JOURNAL OF TIME SERIES ANALYSIS
J. Time Ser. Anal. (2023)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12702

SPECIAL ISSUE

TESTING FOR SYMMETRIC CORRELATION MATRICES WITH APPLICATIONS TO FACTOR MODELS

NAN-JUNG HSU[a] LAI HENG SIM[a] AND RUEY S. TSAY[b]
[a]*Institute of Statistics, National Tsing-Hua University, Hsinchu, China*
[b]*Booth School of Business, University of Chicago, Chicago, IL, USA*

Factor models have been widely used in recent years to model high-dimensional spatio-temporal data. However, the validity of employing factor models in a specific application has received less attention. This article proposes test statistics for testing the symmetry in cross-correlation matrices of a high-dimensional stochastic process implied by exact factor models. A rejection of symmetry indicates that the use of an exact factor model is questionable. Both simulations and real examples are used to demonstrate the applications and to study the finite-sample performance of the proposed test statistics. Empirical results show that the proposed test statistics are effective in identifying cases where exact factor models are not appropriate, providing valuable guidance for choosing factor models in a high-dimensional setting.

Received 24 August 2022; Revised 23 April 2023; Accepted 24 April 2023

Keywords: Cross-correlation function; exact factor model; false discovery rate; high-dimensional time series; multiple hypotheses testing; Wald test.

MOS subject classification: 62M10; 62M20.

1. INTRODUCTION

Dimension reduction or structure regularization becomes a necessity in high-dimensional statistical analysis. One of the commonly used techniques in dimension reduction is to employ factor models in which a complex system is driven by a small number of common factors. In time series analysis, factor models have been widely used in recent years as an effective approach to modeling high-dimensional data. Anderson (1963) and Brillinger (1964) marked the early study of factor models in time series analysis. Geweke (1977) proposed a dynamic factor model for a vector time series which assumes that each individual series is a sum of two independent components. The first component is common and is generated by a small number of factors, and the second component is specifically referred to as noise. The author also proposed a frequency-domain procedure to estimate the model. Peña and Box (1987) proposed a factor model in which all dynamics are driven by the factors and the noises are serially uncorrelated. In addition, the factors are assumed to be independent of each other. This model is often referred to as an exact factor model, because it closely follows that of independent data. Chamberlain (1983) and Chamberlain and Rothschild (1983) introduced the approximate factor models, where the noises may be serially dependent. Certain conditions are needed on the covariance matrix of the noises and the factors to render the model asymptotically identifiable. Bai and Ng (2002) investigated high-dimensional information criteria for determining the number of factors of an approximate factor model. Onatski (2009) studied hypothesis testing about the number of factors in large factor models. Forni *et al.* (2000) proposed the generalized dynamic factor models addressing the issues of identification and estimation. Forni *et al.* (2005) further studied one-sided estimation and forecasting of generalized dynamic factor models. Other types of factor models have also been proposed in the literature for high-dimensional time series analysis. See, for instance, Peña and Poncela (2006), Lam and Yao (2012), Lam

*Correspondence to: Ruey S. Tsay, Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue Chicago, IL 60637, USA. Email: ruey.tsay@chicagobooth.edu

et al. (2011), and Gao and Tsay (2019, 2022), among others. Interested readers are referred to Peña and Tsay (2021, Ch. 6) and the references therein for further information.

The applicability of factor models to a given data set of vector time series is less studied, however. One often takes it for granted that factor models are applicable to all high-dimensional time series. In addition, there are no clear guidelines available on which class of factor models is suitable for a given data set. The goal of this article is to close this gap by assessing the validity of applying exact factor models in an application. Specifically, we propose test statistics to test the null hypothesis of symmetry in the cross-correlation matrices of a high-dimensional time series. For an exact factor model, the cross-correlation matrices are symmetric. Therefore, if some cross-correlation matrices are not symmetric, then employing an exact factor model for the data set under study becomes questionable. Consider a k-dimensional time series $\boldsymbol{y}_t = (y_{1t}, \dots, y_{kt})'$, where the superscript $'$ denotes transpose. We said that $\boldsymbol{y}_t$ follows an exact factor model if it satisfies

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \boldsymbol{L}\boldsymbol{f}_t + \boldsymbol{\epsilon}_t, \tag{1}$$

where $\boldsymbol{\mu}$ is a constant vector, $\boldsymbol{f}_t = (f_{1t}, \dots, f_{rt})'$ is an r-dimensional factor process with $r \ll k$ and $E(\boldsymbol{f}_t) = \boldsymbol{0}$, f_{jt} and $f_{i,t-\ell}$ are uncorrelated for all ℓ and $i \neq j$, $\boldsymbol{L}$ is a $k \times r$ loading matrix, and $\boldsymbol{\epsilon}_t$ is a white noise series such that $E(\boldsymbol{\epsilon}_t) = \boldsymbol{0}$ and $\text{cov}(\boldsymbol{\epsilon}_t) = \boldsymbol{D}_\epsilon$, which is a positive-definite diagonal matrix. In model (1), $\{\boldsymbol{f}_t\}$ and $\{\boldsymbol{\epsilon}_t\}$ are uncorrelated processes. Let $\boldsymbol{\Sigma}_\ell^z = \text{cov}(z_t, z_{t-\ell})$ be the lag-ℓ autocovariance matrix of a vector time series z_t. For the exact factor model (1), we have

$$\boldsymbol{\Sigma}_\ell^y = \boldsymbol{L}\boldsymbol{\Sigma}_\ell^f \boldsymbol{L}', \quad \ell = 1, 2, \dots . \tag{2}$$

Since $\boldsymbol{\Sigma}_\ell^f$ is a diagonal matrix, $\boldsymbol{\Sigma}_\ell^y$ is symmetric if $\boldsymbol{y}_t$ follows an exact factor model. Let $\boldsymbol{\Gamma}_\ell = \text{diag}(\boldsymbol{\Sigma}_0^y)^{-1/2}\boldsymbol{\Sigma}_\ell^y \text{diag}(\boldsymbol{\Sigma}_0^y)^{-1/2}$ be the lag-ℓ cross-correlation matrix of $\boldsymbol{y}_t$, where $\text{diag}(\boldsymbol{A})$ is a diagonal matrix consisting of the diagonal elements of the matrix $\boldsymbol{A}$, and we omit the superscript ^{y} from the correlation matrix as there is no confusion. In this article, we focus on testing the null hypothesis that $\boldsymbol{\Gamma}_\ell$ is symmetric to assess the applicability of exact factor models.

We proposed three approaches to test the hypothesis of symmetry in cross-correlations. The first approach is the Wald test based on asymptotic distributions of the sample cross-correlations. The second approach utilizes the maximum statistic of sample cross-correlation matrices and applies the extreme value theory to make statistical inference. The third approach adopts the idea of multiple-hypothesis testing based on Benjamini–Hochberg (BH) procedures to control the false discovery rate. Failure to reject the null hypothesis of symmetry in $\boldsymbol{\Gamma}_\ell$ would support the use of an exact factor model for $\boldsymbol{y}_t$. Such information should be helpful in applying exact factor models because they can dramatically simplify the modeling process, especially when the dimension k is large and the number of common factors r is small.

Asymptotic normality of sample cross-correlations of a stationary vector time series has been established in the literature (Roy, 1989). However, since the dynamic dependence of $\boldsymbol{y}_t$ can be complicated, it would be very time-consuming to compute the asymptotic variances of many sample cross-correlations when the dimension is high. To simplify the computation, we adopt a block bootstrap procedure to estimate the variances of sample cross-correlations used in our testing. The use of block bootstrap methods in time series analysis is common in the literature, see for instance, Kunsch (1989), Bühlmann (2002), Lahiri (2003), and Politis (2003).

The finite-sample performance of the proposed tests is investigated via a simulation study, which shows a fairly accurate type-I error control and good testing powers against a variety of asymmetric scenarios. In particular, the Wald test is effective in low-dimensional settings but it is computationally infeasible in high-dimensional settings. The maximum approach is effective for high-dimensional settings, but not effective for low-dimensional series, say $k < 10$. The BH approach is effective for both low- and high-dimensional settings, but it requires intensive computing when the dimension is high, say $k > 100$. In terms of computation costs, the maximum approach is the most efficient one when the dimension k is large, for example, $k \geq 100$.

DOI: 10.1111/jtsa.12702

To demonstrate the applicability of the proposed test statistics, we consider three economic datasets often used in the literature for factor modeling. Based on the test results of symmetric cross-correlation functions (CCF), we are able to choose an adequate factor model for each application and obtain satisfactory fitting results.

The article is organized as follows. Section 2 introduces the hypothesis testing for the symmetry of CCF and proposes test statistics that have sound theoretical justifications. Section 3 presents a simulation study to validate the proposed methodology in finite samples. In Section 4, we apply the proposed methodology to three economic datasets. Section 5 provides some concluding remarks.

2. TESTING SYMMETRIC CROSS-CORRELATIONS

Consider a k-dimensional stationary process with a sample realization $\{\boldsymbol{y}_t\}_{t=1}^T$ of T observations. We can estimate $\boldsymbol{\Sigma}_\ell^y$ and $\boldsymbol{\Gamma}_\ell$ by their sample counterparts:

$$\widehat{\boldsymbol{\Sigma}}_\ell^y = \frac{1}{T}\sum_{t=\ell+1}^{T}\left(\boldsymbol{y}_t - \bar{\boldsymbol{y}}\right)\left(\boldsymbol{y}_{t-\ell} - \bar{\boldsymbol{y}}\right)', \tag{3}$$

$$\widehat{\boldsymbol{\Gamma}}_\ell \equiv [\hat{\rho}_{ij}(\ell)] = \mathrm{diag}\{\widehat{\boldsymbol{\Sigma}}_0^y\}^{-\frac{1}{2}}\ \widehat{\boldsymbol{\Sigma}}_\ell^y\ \mathrm{diag}\{\widehat{\boldsymbol{\Sigma}}_0^y\}^{-\frac{1}{2}}, \tag{4}$$

where $\bar{\boldsymbol{y}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t$ is the sample mean. Furthermore, define

$$\boldsymbol{\rho}_\ell = \mathrm{vec}(\boldsymbol{\Gamma}_\ell), \quad \hat{\boldsymbol{\rho}}_\ell = \mathrm{vec}(\widehat{\boldsymbol{\Gamma}}_\ell),$$
$$\boldsymbol{\delta}_\ell \equiv \mathrm{vech}(\boldsymbol{\Gamma}_\ell - \boldsymbol{\Gamma}'_\ell) = \boldsymbol{D}_k\,\boldsymbol{\rho}_\ell, \quad \hat{\boldsymbol{\delta}}_\ell = \boldsymbol{D}_k\,\hat{\boldsymbol{\rho}}_\ell,$$

where $\boldsymbol{\delta}_\ell$ and $\hat{\boldsymbol{\delta}}_\ell$ are $k(k-1)/2$-dimensional vectors, vech($\boldsymbol{A}$) denotes the vectorized operation of the lower-triangular sub-matrix of $\boldsymbol{A}$, and $\boldsymbol{D}_k$ denotes the $k(k-1)/2\times k^2$ matrix in which each row vector has exactly one entry 1 and one entry -1 to carry out the differencing $\hat{\rho}_{ij}(\ell) - \hat{\rho}_{ji}(\ell)$ for a given (i,j) of indexes with $k \geq i > j \geq 1$.

