/2}). \tag{A.1}$$

From (A.1), $\|\boldsymbol{\Sigma}_{12}\|_F = O(\sqrt{p_1 p_2})$ and $\|\boldsymbol{\Sigma}_{gg}\|_F \asymp p_g$ $(g \in \{1,2\})$ under (A1) and (A2),

$$\begin{aligned}
HRV_{12} =& \frac{\widehat{\|\boldsymbol{\Sigma}_{12}\|_F^2}/(p_1p_2)}{\sqrt{\widehat{\|\boldsymbol{\Sigma}_{11}\|_F^2}/p_1^2}\sqrt{\widehat{\|\boldsymbol{\Sigma}_{22}\|_F^2}/p_2^2}} \\
=& \frac{\|\boldsymbol{\Sigma}_{12}\|_F^2/(p_1p_2) + O_p(n^{-1/2})}{\{\|\boldsymbol{\Sigma}_{11}\|_F/p_1 + O_p(n^{-1/2})\}\{\|\boldsymbol{\Sigma}_{22}\|_F/p_2 + O_p(n^{-1/2})\}} \\
=& \rho V_{12} + o_p(1).
\end{aligned}$$

B. Proof of Theorem 2

First, we derive stochastic asymptotic expansion of $n\widehat{\|\boldsymbol{\Sigma}_{12}\|_F^2}/(p_1p_2)$.

$$\begin{aligned}
\frac{n}{p_1p_2}\widehat{\|\boldsymbol{\Sigma}_{12}\|_F^2} =& \frac{1}{np_1p_2}\sum_{\substack{i,j=1\\ i\neq j}}^{n} \mathbf{f}_i^\top \mathbf{B}_1^\top \mathbf{B}_1 \mathbf{f}_j \mathbf{f}_i^\top \mathbf{B}_2^\top \mathbf{B}_2 \mathbf{f}_j + o_p(1) \\
=& \frac{1}{n}\mathrm{tr}\left(\sum_{\substack{i,j=1\\ i\neq j}}^{n} \mathbf{f}_i \mathbf{f}_i^\top \mathbf{B}_{11}^* \mathbf{f}_j \mathbf{f}_j^\top \mathbf{B}_{22}^*\right) + o_p(1).
\end{aligned} \tag{B.1}$$

Here, when the null hypothesis $\boldsymbol{\Sigma}_{12} = \mathbf{B}_1\mathbf{B}_2^\top = \mathbf{O}$ is true, $\mathbf{B}_{11}^*\mathbf{B}_{22}^* = \mathbf{O}$. Therefore, when the null hypothesis is true, $\mathbf{B}_{11}^*$ and $\mathbf{B}_{22}^*$ can be simultaneous diagonalization. In other words, an appropriate orthogonal matrix $\mathbf{H}$ that satisfies the following decomposition exists

$$\mathbf{B}_{11}^* = \mathbf{H}\widetilde{\boldsymbol{\Lambda}}_{11}\mathbf{H}^\top, \mathbf{B}_{22}^* = \mathbf{H}\widetilde{\boldsymbol{\Lambda}}_{22}\mathbf{H}^\top,$$

where

$$\widetilde{\boldsymbol{\Lambda}}_1 = \begin{pmatrix} \boldsymbol{\Lambda}_1 & \mathbf{O}_{d_1,d_2} \\ \mathbf{O}_{d_2,d_1} & \mathbf{O}_{d_2,d_2} \end{pmatrix}, \widetilde{\boldsymbol{\Lambda}}_2 = \begin{pmatrix} \mathbf{O}_{d_1,d_1} & \mathbf{O}_{d_1,d_2} \\ \mathbf{O}_{d_2,d_1} & \boldsymbol{\Lambda}_2 \end{pmatrix}.$$

Under (A4), because the distribution of $\mathbf{f}_i$ is invariant under any orthogonal transformation, the principal term in (B.1) follows the same distribution as:

$$\frac{1}{n}\mathrm{tr}\left(\sum_{\substack{i,j=1\\ i\neq j}}^{n} \mathbf{f}_i \mathbf{f}_i^\top \widetilde{\boldsymbol{\Lambda}}_1 \mathbf{f}_j \mathbf{f}_j^\top \widetilde{\boldsymbol{\Lambda}}_2\right).$$