We are interested in testing the hypothesis of symmetry in cross-correlation matrices:

$$H_0 : \boldsymbol{\Gamma}_\ell = \boldsymbol{\Gamma}'_\ell \quad \text{vs} \quad H_a : \boldsymbol{\Gamma}_\ell \neq \boldsymbol{\Gamma}'_\ell, \quad \ell = 1, 2, \ldots, \tag{5}$$

which is equivalent to checking

$$H_0 : \boldsymbol{\delta}_\ell = \boldsymbol{0} \quad \text{vs} \quad H_a : \boldsymbol{\delta}_\ell \neq \boldsymbol{0}, \quad \ell = 1, 2, \ldots. \tag{6}$$

In what follows, we consider three approaches to conducting the testing. The first approach uses the Wald test statistic based on the asymptotic distributions of sample cross-correlations mentioned in the *Introduction*. The second approach utilizes the maximum statistic of sample cross-correlations and applies the extreme value theory to make statistical inference. Details are given in Section 2.2. The third approach adopts the idea of multiple hypothesis testing based on Benjamini–Hochberg procedures and is given in Section 2.3.

2.1. Wald Test

The first approach proposed is based on the central limit theorem for sample cross-correlations $\hat{\boldsymbol{\rho}}_\ell$ given by Roy (1989) and Hannan (1976):

$$\sqrt{T}\left(\hat{\boldsymbol{\rho}}_\ell - \boldsymbol{\rho}_\ell\right) \rightarrow \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{V}_\ell\right), \quad \text{as} \quad T \rightarrow \infty, \tag{7}$$

under the regularity conditions in Assumption 1 below, where the formula of $\boldsymbol{V}_\ell \equiv \lim_{T\to\infty} \mathrm{var}\left(\sqrt{T}\hat{\boldsymbol{\rho}}_\ell\right)$ is specified in Roy (1989).

Assumption 1. (Regularity conditions)**.** (a) The process $\boldsymbol{y}_t$ is a stationary and ergodic process with a linear process representation having finite fourth moments, and the spectral density of each univariate series is square integrable. (b) All fourth-order cumulants of $\boldsymbol{y}_t$ are zero.

Accordingly, we have the following theorem.

Theorem 1. Under H_0 in (6) and Assumption 1, the Wald test statistic

$$Q_\ell \equiv T\, \hat{\boldsymbol{\delta}}_\ell' \left(\boldsymbol{D}_k \hat{\boldsymbol{V}}_\ell \boldsymbol{D}_k'\right)^{-1} \hat{\boldsymbol{\delta}}_\ell, \quad \ell = 1, 2, \ldots \tag{8}$$

has an asymptotic Chi-square distribution with degrees of freedom $k(k-1)/2$, where $\hat{\boldsymbol{V}}_\ell$ is a consistent estimator of $\boldsymbol{V}_\ell$.

We may define a Portmanteau test statistic by taking into account simultaneously the first ℓ lags of cross-correlations being symmetric. See the corollary below.

Corollary 1. Under H_0 in (6) and Assumption 1, the test statistic, for a given ℓ,

$$Q_{1:\ell} \equiv T\, \hat{\boldsymbol{\delta}}_{1:\ell}' \left[(\boldsymbol{I}_\ell \otimes \boldsymbol{D}_k)\, \hat{\boldsymbol{V}}_{1:\ell}\, (\boldsymbol{I}_\ell \otimes \boldsymbol{D}_k')\right]^{-1} \hat{\boldsymbol{\delta}}_{1:\ell}, \quad \ell = 1, 2, \ldots \tag{9}$$

has an asymptotic Chi-square distribution with degrees of freedom $\ell k(k-1)/2$, where $\hat{\boldsymbol{\delta}}_{1:\ell} = (\hat{\boldsymbol{\delta}}_1', \ldots, \hat{\boldsymbol{\delta}}_\ell')'$ and $\hat{\boldsymbol{V}}_{1:\ell}$ is a consistent estimator of $\boldsymbol{V}_{1:\ell} \equiv \lim_{T\to\infty} \mathrm{var}\left(\sqrt{T}(\hat{\boldsymbol{\rho}}_1', \ldots, \hat{\boldsymbol{\rho}}_\ell')'\right)$.

Similarly to the Portmanteau test for serial correlations, the joint test statistic $Q_{1:\ell}$ should be more powerful than the individual test statistics Q_ℓ. Based on Theorem 1 and Corollary 1, we reject H_0 at the significance level α if $Q_\ell > \chi^2_{k(k-1)/2}(\alpha)$ and $Q_{1:\ell} > \chi^2_{\ell k(k-1)/2}(\alpha)$ for a given ℓ, where $\chi^2_v(\alpha)$ denotes the upper tail α quantile of the Chi-square distribution with v degrees of freedom. In our data analysis later, we adopt the consistent estimates $\hat{\boldsymbol{V}}_\ell$ and $\hat{\boldsymbol{V}}_{1:\ell}$ derived in Melard *et al.* (1991) to implement the Wald tests and found that the test is effective in low-dimensional settings with a small dimension k and large sample size T, but its performance deteriorates quickly in the high-dimensional setting when T is not sufficiently large relative to the dimension k. This is caused mainly by the fact that both $\hat{\boldsymbol{V}}_\ell$ and $\hat{\boldsymbol{V}}_{1:\ell}$ become inaccurate estimates when the sample size is not large. In addition, the computation involved becomes very intensive.

2.2. Maximum Statistics

The second approach proposed is based on extreme value theory for the maximum statistic, which has been widely used for hypothesis testing in high-dimensional time series. See, for example, Chang *et al.* (2017), Xiao and Wu (2014), and Tsay (2020). For the testing problem considered, define the standardized statistics and the individual and joint test statistics:

$$\hat{\delta}_{ij}(\ell) \equiv \hat{\rho}_{ij}(\ell) - \hat{\rho}_{ji}(\ell), \quad Z_{\ell ij} = \frac{\hat{\delta}_{ij}(\ell)}{\widehat{\mathrm{se}}(\hat{\delta}_{ij}(\ell))}, \quad 1 \le j < i \le k, \tag{10}$$

$$\begin{aligned} Z_\ell^{(\max)} &\equiv \max_{1\le j<i\le k} Z_{\ell ij}, \quad Z_{1:\ell}^{(\max)} \equiv \max_{1\le r\le \ell} Z_r^{(\max)}, \\ Z_\ell^{(\min)} &\equiv \min_{1\le j<i\le k} Z_{\ell ij}, \quad Z_{1:\ell}^{(\min)} \equiv \min_{1\le r\le \ell} Z_r^{(\min)}, \\ M_\ell &\equiv \max\left\{Z_\ell^{(\max)}, -Z_\ell^{(\min)}\right\}, \quad M_{1:\ell} \equiv \max\left\{Z_{1:\ell}^{(\max)}, -Z_{1:\ell}^{(\min)}\right\}, \end{aligned} \tag{11}$$

DOI: 10.1111/jtsa.12702

for $\ell = 1, 2, \dots$, where the denominator $\widehat{\text{se}}(\hat{\delta}_{ij}(\ell))$ in (10) is obtained by a block bootstrap method discussed shortly after. Since $Z_{\ell ij}$'s correspond to the marginal elements of the standardized $\hat{\boldsymbol{\delta}}_\ell$, by Theorem 1, $Z_{\ell ij}$ is asymptotically $\mathcal{N}(0,1)$ distributed under H_0. Accordingly, $Z_\ell^{(\max)}$ and $Z_\ell^{(\min)}$ are the extreme values of $k(k-1)/2$ random variates that are asymptotically $\mathcal{N}(0,1)$; $Z_{1:\ell}^{(\max)}$ and $Z_{1:\ell}^{(\min)}$ are the extreme values of $\ell k(k-1)/2$ random variates that are asymptotically $\mathcal{N}(0,1)$. For completeness, we first review the extreme value theorem for independent normal variates in the following lemma:

Lemma 1. Assume $Z_i \overset{i.i.d.}{\sim} \text{N}(0,1)$, for $i = 1, 2, \dots, n$, and define the maximum and minimum statistics: $Z_{(n)} \equiv \max_{1 \le i \le n} Z_i$ and $Z_{(1)} \equiv \min_{1 \le i \le n} Z_i$. Then, as $n \to \infty$,

$$a_n \left(Z_{(n)} - b_n \right) \to_d \Lambda, \quad a_n \left(-Z_{(1)} - b_n \right) \to_d \Lambda,$$

where $\to_d$ denotes convergence in distribution, Λ has the standard Gumbel distribution $G(x) = \exp(-\exp(-x))$ for $x \in \mathbb{R}$, a_n and b_n are normalizing constants (Embrechts *et al.*, 2013) given by

$$a_n = \sqrt{2 \ln n}, \quad b_n = a_n - \frac{\ln 4\pi + \ln \ln n}{2a_n}.$$

In this study, we employ the test statistics M_ℓ and $M_{1:\ell}$ based on the extremes of $Z_{\ell ij}$'s. Although $Z_{\ell ij}$'s are dependent variates in our case, Lemma 1 still provides a good approximation for deriving the null distribution for M_ℓ and $M_{1:\ell}$ under H_0 whenever the dimension k of the series is sufficiently large. Some justifications for this claim and an adjustment for the extreme value distribution to account for the dependence among $Z_{\ell ij}$'s are provided in the Appendix, adopting some results of Afonja (1972). Following Lemma 1, we have the following theorem.