We let $\mathbf{V}_i = \mathbf{f}_i\mathbf{f}_i^\top - \mathbf{I}_d$, $\mathbf{W} = n^{-1/2}\sum_{i=1}^n \mathbf{V}_i$, and $\widetilde{\mathbf{z}} = \mathrm{vec}(\mathbf{W})$. As $\widetilde{\mathbf{\Lambda}}_1\widetilde{\mathbf{\Lambda}}_2 = \widetilde{\mathbf{\Lambda}}_2\widetilde{\mathbf{\Lambda}}_1 = \mathbf{O}$ under the null hypothesis, the main term can be rewritten as follows:

$$\begin{aligned}
\frac{1}{n}\mathrm{tr}\left(\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbf{f}_i\mathbf{f}_i^\top\widetilde{\mathbf{\Lambda}}_1\mathbf{f}_j\mathbf{f}_j^\top\widetilde{\mathbf{\Lambda}}_2\right) =& \frac{1}{n}\mathrm{tr}\left(\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbf{V}_i\widetilde{\mathbf{\Lambda}}_1\mathbf{V}_j\widetilde{\mathbf{\Lambda}}_2\right)\\
=&\widetilde{\mathbf{z}}^\top\left(\widetilde{\mathbf{\Lambda}}_1\otimes\widetilde{\mathbf{\Lambda}}_2\right)\widetilde{\mathbf{z}}\\
&-\frac{1}{n}\sum_{i=1}^n\left\{\mathrm{tr}(\widetilde{\mathbf{\Lambda}}_1\mathbf{V}_i\widetilde{\mathbf{\Lambda}}_2\mathbf{V}_i) - \mathrm{tr}(\widetilde{\mathbf{\Lambda}}_1)\mathrm{tr}(\widetilde{\mathbf{\Lambda}}_2)\right\}\\
&-\mathrm{tr}(\widetilde{\mathbf{\Lambda}}_1)\mathrm{tr}(\widetilde{\mathbf{\Lambda}}_2)\\
=&\widetilde{\mathbf{z}}^\top\left(\widetilde{\mathbf{\Lambda}}_1\otimes\widetilde{\mathbf{\Lambda}}_2\right)\widetilde{\mathbf{z}} - \mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2) + o_p(1).
\end{aligned}$$

From CLT, under (A1)–(A4),

$$\widetilde{\mathbf{z}} \rightsquigarrow \mathcal{N}_{d^2}\left(\mathbf{0}, \mathbf{I}_{d^2} + \mathbf{K}_{d^2}\right)\ (n, p_1, p_2 \to \infty).$$

From the above discussion, under (A1)–(A4) and $\mathcal{H}$,

$$\frac{n\|\widehat{\mathbf{\Sigma}_{12}}\|_F^2}{p_1p_2} + \mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2) \rightsquigarrow \mathbf{z}^\top(\mathbf{I}_{d^2} + \mathbf{K}_{d^2})\left(\widetilde{\mathbf{\Lambda}}_1\otimes\widetilde{\mathbf{\Lambda}}_2\right)\mathbf{z}\ (n, p_1, p_2 \to \infty), \tag{B.2}$$

where $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{d^2})$. Furthermore, because $(\mathbf{I}_{d^2} + \mathbf{K}_{d^2})(\widetilde{\mathbf{\Lambda}}_1 \otimes \widetilde{\mathbf{\Lambda}}_2)$ is a lower-triangular matrix, the eigenvalue of $(\mathbf{I}_{d^2} + \mathbf{K}_{d^2})(\widetilde{\mathbf{\Lambda}}_1 \otimes \widetilde{\mathbf{\Lambda}}_2)$ is $\lambda_{1i}\lambda_{2j}$ for $i \in \{1, \ldots, d_1\}, j \in \{1, \ldots, d_2\}$. As the distribution of $\mathbf{z}$ is invariant with respect to any orthogonal transformation, the quadratic form in (B.2) converges to a weighted chi-square distribution with $\lambda_{1i}\lambda_{2j}$ as a weight. Therefore, under (A1)–(A4),

$$\frac{n\|\widehat{\mathbf{\Sigma}_{12}}\|_F^2}{p_1p_2} + \mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2) \rightsquigarrow \sum_{i=1}^{d_1}\sum_{j=1}^{d_2}\lambda_{1i}\lambda_{2j}\chi_{ij}^2\ (n, p_1, p_2 \to \infty). \tag{B.3}$$

From (A.1) and $\|\mathbf{\Sigma}_{gg}\|_F/p_g = \|\mathbf{\Lambda}_g\|_F + o(1)$,

$$nMRV_{12} = \frac{n\|\widehat{\mathbf{\Sigma}_{12}}\|_F^2/(p_1p_2)}{\|\mathbf{\Lambda}_1\|_F\|\mathbf{\Lambda}_2\|_F} + o_p(1)\ (n, p_1, p_2 \to \infty). \tag{B.4}$$

Theorem 2 is proven by applying Slutsky's theorem to (B.3) and (B.4).