Theorem 2. Assume $\boldsymbol{y}_t$ satisfies Assumption 1(a). For a sufficiently large k, we reject H_0 at significance level α if

$$M_\ell > b_{1,k} - \frac{1}{a_{1,k}} \ln[-\ln(1-\alpha/2)], \tag{12}$$

$$M_{1:\ell} > b_{\ell,k} - \frac{1}{a_{\ell,k}} \ln[-\ln(1-\alpha/2)], \tag{13}$$

for a given $\ell \in \{1, 2, \dots\}$, where

$$a_{\ell,k} = \sqrt{2 \ln[\ell k(k-1)/2]}, \quad b_{\ell,k} = a_{\ell,k} - \frac{\ln 4\pi + \ln \ln[\ell k(k-1)/2]}{2a_{\ell,k}}.$$

Theorem 2 can be obtained directly from Lemma 1 with $n = k(k-1)/2$ for individual test M_ℓ and $n = \ell k(k-1)/2$ for joint test $M_{1:\ell}$, or refer to the proof in Tsay (2020). Although the normalized $Z_{\ell ij}$ are not independent $\mathcal{N}(0,1)$ in general, the proposed decision rules in Theorem 2 work very well as shown in our simulation study. Again, more discussions regarding this issue are provided in the Appendix.

Return to the estimation of $\widehat{\text{se}}(\hat{\delta}_{ij}(\ell))$ in (10). Since $\hat{\delta}_{ij}(\ell)$ corresponds to a single element in $\boldsymbol{\delta}_\ell$, its variance estimate could be directly adopted from the corresponding diagonal entry of the asymptotic variance $\frac{1}{T}\boldsymbol{D}_\ell \hat{\boldsymbol{V}}_\ell \boldsymbol{D}_k'$ described in (8). However, this theoretical estimate often fares poorly in a high-dimensional setting. We recommend using a block bootstrap (Kunsch, 1989) to obtain a data-driven estimate $\widehat{\text{se}}_{\text{BT}}(\hat{\delta}_{ij}(\ell))$. Block bootstrap has been commonly used in time series analysis. See, for instance, Lahiri (2003) and the references therein. The bootstrap procedure used is given in Algorithm 1. Regarding the block size, m_{BT} can be chosen following two rules in practice. One is to set at the magnitude of $T^{1/3}$ (Hall *et al.*, 1995). The other choice is the smallest lag such that the sample CCFs of $\boldsymbol{y}_t$ are insignificant beyond that lag.

DOI: 10.1111/jtsa.12702 *Journal of Time Series Analysis* published by John Wiley & Sons Ltd.

Algorithm 1. Bootstrap procedure for estimating se($\hat{\delta}_{ij}(\ell)$) in (10)

Specify the block size m_{BT} and the bootstrap replicates B.

Define overlapping sub-block data with size m_{BT}:

$$\mathcal{B}_j \equiv \{\boldsymbol{y}_j, \dots, \boldsymbol{y}_{j+m_{\mathrm{BT}}-1}\}, \quad j \in \mathcal{J} \equiv \{1, 2, \dots, T+1-m_{\mathrm{BT}}\}.$$

for $b = 1, \dots, B$ **do**

  Generate a random sample via block bootstrap from $\{\boldsymbol{y}_t\}$, that is,

$$\boldsymbol{Y}^{(b)} \equiv \{\mathcal{B}_{j_1}, \dots, \mathcal{B}_{j_L}\}, \quad L = [T/m_{\mathrm{BT}}], \quad j_\ell \overset{i.i.d.}{\sim} \mathrm{Unif}(\mathcal{J}),$$

  where $\mathcal{B}_{j_\ell}$'s are sequentially ordered according to the index ℓ to form a vector time series with sample size T.

  Compute sample CCFs based on $\boldsymbol{Y}^{(b)}$, denoted as $\hat{\rho}_{ij}^{(b)}(\ell)$.

  Compute $\hat{\delta}_{ij}^{(b)}(\ell) = \hat{\rho}_{ij}^{(b)}(\ell) - \hat{\rho}_{ji}^{(b)}(\ell)$ for $1 \le j < i \le k$, $\ell = 1, 2, \dots$

end for

Compute the bootstrap standard error:

$$\widehat{\mathrm{se}}_{\mathrm{BT}}\left(\hat{\rho}_{ij}(\ell) - \hat{\rho}_{ji}(\ell)\right) = \left\{ \frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\delta}_{ij}^{(b)}(\ell) - \bar{\delta}_{ij}(\ell) \right)^2 \right\}^{1/2},$$

where $\bar{\delta}_{ij}(\ell) = B^{-1} \sum_{b=1}^{B} \hat{\delta}_{ij}^{(b)}(\ell)$.

2.3. Multiple Testings Using Benjamini–Hochberg Procedure

In contrast to the Wald test and the maximum statistics that consider a single hypothesis testing, we may consider a multiple-testing approach by breaking down the null hypothesis into multiple hypotheses:

$$H_0^{(\ell,i,j)} : \delta_{ij}(\ell) = 0 \quad \text{vs} \quad H_a^{(\ell,i,j)} : \delta_{ij}(\ell) \ne 0, \quad \text{for } 1 \le j < i \le k,\ \ell = 1, 2, \dots$$

which is asymptotically equivalent to

$$\tilde{H}_0^{(\ell,i,j)} : Z_{\ell ij} \sim \mathcal{N}(0,1) \quad \text{vs} \quad \tilde{H}_a^{(\ell,i,j)} : Z_{\ell,ij} \text{ not } \mathcal{N}(0,1), \quad \text{for } 1 \le j < i \le k,\ \ell = 1, 2, \dots, \tag{14}$$

where $Z_{\ell ij}$ is defined in (10). Similarly to the maximum statistics, this approach focuses on the marginal distribution of each $Z_{\ell ij}$, but ignores the dependence among $Z_{\ell ij}$'s to gain computational efficiency. We adopt the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001) to perform the multiple hypotheses testing for (14) in the following steps:

Step 1: Compute the p-value associated with the individual test $\tilde{H}_0^{(\ell,i,j)}$, that is,

$$p_{\ell ij} = 2\left\{1 - \Phi^{-1}\left(|Z_{\ell ij}|\right)\right\}, \quad 1 \le j < i \le k,\ \ell = 1, 2, \dots, \tag{15}$$

where $\Phi^{-1}(\cdot)$ is the cumulative distribution function of $\mathcal{N}(0, 1)$.

DOI: 10.1111/jtsa.12702

Step 2a (denoted as BH_{ℓ} procedure): For an individual lag ℓ, order the p-values $\{p_{\ell ij} : 1 \le j < i \le k\}$ in (15) by ascending order and denote them as

$$p_{(1)} \le p_{(2)} \le \cdots \le p_{(m)}, \quad \text{where} \quad m = k(k-1)/2.$$

Step 2b (denoted as $\mathrm{BH}_{1:\ell}$ procedure): For multiple lags 1 to ℓ, order the p-values $\{p_{rij} : 1 \le j < i \le k, 1 \le r \le \ell\}$ in (15) by ascending order and denote them as

$$p_{(1)} \le p_{(2)} \le \cdots \le p_{(m)}, \quad \text{where} \quad m = \ell k(k-1)/2.$$

Step 3: Let $j_{\max}$ be the largest index j such that $p_{(j)} \le \frac{j}{m}q$, that is,

$$j_{\max} \equiv \max_{1 \le j \le m} \left\{ j : p_{(j)} \le \frac{j}{m} q \right\},$$

where q is a user-specified upper threshold for the false discovery rate (FDR).
Step 4: Reject the hypothesis in (6) if $j_{\max}$ is not an empty set.
Step 5: Identify the pairs of series associated with $p_{(j)}$ satisfying $j \le j_{\max}$, if any.

When considering the CCF information of a single lag ℓ, the BH_{ℓ} procedure in Step 2a is implemented. When considering the CCF information of multiple lags from 1 to ℓ, the $\mathrm{BH}_{1:\ell}$ procedure in Step 2b is implemented. Besides the rejection decision, BH procedures also identify the alternative set which contains pairs of series with asymmetric CCF structure, under the control of FDR $\le q$ (Benjamini and Hochberg, 1995) at a prespecified level q. These pairs could be informative in an application. See an illustration in Section 4.

3. SIMULATION STUDY

This section presents six experiments to study the proposed methods in testing symmetric cross-correlations for a vector time series. Experiments 1–3 discuss the low-dimensional scenarios whereas Experiment 4 focuses on high-dimensional scenarios. Experiment 5 is a sensitivity analysis of the fourth cumulant assumption. Experiment 6 examines a more general scenario for applying the proposed test in which the idiosyncratic term $\boldsymbol{\epsilon}_t$ is serial dependent while maintaining the absence of cross-sectional dependence. This simulation shows that the proposed tests are applicable to approximate factor models so long as the factor series are mutually independent.

We consider the following data-generating process for Experiments 1, 2, and 4:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{L}\boldsymbol{f}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_k), \\ \boldsymbol{f}_t &= \boldsymbol{\Phi}\boldsymbol{f}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_r), \end{aligned}$$

where $k \ge r$, and the two noise processes $\{\boldsymbol{\epsilon}_t\}$ and $\{\boldsymbol{\eta}_t\}$ are independent. Experiments 3, 5 and 6 use slightly different forms for data generation specified in each corresponding subsection. To generate various CCF structures in the simulation, we consider a random setting for the loading matrix $\boldsymbol{L}$ and three specifications for the AR matrix $\boldsymbol{\Phi}$ as follows:

- The loading matrix $\boldsymbol{L}$:

$$\boldsymbol{L} = [L_{ij}], \quad L_{ij} \equiv \frac{\tilde{L}_{ij}}{\max_{1 \le i \le k, 1 \le j \le r} |\tilde{L}_{ij}|}, \quad \tilde{L}_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1). \tag{16}$$

- Setting for $\mathbf{\Phi}$:
  - (a) Case A (under H_0),

$$\mathbf{\Phi} = \text{diag}[\Phi_{ii}], \quad \Phi_{ii} \overset{i.i.d.}{\sim} U[-0.9, 0.9]. \tag{17}$$

  - (b) Case B (under the alternative H_a),

$$\mathbf{\Phi} = [\Phi_{ij}], \quad \Phi_{ij} = \frac{\tilde{\Phi}_{ij}}{\sqrt{\sum_{1 \le i,j \le r} \tilde{\Phi}_{ij}^2}}, \quad \tilde{\Phi}_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0,1). \tag{18}$$

  - (c) Case C (lower triangular),

$$\mathbf{\Phi} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.4 & 0 \\ \phi & \phi & 0.3 \end{bmatrix}, \quad \phi = 0, 0.1, 0.2, 0.3, 0.4, 0.5. \tag{19}$$

We compare the performance of various test statistics proposed in Section 2, including

- Wald test statistics: $Q_1, Q_2, Q_3, Q_{1:2}, Q_{1:3}$;
- Maximum test statistics: $M_1, M_2, M_3, M_{1:2}, M_{1:3}$;
- Multiple testing procedures: BH_1, BH_2, BH_3, $\text{BH}_{1:2}$, $\text{BH}_{1:3}$.