C. Proof of Theorem 3

The difference between the random variables in Theorem 2 and T is given by:

$$nMRV_{12} + \frac{\mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2)}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}_1^2)\mathrm{tr}(\mathbf{\Lambda}_2^2)}} - T = \frac{\mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2)}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}_1^2)\mathrm{tr}(\mathbf{\Lambda}_2^2)}} - \frac{\widehat{\mathrm{tr}(\mathbf{\Lambda}_1)}\widehat{\mathrm{tr}(\mathbf{\Lambda}_2)}}{\sqrt{\widehat{\mathrm{tr}(\mathbf{\Lambda}_1^2)}\widehat{\mathrm{tr}(\mathbf{\Lambda}_2^2)}}}. \tag{C.1}$$

Using Lemmas 1 and 2 in [8], the following two properties are established.

1. Under (A1)–(A4), $\widehat{\lambda}_{gi} = \lambda_{gi} + o_p(1)$ for $i \in \{1, \ldots, d_g\},\ g \in \{1, 2\}$.
2. There exists $c_g \in (0, 1]$ such that $\Pr(\widehat{d}_g = d_g) \to 1$ under (A1)–(A4), for any $i_{g,\max} \in (d_g, \lfloor c_g \min(p_g, n) \rfloor - d_g - 1]$, where $\lfloor \cdot \rfloor$ denotes the floor function.

These two properties yield the following conclusions. Under (A1)–(A3),

$$\widehat{\mathrm{tr}(\mathbf{\Lambda}_g)} = \mathrm{tr}(\mathbf{\Lambda}_g) + o_p(1), \quad \widehat{\mathrm{tr}(\mathbf{\Lambda}_g^2)} = \mathrm{tr}(\mathbf{\Lambda}_g^2) + o_p(1). \tag{C.2}$$

By integrating (C.1) and (C.2) under (A1)–(A3), we obtain

$$nMRV_{12} + \frac{\mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2)}{\sqrt{\mathrm{tr}(\mathbf{\Lambda}_1^2)\mathrm{tr}(\mathbf{\Lambda}_2^2)}} - T = o_p(1). \tag{C.3}$$

Theorem 3 is proven by integrating Theorems 2 and (C.3).

D. Proof of Theorem 4

Under fixed alternative $\mathcal{A}_F$ and (A1)–(A4),

$$\sqrt{n}(MRV_{12} - \rho V_{12}) = \rho V_{12}\{\mathrm{vec}(\mathbf{E})\}^\top \widetilde{\mathbf{z}} + o_p(1), \tag{D.1}$$

where

$$\mathbf{E} = \frac{p_1 p_2}{\|\mathbf{\Sigma}_{12}\|_F^2}(\mathbf{B}_{22}^* \mathbf{B}_{11}^* + \mathbf{B}_{11}^* \mathbf{B}_{22}^*) - \frac{p_1^2}{\|\mathbf{\Sigma}_{11}\|_F^2}\mathbf{B}_{11}^{*2} - \frac{p_2^2}{\|\mathbf{\Sigma}_{22}\|_F^2}\mathbf{B}_{22}^{*2}.$$

Additionally, from the central limit theorem, because the distribution of $\widetilde{\mathbf{z}}$ converges to $\mathcal{N}_{d^2}(\mathbf{0}, \mathbf{I}_{d^2} + \mathbf{K}_{d^2})$, the first term on the right-hand side of (D.1) converges to a normal distribution with mean 0 and variance $2\rho V_{12}^2 \mathrm{tr}(\mathbf{E}^2)$. Here, $2\mathrm{tr}(\mathbf{E}^2)$ is expanded as follows.

$$\begin{aligned}2\mathrm{tr}(\mathbf{E}^2) =& 2\left[\frac{p_1^4\|\mathbf{B}_{11}^{*2}\|_F^2}{\|\mathbf{\Sigma}_{11}\|_F^4} + \frac{p_2^4\|\mathbf{B}_{22}^{*2}\|_F^2}{\|\mathbf{\Sigma}_{22}\|_F^4} + 2\frac{p_1^2 p_2^2 \mathrm{tr}\{(\mathbf{B}_{11}^*\mathbf{B}_{22}^*)^2\} + \mathrm{tr}(\mathbf{B}_{11}^{2*}\mathbf{B}_{22}^{2*})}{\|\mathbf{\Sigma}_{12}\|_F^4}\right.\\ &\left.+\frac{2p_1^2p_2^2\mathrm{tr}(\mathbf{B}_{11}^{*2}\mathbf{B}_{22}^{*2})}{\|\mathbf{\Sigma}_{11}\|_F^2\|\mathbf{\Sigma}_{22}\|_F^2} - \frac{4p_1^3p_2\mathrm{tr}(\mathbf{B}_{11}^{*3}\mathbf{B}_{22}^*)}{\|\mathbf{\Sigma}_{11}\|_F^2\|\mathbf{\Sigma}_{12}\|_F^2} - \frac{4p_1p_2^3\mathrm{tr}(\mathbf{B}_{22}^{*3}\mathbf{B}_{11}^*)}{\|\mathbf{\Sigma}_{22}\|_F^2\|\mathbf{\Sigma}_{21}\|_F^2}\right].\end{aligned}$$

Therefore, Theorem 4 is proven.

E. Proof of Corollary 1

We verify that $\hat{t}_\alpha$ is a consistent estimator of t_α, where t_α is $(1-\alpha)$-quantile of the weighted chi-squared distribution in (2.3). Therefore, we need only evaluate $\Pr(T > t_\alpha | \mathcal{A}_F)$. From Theorem 4, under (A1)–(A4), we obtain

$$\begin{aligned}\Pr(T > t_\alpha|\mathcal{A}_F) =& 1 - \Phi\left(\frac{t_\alpha}{\sigma\sqrt{n}} - \frac{\mathrm{tr}(\mathbf{\Lambda}_1)\mathrm{tr}(\mathbf{\Lambda}_2)}{\sigma\sqrt{n\mathrm{tr}(\mathbf{\Lambda}_1^2)\mathrm{tr}(\mathbf{\Lambda}_2^2)}} - \frac{\sqrt{n}\rho V_{12}}{\sigma}\right) + o(1)\\ =& 1 - \Phi(-\infty) + o(1) = 1 + o(1).\end{aligned}$$

Therefore, Corollary 1 is proven.

F. Proof of Theorem 5

First, we derive stochastic asymptotic expansion of $n\|\widehat{\boldsymbol{\Sigma}_{12}}\|_F^2/(p_1p_2)$. Under (A1)–(A4) and $\mathcal{A}_L$,

$$\frac{n\|\widehat{\boldsymbol{\Sigma}_{12}}\|_F^2}{p_1p_2} + \mathrm{tr}(\boldsymbol{\Lambda}_1)\mathrm{tr}(\boldsymbol{\Lambda}_2) = \widetilde{\mathbf{z}}^\top \widetilde{\mathbf{C}}\widetilde{\mathbf{z}} + \widetilde{\mathbf{c}}^\top\widetilde{\mathbf{z}} + \frac{n\|\boldsymbol{\Sigma}_{12}\|_F^2}{p_1p_2} + o_p(1), \tag{F.1}$$

where

$$\widetilde{\mathbf{C}} = \frac{1}{p_1p_2}(\mathbf{B}_1^\top\mathbf{B}_1)\otimes(\mathbf{B}_2^\top\mathbf{B}_2),\ \ \widetilde{\mathbf{c}} = \frac{\sqrt{n}}{p_1p_2}\mathrm{vec}(\mathbf{B}_1^\top\mathbf{B}_1\mathbf{B}_2^\top\mathbf{B}_2 + \mathbf{B}_2^\top\mathbf{B}_2\mathbf{B}_1^\top\mathbf{B}_1).$$