Other settings in the simulation study include

- type-I error for the Wald tests and the maximum test statistics is set to $\alpha = 0.05$ and the FDR for the BH procedure is set at $q = 0.05$;
- the block bootstrap used to compute the test statistics is set with $B = 100$ and $m_{\text{BT}} = 5$;
- the number of simulation replicates is 500 for each simulation study.

We evaluate both the empirical size and power of the test statistics under the six experiments described below and summarize an overall finding at the end.

3.1. Experiment 1

Experiment 1 considers the scenarios with a fixed $r = \dim(\boldsymbol{f}_t) = 3$, coupled with various dimensions $k = 3, 5, 10, 15, 20$. The $\boldsymbol{L}$ matrix is given in (16), and the data are generated under H_0 using $\mathbf{\Phi}$ in (17) and under H_a using $\mathbf{\Phi}$ in (18) with the sample sizes $T = 300$ and $T = 1000$. The empirical sizes and powers of the test statistics considered, evaluated by the proportion of rejecting H_0 among 500 simulation replicates, are summarized in Figures 1 and 2 for $T = 300$. Similar results for $T = 1000$ are reported in Supporting information.

The empirical size of the BH procedure is fairly close to $q = 0.05$. The maximum statistics also have the empirical size around 0.05 for time series with $k \ge 10$, but they are too conservative for series with $k < 10$. This is understandable because in such cases the number of $Z_{\ell ij}$'s ($k(k-1)/2$) is too small to apply the extreme value approximation in Theorem 2. In contrast, the Wald test statistic has an accurate size at 0.05 for series with small k, but it becomes too conservative for $k \ge 10$. As expected, the sizes of all three test statistics improve as the sample size increases. Regarding testing power, the BH procedure is the most powerful among the three methods for all cases. The maximum statistics are competitive with the BH procedures when $k > 10$. An important finding of this study is that the power of the Wald statistics declines quickly as k increases, suggesting that the Wald statistic is ineffective when $k > 10$.

J. Time Ser. Anal. (2023)
DOI: 10.1111/jtsa.12702

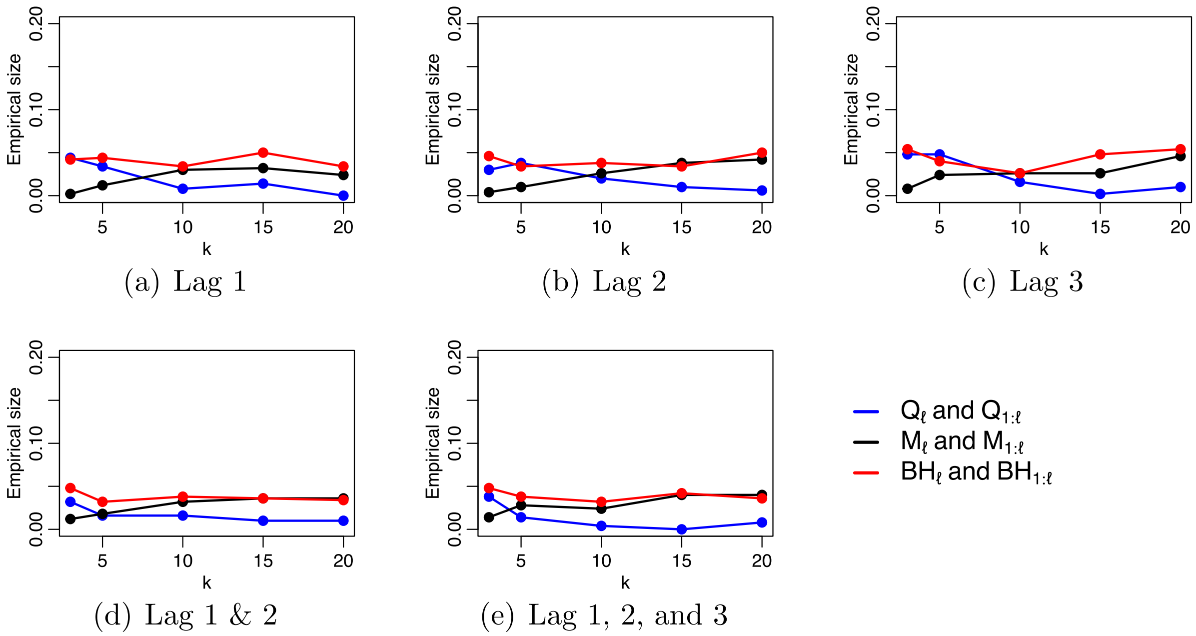


Figure 1. Empirical sizes of test statistics in Experiment 1 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$)

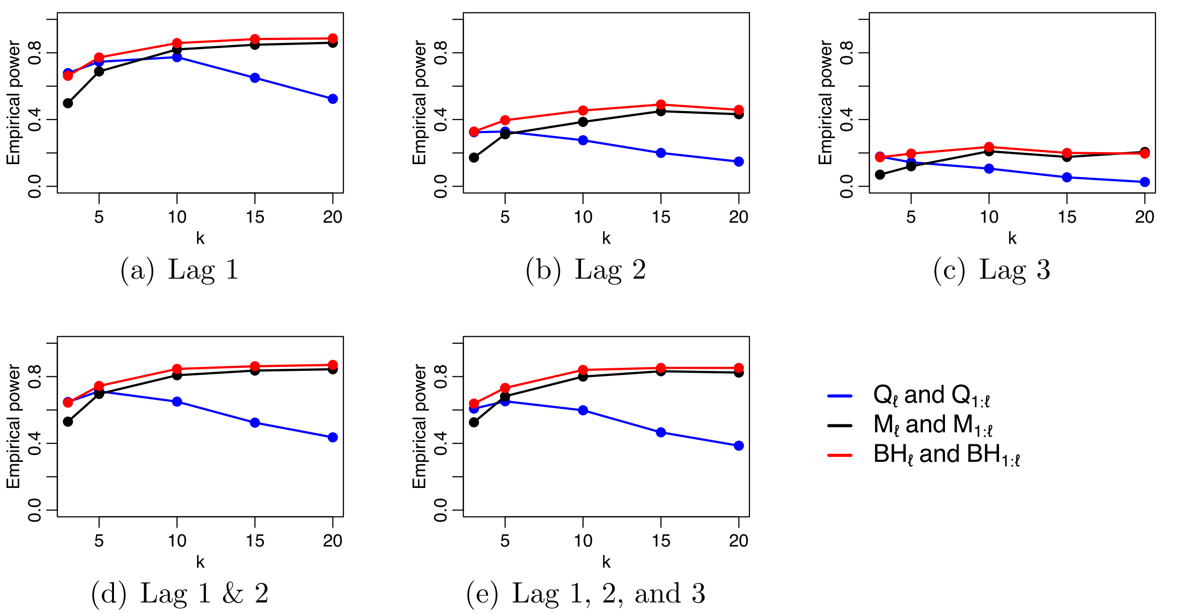
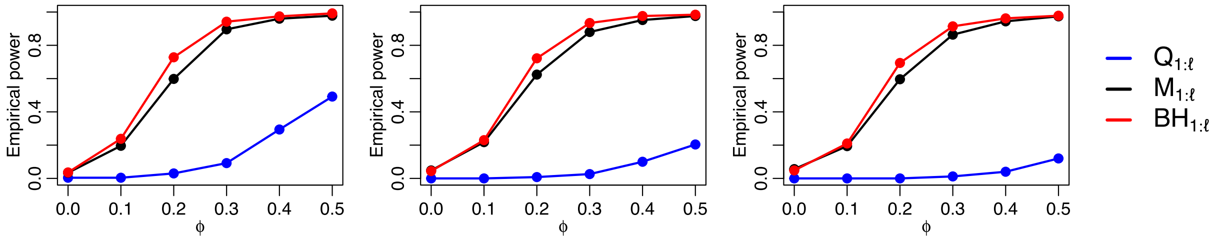


Figure 2. Empirical powers of test statistics in Experiment 1 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$)

DOI: 10.1111/jtsa.12702



(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$

Figure 3. Empirical powers of joint tests in Experiment 2 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$)

3.2. Experiment 2

Experiment 2 considers the scenarios with fixed $r = 3$ and $k = 20$. The data are generated using $\boldsymbol{L}$ in (16) and $\boldsymbol{\Phi}$ in (19). In particular, $\phi = 0$ corresponds to a case of H_0, $\phi \neq 0$ corresponds to a case of H_a, and the magnitude of $|\phi|$ indicates the level of deviating from H_0. Similarly to Experiment 1, two different sample sizes $T = 300$ and $T = 1000$ are considered. Since the performance of tests based on cumulative statistics fares generally better than that based on single-lag statistics, we only report the results of $Q_{1:\ell}$, $M_{1:\ell}$, and $\text{BH}_{1:\ell}$ for the experiments hereafter.

The empirical powers for Experiment 2 are summarized in Figure 3 for $T = 300$. The left-end-point of the power functions indicates that the test size under H_0 is well controlled with $\alpha = 0.05$. Again, the maximum statistics and the BH procedures are competitive. Both methods have good testing powers compared to the Wald test statistics. Simulation results for data with $T = 1000$ show similar conclusions and are reported in Supporting information.

3.3. Experiment 3

Experiment 3 considers dynamic factor models with $k = 10$ and $r = 3$ as an alternative model scenario. The model setting is as follows:

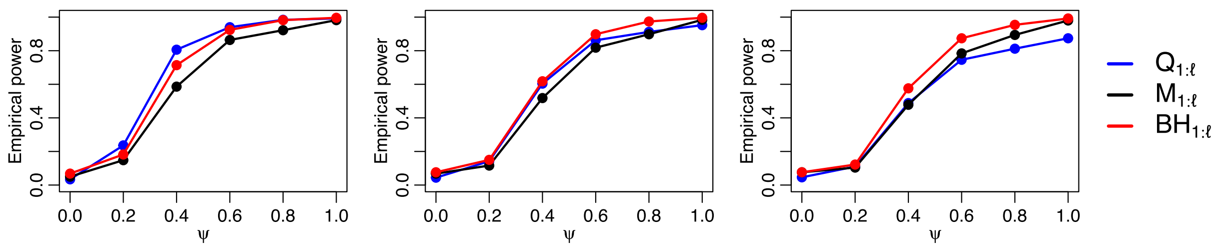
$$\boldsymbol{y}_t = \boldsymbol{L}_1\boldsymbol{f}_t + \psi\boldsymbol{L}_2\boldsymbol{f}_{t-1} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_k),$$
$$\boldsymbol{f}_t = \boldsymbol{\Phi}\boldsymbol{f}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_r).$$

In this experiment, the sample size is set to $T = 300$ and the settings for $\boldsymbol{L} = [\boldsymbol{L}_1, \boldsymbol{L}_2]$ and $\boldsymbol{\Phi}$ follow (16) and Case A in (17) respectively. The value of $|\psi|$ controls the level of deviating from H_0. In particular, $\psi = 0$ corresponds to the case of H_0 and $\psi \neq 0$ corresponds to the case of H_a. The resulting empirical powers for ψ in the range of [0, 1] are summarized in Figure 4. From the plots, the three testing methods are competitive, and all have good testing powers in dynamic factor model scenarios. Similar conclusion holds for data with $T = 1000$ and is reported in Supporting information.

3.4. Experiment 4

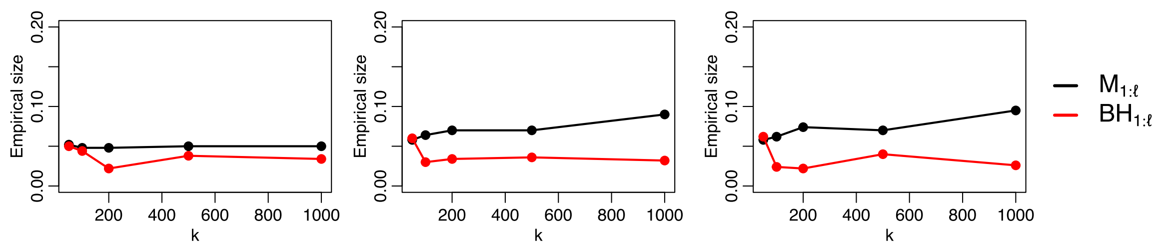
Experiment 4 focuses on high-dimensional scenarios with $k = 50{,}100, 200{,}500, 1000$ and $r = \sqrt{k}/2$ such that $r/k \to 0$ as $k \to \infty$. The sample size is set to $T = 500$, and the settings for $\boldsymbol{L}$ and $\boldsymbol{\Phi}$ are identical to those of Experiment 1. Note that, in this experiment, the series dimension k could be larger than the sample size T. In Experiment 4, we only report the results for the maximum statistics and the BH procedures since implementing the Wald test is too time-consuming for $k > 20$. The resulting empirical sizes and powers are summarized in Figures 5 and 6 respectively. From the plots, the maximum statistics might have some upward size distortion while the BH procedures might have downward size distortion when k is large. This is not surprising because the factor models

   DOI: 10.1111/jtsa.12702



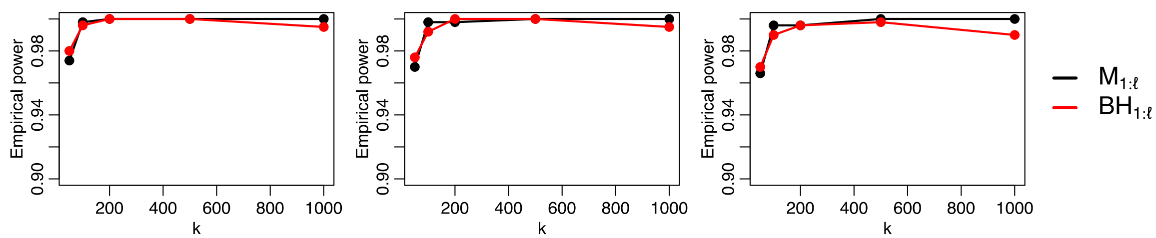
(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$

Figure 4. Empirical powers of joint statistics in Experiment 3 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$)



(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$

Figure 5. Empirical sizes of joint test statistics in Experiment 4 for sample size $T = 500$, based on 500 simulation replicates ($\alpha = q = 0.05$), where k denotes dimension



(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$

Figure 6. Empirical powers of joint test statistics in Experiment 4 for sample size $T = 500$, based on 500 simulation replicates ($\alpha = q = 0.05$), where k denotes dimension

used are AR(1) so that some of the lag-3 cross-correlations might not be statistically significant. The plots also show that both the maximum statistics and the BH procedure have good power in the high-dimensional cases with the maximum statistics faring slightly better when the dimension is 1000.

3.5. Experiment 5

Experiment 5 is a sensitivity analysis to validate the performance of the Wald test when the fourth cumulant assumption is violated. In this experiment, we only consider the H_0 scenario and examine the test size. The data $\boldsymbol{y}_t$ are generated with $r = 3$ and $k = 10$ using $\boldsymbol{L}$ in (16) and $\boldsymbol{\Phi}$ in (17). Instead of Gaussian noises, $\boldsymbol{\epsilon}_t$ follows a Student t-distribution with ν degrees of freedom (denoted as t_ν) in generating $\boldsymbol{y}_t$. As known for t_ν distributions, the fourth moment exists when $\nu > 4$ but the fourth cumulants are not zero. Following this t_ν setting, the cumulant assumption required for the Wald test is violated except for $\nu \to \infty$ corresponding to the Gaussian case that satisfies the cumulant assumption (i.e. all fourth cumulants are zero). In this experiment, the sample size is set to $T = 300$ and ν is ranging from 1 to 10 indicating the level of violation from the cumulant assumption. The results for Experiment 5 are summarized in Figure 7. It turns out that the empirical sizes of all proposed tests are

DOI: 10.1111/jtsa.12702

Journal of Time Series Analysis published by John Wiley & Sons Ltd.

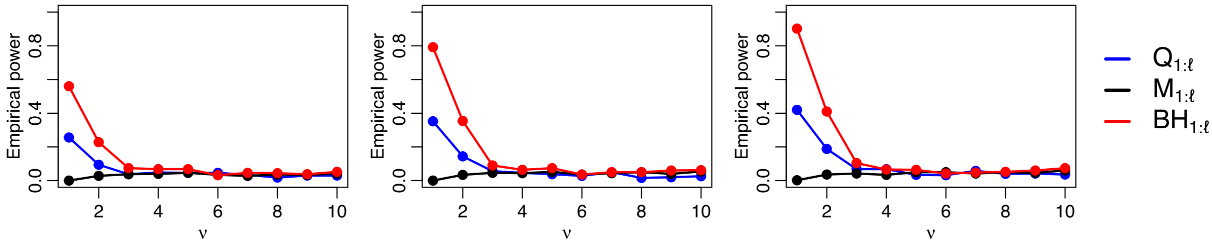


Figure 7. Empirical sizes of test statistics in Experiment 5 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$), where ν denotes the degrees of freedom of t-distribution

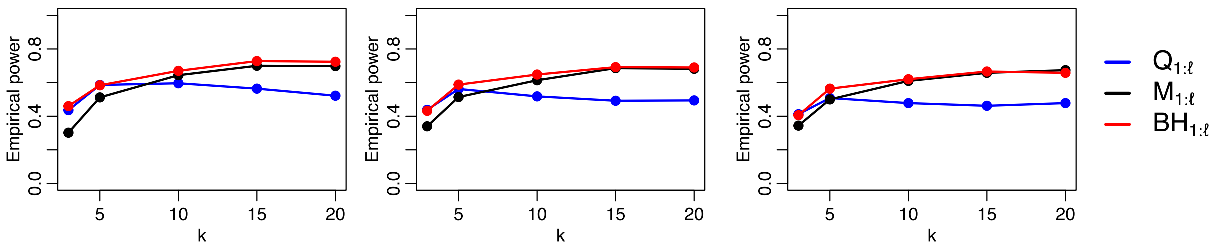


Figure 8. Empirical powers of joint statistics in Experiment 6 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$)

well-controlled at $\alpha = 0.05$ whenever $\nu > 4$. Similar results are found for data with $T = 1000$ and are reported in Supporting information. To conclude, the proposed tests remain effective so long as the fourth moments of the time series $\boldsymbol{y}_t$ exist.