The first term on the right-hand side of (F.1) is always $O_p(1)$ regardless of $\eta = 1/2$ and $\eta > 1/2$. The variance of the second term can be evaluated as follows:

$$\begin{aligned}
\mathrm{var}(\widetilde{\mathbf{c}}^\top\widetilde{\mathbf{z}}) =& \frac{4n}{p_1^2p_2^2}\left[\mathrm{tr}\{(\mathbf{B}_1^\top\mathbf{B}_1\mathbf{B}_2^\top\mathbf{B}_2)^2\} + \mathrm{tr}\{(\mathbf{B}_1^\top\mathbf{B}_1)^2(\mathbf{B}_2^\top\mathbf{B}_2)^2\}\right]\\
\le& \frac{4n}{p_1^2p_2^2}\mathrm{tr}\{(\mathbf{B}_1^\top\mathbf{B}_1\mathbf{B}_2^\top\mathbf{B}_2)^2\}\\
&+ \frac{4n}{p_1^2p_2^2}\lambda_1(\mathbf{B}_1^\top\mathbf{B}_1)\lambda_1(\mathbf{B}_2^\top\mathbf{B}_2)\mathrm{tr}(\mathbf{B}_1^\top\mathbf{B}_1\mathbf{B}_2^\top\mathbf{B}_2)\\
=& 4n^{1-2\eta}\{\mathrm{tr}(\boldsymbol{\Xi}^2) + \lambda_1(\mathbf{B}_{11}^*)\lambda_1(\mathbf{B}_{22}^*)\Delta\} + o(1) = O(n^{1-2\eta}).
\end{aligned}$$

The third term is evaluated as $n\|\boldsymbol{\Sigma}_{12}\|_F^2/(p_1p_2) = n^{1-2\eta}\Delta + o(1)$. Therefore, the second and third terms on the right side of (F.1) are related to $\eta = 1/2$ and $\eta > 1/2$.

Under $\mathcal{A}_L$ and $\eta > 1/2$, because $n\|\boldsymbol{\Sigma}_{12}\|_F^2/(p_1p_2) = o(1)$ and $\mathrm{var}(\widetilde{\mathbf{c}}^\top\widetilde{\mathbf{z}}) = o(1)$, the second and third terms on the right side of (F.1) are negligible. Additionally, from the central limit theorem, $\widetilde{\mathbf{z}}$ is asymptotically distributed $\mathcal{N}_{d^2}(\mathbf{0}, \mathbf{I}_{d^2} + \mathbf{K}_{d^2})$. Therefore, under (A1)–(A4), $\mathcal{A}_L$, and $\eta > 1/2$,

$$\frac{n\|\widehat{\boldsymbol{\Sigma}_{12}}\|_F^2}{p_1p_2} + \mathrm{tr}(\boldsymbol{\Lambda}_1)\mathrm{tr}(\boldsymbol{\Lambda}_2) \rightsquigarrow \mathbf{z}^\top\widetilde{\mathbf{C}}^*\mathbf{z}, \tag{F.2}$$

where $\widetilde{\mathbf{C}}^* = \mathbf{B}_{11}^*\otimes\mathbf{B}_{22}^*$.

Under $\mathcal{A}_L$ and $\eta = 1/2$, because $n\|\boldsymbol{\Sigma}_{12}\|_F^2/(p_1p_2) \asymp 1$ and $\mathrm{var}(\widetilde{\mathbf{c}}^\top\widetilde{\mathbf{z}}) \asymp 1$, the second and third terms on the right side of (F.1) are nonnegligible. Therefore, under (A1)–(A4), $\mathcal{A}_L$ and $\eta = 1/2$,

$$\frac{n\|\widehat{\boldsymbol{\Sigma}_{12}}\|_F^2}{p_1p_2} + \mathrm{tr}(\boldsymbol{\Lambda}_1)\mathrm{tr}(\boldsymbol{\Lambda}_2) \rightsquigarrow \mathbf{z}^\top\widetilde{\mathbf{C}}^*\mathbf{z} + \widetilde{\mathbf{c}}^{*\top}\mathbf{z} + \Delta, \tag{F.3}$$

where

$$\widetilde{\mathbf{c}}^* = \boldsymbol{\Xi} + \boldsymbol{\Xi}^\top.$$

Furthermore, from (A.1) and the basic properties of the stochastic convergence under (A1)–(A3),

$$nMRV_{12} = \frac{n\|\widehat{\boldsymbol{\Sigma}_{12}}\|_F^2/(p_1p_2)}{\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F} + o_p(1). \tag{F.4}$$

Theorem 5 is proven by integrating (F.2)–(F.4).