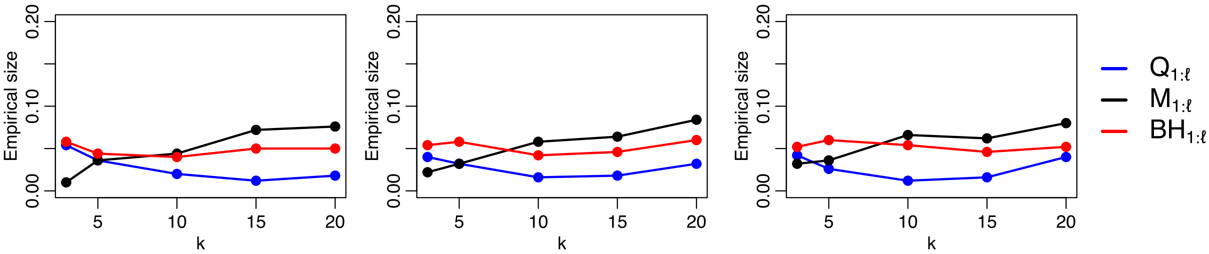
3.6. Experiment 6

Experiment 6 investigates the performance of the test statistics under a broader class in which the idiosyncratic term $\boldsymbol{\epsilon}_t$ is serially dependent, but has no cross-sectional dependence, that is, $\text{cov}(\epsilon_{it}, \epsilon_{jt'}) = 0$ for all $i \neq j$ and $t \neq t'$, and $\text{cov}(\epsilon_{it}, \epsilon_{it'})$ could be non-zero for some $t \neq t'$. The data generating model is identical to that of Experiment 1, except that $\boldsymbol{\epsilon}_t$ is assumed to follow a first-order moving-average (MA(1)) model, given by $\boldsymbol{\epsilon}_t = \boldsymbol{\zeta}_t + 0.5\boldsymbol{\zeta}_{t-1}$, where $\boldsymbol{\zeta}_t$ is an i.i.d. white noise process. In this experiment, the sample size is set to $T = 300$, and $k \in \{3, 5, 10, 15, 20\}$. The empirical sizes and powers of the joint test statistics considered, evaluated by the proportion of rejecting H_0 among 500 simulation replicates, are summarized in Figures 8 and 9 for $T = 300$. The performance of the test statistics used is very similar to that of Experiment 1. This is expected since the proposed test statistics are only relevant to the underlying cross-sectional CCF which is absent under the idiosyncratic MA(1) setting. The simulation results for $T = 1000$ are reported in Supporting information.

3.7. Summary of Simulation

Finally, we use the simulation settings in Experiment 1 to evaluate the computing time for the proposed testing procedures. The time comparison is shown in Figure 10, where the x-axis indicates the dimension k and the y-axis indicates the computing time in minutes (log scale) for a single simulation replicate. From the plot, the BH procedures are faster when the dimension k is small, but the maximum statistics become more computationally efficient as k increases. On the other hand, the computing time for the Wald test grows too fast as k increases and becomes almost infeasible for $k > 40$.

   DOI: 10.1111/jtsa.12702



(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$

Figure 9. Empirical sizes of joint statistics in Experiment 6 for $T = 300$, based on 500 simulation replicates ($\alpha = q = 0.05$)

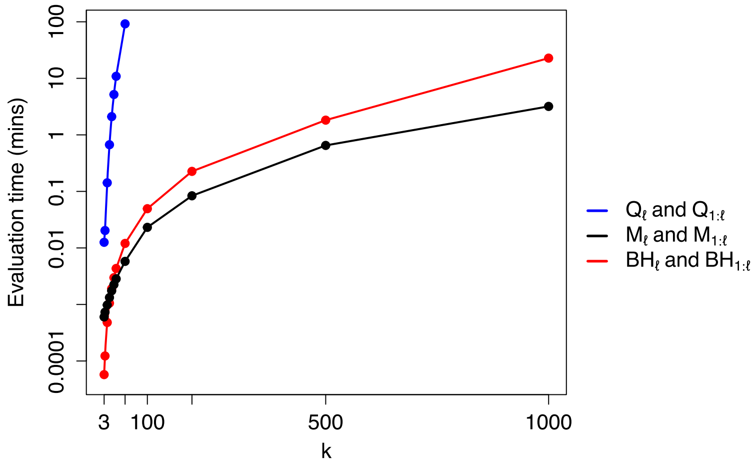


Figure 10. Computing-time evaluation for implementing the proposed testing procedures, where k is the dimension and the y-axis is in log schedule

To summarize, the three proposed methods perform reasonably well in testing CCF symmetry in vector time series. The Wald tests are most useful for low-dimensional time series, say $k < 10$, but become less powerful and even computationally infeasible for high-dimensional time series. The maximum statistics are effective when the series dimension k is sufficiently large to apply the extreme value theory, for example, $k \geq 10$. In particular, it enjoys an advantage in the computation time for high-dimensional settings. The BH procedures fare well consistently in most scenarios considered in our simulation study, but they require heavier computation when k is large, say $k > 100$. Finally, as a byproduct, the BH procedure also identifies the pairs of series with asymmetric CCFs.

Our test statistics are functions of the sample CCFs $\hat{\boldsymbol{\rho}}_\ell$. As expected, the test statistic is more powerful when the sample CCFs involved are strong, for example, Q_1 is generally more powerful than Q_2 and Q_3 since the magnitudes of elements in $\hat{\boldsymbol{\rho}}_1$ is generally higher than those in $\hat{\boldsymbol{\rho}}_2$ and $\hat{\boldsymbol{\rho}}_3$ for non-seasonal series. To mitigate the difficulty in selecting the lags, we recommend that the cumulative statistics $M_{1:\ell}$ and $\mathrm{BH}_{1:\ell}$ be used in applications.

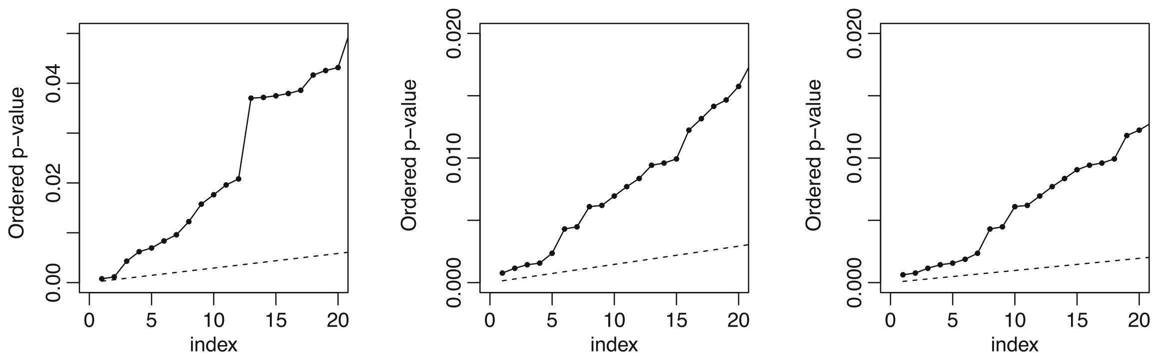
4. APPLICATIONS

We demonstrate the proposed methodology in three applications to assist factor modeling. For all data sets, the block size in the bootstrap steps is set to $m_{\mathrm{BT}} = 5$ in implementing the proposed tests. Our testing results are further confirmed by a parametric approach under a factor model framework with model selection followed by the likelihood ratio test.

Table I. The p-values of Wald test statistics and the maximum statistics for Dataset 1

	Lag ℓ				
Test statistics	1	2	3	4	5
Q_ℓ	0.9935	0.9666	0.9989	0.9981	1.0000
$Q_{1:\ell}$	0.9935	1.0000	1.0000	1.0000	1.0000
M_ℓ	0.2771	0.1750	0.1862	0.0232†	0.0618
$M_{1:\ell}$	0.2771	0.3044	0.4161	0.0650	0.0779

Note: † indicates significance at $\alpha = 0.05$.



(a) BH_1 (b) $\mathrm{BH}_{1:2}$ (c) $\mathrm{BH}_{1:3}$

Figure 11. Ordered p-values of BH procedures for Dataset 1 (FDR controlled at $q = 0.05$)

4.1. Dataset 1

The first dataset used consists of quarterly Gross Domestic Product (GDP) series of 19 European countries from 2000.II to 2018.IV. The data are log-transformed and differenced to remove trends; time plots of the data are shown in Supporting information. These data were used in Peña and Tsay (2021, Example 6.4) to demonstrate dynamic factor model analysis.

We apply the proposed testing methods to the growth rates of sample size $T = 75$. The p-values based on Wald tests and the maximum statistics are reported in Table I. Except for the test statistic M_4, all p-values are greater than 0.05, indicating there is insufficient evidence to reject the CCF symmetry.

We further implement the multiple testing procedure to Dataset 1 to explore the sources of CCF asymmetry in the GDP growth series under FDR controlled at $q = 0.05$. The ordered p-values $p_{(j)}$ based on the test statistics $\{Z_{\ell ij}\}$ compared with the corresponding threshold $(j/m)q$ (dashed line) are shown in Figure 11. In all BH procedures, there is no p-value below the threshold line showing no evidence of CCF asymmetry, which shows an agreement with the results of Wald tests and the maximum statistics.

All testing results lead to the same conclusion that the GDP growth rates do not exhibit strong evidence of asymmetric CCF structure between the series. Consequently, we use an exact factor model for further analysis. This viewpoint is verified via a comprehensive model fitting with model selection. We consider the factor model:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{L}\boldsymbol{f}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\boldsymbol{0}, \sigma_\epsilon^2 \boldsymbol{I}), \\ \boldsymbol{f}_t &= \boldsymbol{\Phi}_1 \boldsymbol{f}_{t-1} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{f}_{t-p} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \end{aligned} \tag{20}$$

where $\boldsymbol{L}'\boldsymbol{L}$ is diagonal and the covariance matrix for $\boldsymbol{\eta}_t$ is set to $\boldsymbol{I}$ for model identifiability. The best fitted model is selected among $p \in \{1, 2, 3, 4\}$ and $r \in \{1, \dots, 5\}$ via Bayesian Information Criterion (BIC) (Schwarz, 1978)

Table II. BIC values of maximum likelihood fittings of factor models for Dataset 1

The number of common factor r	The order of VAR model			
	$p = 1$	$p = 2$	$p = 3$	$p = 4$
$r = 1$	3633.20	3636.92	3640.17	3642.39
$r = 2$	3662.85	3663.50	3677.15	3691.87
$r = 3$	3703.11	3692.37	3722.03	3742.57
$r = 4$	3762.98	3747.06	3776.04	3834.43
$r = 5$	3792.00	3795.24	3865.12	3981.94

Table III. The p-values of the Wald test and the maximum statistics for Dataset 2

Test statistics	Lag ℓ				
	1	2	3	4	5
Q_ℓ	0.1693	0.9662	0.6767	0.9970	0.9900
$Q_{1:\ell}$	0.1693	0.8827	0.9999	1.0000	1.0000
M_ℓ	0.0099†	0.1531	0.0674	0.0552	0.3304
$M_{1:\ell}$	0.0099†	0.0156†	0.0208†	0.0257†	0.0305†

Note: † indicates significance at $\alpha = 0.05$.