G. Proof of Corollary 2

We verify that $\hat{t}_\alpha$ is a consistent estimator of t_α, where t_α is $(1-\alpha)$-quantile of the weighted chi-squared distribution in (2.3). Therefore, we only evaluate $\Pr(T > t_\alpha|\mathcal{A}_L)$.

From Theorem 5, under (A1)–(A4) and $\eta = 1/2$,

$$\Pr(T > t_\alpha|\mathcal{A}_L) = G\{t_\alpha - \Delta/(\|\boldsymbol{\Lambda}_1\|_F\|\boldsymbol{\Lambda}_2\|_F)\} + o(1).$$

Under $\mathcal{A}_L$,

$$\begin{aligned}\left\|\frac{1}{p_1p_2}\mathbf{B}_1^\top\mathbf{B}_1\mathbf{B}_2^\top\mathbf{B}_2\right\|_F \leq& \frac{1}{n^\eta}\left\|\frac{n^\eta}{p_1p_2}\mathbf{B}_1^\top\mathbf{B}_1\mathbf{B}_2^\top\mathbf{B}_2 - \boldsymbol{\Xi}\right\|_F + \frac{1}{n^\eta}\|\boldsymbol{\Xi}\|_F \\ =& o(1) \quad (n, p_1, p_2 \to \infty).\end{aligned}$$

Therefore, $\|\mathbf{B}_{11}^*\mathbf{B}_{22}^*\|_F = o(1)$ holds under $\mathcal{A}_L$. Because $\mathbf{B}_{11}^*\mathbf{B}_{22}^* = \mathbf{B}_{22}^*\mathbf{B}_{11}^* = \mathbf{O}$, $\mathbf{B}_{11}^*$ and $\mathbf{B}_{22}^*$ can be diagonalized simultaneously. Therefore, $\mathbf{z}^\top\mathbf{C}^*\mathbf{z}$ shows the same asymptotic distribution as in (2.3). The asymptotic distribution of T under $\mathcal{A}_L$ and $\eta > 1/2$ is similar to that of T under the null hypothesis $\mathcal{H}$, Hence the asymptotic power is equal to α.

Acknowledgements

The authors thank H. Watanabe for the technical assistance with the simulation experiments. Also, we would like to thank Editage (www.editage.jp) for English language editing. The research of the first and second authors were supported by JSPS KAKENHI Grant number 24K14867 and 20K11714, respectively.

References

[1] Ahmad, M. R. (2019). A significance test of the RV coefficient in high dimensions. *Comput. Stat. Data Anal.* **131**, 116–130.

[2] Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica* **81**, 1203–1227.

[3] Cornillon, P. A. (1998). Prise en compte de proximites en analyse factorielle et comparative. *Thesis, Montpellier*, 101–118.

[4] Escoufier, Y. (1973). Le Traitement des variables vectorielles. *Biometrics* **29**, 751–760.

[5] Hyodo, M. and Nishiyama, T. (2018). Simultaneous testing of the mean vector and the covariance matrix for high-dimensional data. *TEST* **27**, 680–699.

[6] Hyodo, M. and Nishiyama, T. (2021). Simultaneous testing of the mean vector and covariance matrix among k populations for high-dimensional data. *Commun. Stat.-Theory Methods* **50**, 663–684.

[7] Hyodo, M., Nishiyama, T., and Pavlenko, T. (2020). Testing for independence of high-dimensional variables:ρV-coefficient based approach. *J. Multivariate Anal.* **178**, 104627.

[8] Hyodo, M., Nishiyama, T., and Pavlenko, T. (2023). A Behrens–Fisher problem for general factor models in high dimensions. *J. Multivariate Anal.* **195**, 105162.

[9] Josse, J. and Holmes, S. (2016). Measuring multivariate association and beyond. *Statistics Surveys* **10**, 132–167.

[10] Josse, J., Pages, J., and Husson, F. (2008). Testing the significance of the RV coefficient. *Comput. Stat. Data Anal.* **53**, 82–91.

[11] Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *Ann. Stat.* **40(2)**, 908–940.

[12] Sun, P., Tang, Y., Cao, M. (2022). Homogeneity test of multi-sample covariance matrices in high dimensions. *Mathematics* **10(22)**, 4339.

[13] Yamadaa, Y., Hyodo, M., and Nishiyama, T. (2017). Testing block-diagonal covariance structure for high-dimensional data under non-normality. *J. Multivariate Anal.* **155**, 305–316.