calculated by

$$BIC(\mathcal{M}) = -2 \log L(\hat{\theta}_{\mathcal{M}}) + \log(T)\ \mathrm{df}(\mathcal{M}), \tag{21}$$

$$\mathrm{df}(\mathcal{M}) = \mathrm{df}(\boldsymbol{L}) + \sum_{i=1}^{p} \mathrm{df}(\boldsymbol{\Phi}_i) + \mathrm{df}(\boldsymbol{\Sigma}_\epsilon) = kr - r(r-1)/2 + pr^2 + 1, \tag{22}$$

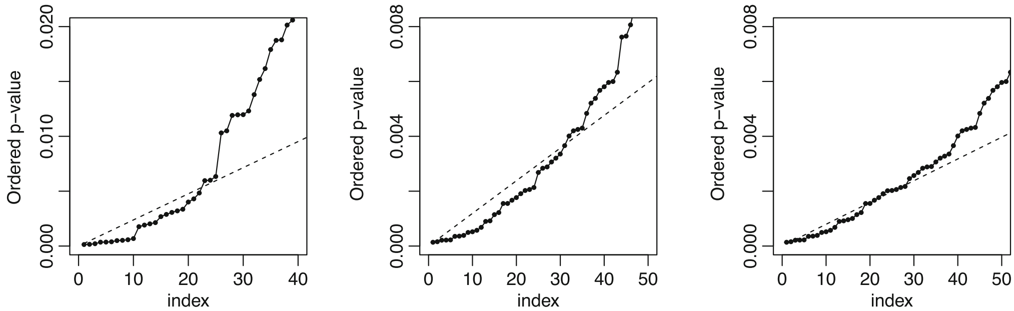
where T is the sample size, $L(\hat{\theta}_{\mathcal{M}})$ is the maximum likelihood (ML) for model $\mathcal{M}$, and $\mathrm{df}(\mathcal{M})$ is the number of parameters in model $\mathcal{M}$.

We implement the ML fittings to the candidate models with BIC selection using the toolbox in R: `MARSS` package (Holmes *et al.*, 2021), resulting in Table II. The best model selected by BIC is $(r, p) = (1, 1)$, which has a single latent factor and verifies the symmetric CCF structure in $\boldsymbol{y}_t$. This comprehensive analysis leads to a consistent result with that obtained by the proposed test statistics.

4.2. Dataset 2

Dataset 2 consists of three variables, GDP, Consumer Price Index (CPI), and the unemployment rate (UR) of G7 countries. The data are quarterly statistics from January 1991 to July 2019, resulting in sample size $T = 115$, and are retrieved from the Federal Reserve Economic Data (FRED) at https://fred.stlouisfed.com. Similarly to Dataset 1, we took the first difference of the series to achieve stationarity, again, shown in Supporting information. The proposed testing methods are applied to the differenced series of Dataset 2 with $k = 21$ and $T = 114$. The p-values based on Wald tests and the maximum statistics are reported in Table III. The maximum statistics find strong evidence of CCF asymmetry for Dataset 2. Although the test statistics Q_ℓ and $Q_{1:\ell}$ find no evidence of rejecting H_0, we realize that, from our simulations, the Wald test is less powerful than the other testing procedures when k is around 20.

To explore the sources of CCF asymmetry in detail, the BH procedures are implemented with FDR controlled at $q = 0.05$. The results are reported in Figure 12. The ordered p-values fall below the threshold (dashed line) showing strong evidence of CCF asymmetry between the series. We further provide the sample CCF plots for the



(a) BH_1 (b) $\mathrm{BH}_{1:2}$ (c) $\mathrm{BH}_{1:3}$

Figure 12. Ordered p-values of BH procedures for Dataset 2 (FDR controlled at $q = 0.05$)

Table IV. BIC values of maximum likelihood fittings of factor models for Dataset 2

The number of common factor r	The order of VAR model			
	$p = 1$	$p = 2$	$p = 3$	$p = 4$
$r = 1$	6309.06	6299.34	6303.37	6307.40
$r = 2$	6054.09	6024.26	6026.52	6022.58
$r = 3$	5973.27	5958.10	5964.57	5980.18
$r = 4$	5943.35	5902.06	5922.95	5965.04
$r = 5$	5909.48	5911.34	5970.47	6046.00
$r = 6$	5916.94	5949.48	6045.38	6162.95

selected pairs of series with the small leading p-values in Supporting information, from which asymmetric CCF structures are easily observed.

The testing results of the maximum statistics and BH procedures reveal strong evidence of asymmetric CCF structure in the series. Therefore, exact factor models are not applicable to this dataset. Finally, we verify this result via a comprehensive model fitting with model selection. Similarly to Dataset 1, we consider the factor model in (20), perform the maximum likelihood fitting, and select the best model using BIC among models with $p \in \{1, 2, 3, 4\}$ and $r \in \{1, \dots, 6\}$. The result is given in Table IV.

The best model selected by BIC is $(r, p) = (4, 2)$. Under this specific model, we further apply the likelihood-ratio test to examine the symmetry of CCFs in a parametric setting:

$$H_0 : \boldsymbol{\theta} \in \Theta_0, \quad \text{vs.} \quad H_a : \boldsymbol{\theta} \in \Theta,$$

where Θ is the parameter space of (20) and Θ_0 is the restricted parameter space to guarantee a symmetric CCF structure for $\{\boldsymbol{y}_t\}$. We reject the null hypothesis if

$$LR = -2\left[\log L(\hat{\boldsymbol{\theta}}_0) - \log L(\hat{\boldsymbol{\theta}})\right] > \chi^2_{\dim(\Theta)-\dim(\Theta_0)}(\alpha),$$

where $\log L(\hat{\boldsymbol{\theta}}_0) = \sup_{\theta\in\Theta_0} \log L(\boldsymbol{\theta})$, $\log L(\hat{\boldsymbol{\theta}}) = \sup_{\theta\in\Theta} \log L(\boldsymbol{\theta})$, and $L(\cdot)$ is the data likelihood. It turns out that this parametric test also leads to a rejection of CCF symmetry for Dataset 2 with the following statistics:

$$\log L(\hat{\boldsymbol{\theta}}) = -2687.7, \quad \dim(\Theta) = 111,$$
$$\log L(\hat{\boldsymbol{\theta}}_0) = -2731.7, \quad \dim(\Theta_0) = kr - r(r-1)/2 + pr + 1 = 87,$$
$$LR = -2\,[-2731.7 - (-2687.7)] = 88 > \chi^2_{111-87}(0.05) = 36.4,$$

DOI: 10.1111/jtsa.12702

Table V. The p-values of the maximum statistics for Dataset 3

	Lag ℓ				
Test statistics	1	2	3	4	5
M_ℓ	0.0000†	0.0002†	0.0034†	0.0000†	0.0027†
$M_{1:\ell}$	0.0000†	0.0000†	0.0000†	0.0000†	0.0000†

Note: † indicates significance at $\alpha = 0.05$.

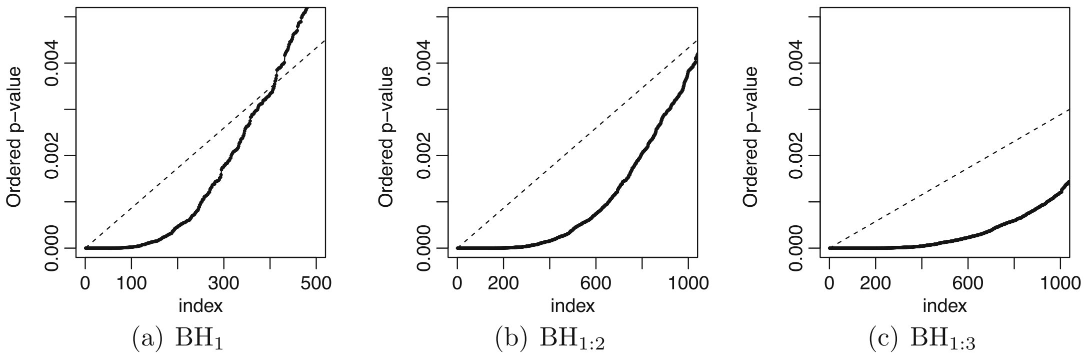


Figure 13. Ordered p-values of the BH procedures for Dataset 3 (FDR controlled at $q = 0.05$)

where dim(Θ) is computed based on (22). In sum, based on either a parametric test or a non-parametric test, the same conclusion suggests that an exact factor model is not suitable for Dataset 2. Other factor models should be employed.

4.3. Dataset 3

The third dataset employed consists of monthly U.S. macroeconomic variables with $k = 108$ time series and $T = 283$ observations. These data were derived from the following resources available in the public domain: FRED at https://fred.stlouisfed.org and Mark Watson at https://www.princeton.edu/~mwatson/publi.html and used in Stock and Watson (2009). The first differenced series of the series, from December 1972 to July 1996, are used in our analysis.

To examine the CCF asymmetry among series, we only apply the maximum statistics and the BH procedures to Dataset 3. The testing results are presented in Table V and Figure 13. Both methods consistently show that there is strong evidence of asymmetric CCF structures among the series. Therefore, exact factor models are not applicable in this particular case; other classes of factor models should be used.

5. CONCLUDING REMARK

In this article, we considered three test statistics for testing the symmetry of the cross-correlation matrices of a vector time series. Based on simulation studies, the Wald test statistics are useful when the dimension k of the time series is small and the sample size T is large. However, they become infeasible when k is large. On the other hand, both the maximum statistics and Benjamini–Hochberg procedures are powerful in detecting the asymmetry in cross-correlation matrices when the dimension is high. They also have good empirical sizes in our simulation studies. The maximum statistics do not fare well when the dimension is low as they depend heavily on the limiting extreme value theory, which requires a large dimension k. On the other hand, the maximum statistics are faster to compute when k is large.

Testing for symmetry in CCFs prior to employing factor models allows users to determine the adequacy of considering orthogonal common factors in a given application. We demonstrated the application of the proposed tests with three real examples, showing that they are indeed useful in assessing the applicability of exact factor models.

Finally, we would like to iterate that various factor models with different parameterizations have been proposed in different scientific fields. The approximate factor models of Chamberlain and Rothschild (1983) allow for weak temporal dependence in the idiosyncratic components, Lam and Yao (2012) and Gao and Tsay (2019, 2022) assume that the idiosyncratic components are serially independent, and Peña and Box (1987) assumes that the factor series are mutually orthogonal. In some factor models, the mutual independence of factor series is not explicitly mentioned. Our simulations show that the proposed test statistics can detect symmetric cross-correlation matrices so long as the factor series are mutually orthogonal. The problem of checking for the applicability of other factor models remains open. We leave it for future research.

DATA AVAILABILITY STATEMENT

These data were derived from the following resources available in the public domain: FRED at https://fred.stlouisfed.org Mark Watson at https://www.princeton.edu/~mwatson/publi.html.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

REFERENCES

Afonja B. 1972. The moments of the maximum of correlated normal and *t*-variates. *Journal of the Royal Statistical Society. Series B (Methodological)* **34**:251–262.

Anderson TW. 1963. The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika* **28**:1–25.

Bai J, Ng S. 2002. Determining the number of factors in approximate factor models. *Econometrica* **70**(1):191–221.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(1):289–300.

Benjamini Y, Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* **29**(4):1165–1188.

Brillinger DR. 1964. *The generalization of the techniques of factor analysis, canonical correlation and principal components to stationary time series*. In *Invited Paper at the Royal Statistical Society Conference in Cardiff, Wales*, UK.

Bühlmann P. 2002. Bootstraps for time series. *Statistical Science* **17**(1):52–72.

Chamberlain G. 1983. Funds, factors, and diversification in arbitrage pricing models. *Econometrica* **51**(5):1305–1323.

Chamberlain G, Rothschild M. 1983. Arbitrage, factor structure in arbitrage pricing models. *Econometrica* **51**(5):1281–1304.

Chang J, Yao Q, Zhou W. 2017. Testing for high-dimensional white noise using maximum cross-correlations. *Biometrika* **104**(1):111–127.

Embrechts P, Klüppelberg C, Mikosch T. 2013. *Modelling Extremal Events: for Insurance and Finance* Springer Science & Business Media, Berlin.

Forni M, Hallin M, Lippi M, Reichlin L. 2000. The generalized dynamic factor model: identification and estimation. *Review of Economics and Statistics* **82**(4):540–554.

Forni M, Hallin M, Lippi M, Reichlin L. 2005. The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association* **100**:830–840.

Gao Z, Tsay RS. 2019. A structural factor approach to modeling high-dimensional time series and space-time data. *Journal of Time Series Analysis* **40**(3):343–362.

Gao Z, Tsay RS. 2022. Modeling high-dimensional time series: a factor model with dynamically dependent factors and diverging eigenvalues. *Journal of the American Statistical Association* **117**:1398–1414.

Geweke J. 1977. *The dynamic factor analysis of economic time series*. In *Latent Variables in Socio-Economic Models*, Aigner DJ, Goldberger AS (eds.). North-Holland, Amsterdam; 365–383.

Hall P, Horowitz JL, Jing B. 1995. On blocking rules for the bootstrap with dependent data. *Biometrika* **82**(3):561–574.

Hannan EJ. 1976. The asymptotic distribution of serial covariances. *Annals of Statistics* **4**(2):396–399.

Holmes E, Ward E, Scheuerell M, Wills K. 2021. *MARSS: multivariate autoregressive state-space modeling*. R package version 3.11.4.

Kunsch HR. 1989. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* **17**(3):1217–1241.
Lahiri SN. 2003. *Resampling Methods for Dependent Data* Springer, New York.
Lam C, Yao Q. 2012. Factor modeling for high dimensional time series: inference for the number of factors. *Annals of Statistics* **40**(2):694–726.
Lam C, Yao Q, Bathia N. 2011. Estimation of latent factors for high-dimensional time series. *Biometrika* **98**(4):901–918.
Melard G, Paesmans M, Roy R. 1991. Consistent estimation of the asymptotic covariance structure of multivariate serial correlations. *Journal of Time Series Analysis* **12**(4):351–361.
Onatski A. 2009. Testing hypotheses about the number of factors in large factor models. *Econometrica* **77**(5):1447–1479.
Peña D, Box G. 1987. Identifying a simplifying structure in time series. *Journal of the American Statistical Association* **82**:836–843.
Peña D, Poncela P. 2006. Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference* **136**(4):1237–1256.
Peña D, Tsay RS. 2021. *Statistical Learning for Big Dependent Data* John Wiley & Son, New York.
Politis DN. 2003. The impact of bootstrap methods on time series analysis. *Statistical Science* **18**(2):219–230.
Roy R. 1989. Asymptotic covariance structure of serial correlations in multivariate time series. *Biometrika* **76**(4):824–827.
Schwarz G. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**(2):461–464.
Stock JH, Watson M. 2009. *Forecasting in dynamic factor models subject to structural instability*. In *The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford; 173–205.
Tsay RS. 2020. Testing serial correlations in high-dimensional time series via extreme value theory. *Journal of Econometrics* **216**(1):106–117.
Xiao H, Wu WB. 2014. Portmanteau test and simultaneous inference for serial covariances. *Statistica Sinica* **24**:577–599.

APPENDIX : ADJUSTMENT FOR DEPENDENCE

In Theorem 2, the normalizing constants $a_{\ell,k}$ and $b_{\ell,k}$ are derived under independent standard normal variates. In our application, the normalized variates $Z_{\ell ij}$'s are dependent in general. We propose an adjustment to the normalizing constants in applying Theorem 2 to account for the dependence among $Z_{\ell ij}$'s for the proposed maximum test statistic. The adjustment is based on a 'working model' for dependent $\mathcal{N}(0,1)$ variates by assuming a compound symmetry dependence structure:

$$\boldsymbol{Z}_n = (Z_1, \dots, Z_n)' \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{C}_n(\tau)\right), \quad \boldsymbol{C}_n(\tau) = \begin{pmatrix} 1 & \tau & \cdots & \tau \\ \tau & 1 & \cdots & \tau \\ \vdots & \ddots & \ddots & \vdots \\ \tau & \cdots & \tau & 1 \end{pmatrix}, \quad \tau \in [0,1).$$

Under this working model, the first and second moments of the maximum statistic $Z_{(n)} = \max_{1\le j\le n} Z_j$ for arbitrary $n > 3$ have closed forms (Afonja, 1972):

$$EZ_{(n)} = \frac{n(n-1)}{2\sqrt{\pi}}\sqrt{1-\tau}\;\Phi_{n-2}(\boldsymbol{0},\ 1/3), \tag{A1}$$

$$EZ^2_{(n)} = 1 + (1-\tau)n(n-1)(n-2)\left(4\pi\sqrt{3}\right)^{-1}\Phi_{n-3}(\boldsymbol{0},\ 1/4), \tag{A2}$$

where $\Phi_h(\cdot, r)$ is the distribution function of a h-dimensional $\mathcal{N}\left(\boldsymbol{0}, \boldsymbol{C}_h(r)\right)$. In particular, $\Phi_{n-2}(\boldsymbol{0},\ 1/3)$ and $\Phi_{n-3}(\boldsymbol{0},\ 1/4)$ can be accurately evaluated via Monte Carlo methods according to

$$\Phi_h(\boldsymbol{0}, r) = E\left[\left(\Phi\left(-\sqrt{\frac{r}{1-r}}Z_1\right)\right)^h\right], \quad Z_1 \sim \mathcal{N}(0,1), \quad \Phi(z) = P(Z_1 \le z).$$

The extreme value theorem generally holds for the maximum $Z_{(n)}$ even under the dependent scenario, that is, with proper normalizing constants $\left(a_n^{\dagger}, b_n^{\dagger}\right)$ such that

$$a_n^{\dagger}\left(Z_{(n)} - b_n^{\dagger}\right) \to_d \Lambda, \quad \text{as} \quad n \to \infty, \tag{A3}$$

where Λ has the standard Gumbel distribution. We suggest to find the suitable $(a_n^{\dagger}, b_n^{\dagger})$ in (A3) by matching the first and second moments in the asymptotics, that is, equating

$$a_n^{\dagger}\left(EZ_{(n)} - b_n^{\dagger}\right) = E\Lambda = \gamma, \quad \left(a_n^{\dagger}\right)^2 var\left(Z_{(n)}\right) = var(\Lambda) = \frac{\pi^2}{6}, \tag{A4}$$

where $\gamma \approx 0.5772$ is the Euler constant. The resulting $a_n^{\dagger}$ and $b_n^{\dagger}$ satisfy

$$a_n^{\dagger} = \frac{\pi}{\sqrt{6}}\left[EZ_{(n)}^2 - (EZ_{(n)})^2\right]^{1/2}, \quad b_n^{\dagger} = EZ_{(n)} - \frac{\gamma}{a_n^{\dagger}}, \tag{A5}$$

in which $EZ_{(n)}$ and $EZ_{(n)}^2$ are given in (A1) and (A2) respectively. The adjusted normalizing constants in (A5) rely on the dimension n and the strength of dependence τ in the working model.

Connecting to our testing problem, our maximum statistic is taken among $Z_{\ell ij}$'s defined in (10) playing the roles of dependent normal variates in $\boldsymbol{Z}_n$ with $n = k(k-1)/2$ for test statistic M_ℓ and $n = \ell k(k-1)/2$ for test statistic $M_{1:\ell}$ respectively. We still need a sensible τ to plug in (A5) for computing $EZ_{(n)}$ and $EZ_{(n)}^2$, discussed after Corollary 2. Now, we are ready to give a modified version of Theorem 2 for adjusting the dependence among $Z_{\ell ij}$'s.

Corollary 2. Consider a sufficiently large k and a small lag $\ell \in \{1, 2, \dots\}$. Let $(a_{\ell,k}^{\dagger}, b_{\ell,k}^{\dagger})$ be $(a_n^{\dagger}, b_n^{\dagger})$ with $n = \ell k(k-1)/2$ in (A5). We reject H_0 at the significance level α if

$$M_\ell > b_{1,k}^{\dagger} - \frac{1}{a_{1,k}^{\dagger}} \ln[-\ln(1-\alpha/2)], \tag{A6}$$

$$M_{1:\ell} > b_{\ell,k}^{\dagger} - \frac{1}{a_{\ell,k}^{\dagger}} \ln[-\ln(1-\alpha/2)]. \tag{A7}$$

To apply Corollary 2 in practice, we need a sensible value of τ from data to resemble the dependence strength among $Z_{\ell ij}$'s. We conducted several experiments to determine τ. Some choices for τ are suggested. Briefly speaking, we first obtain a bootstrap estimate $\boldsymbol{C}^* \equiv \widehat{\text{corr}}_{\text{BT}}(\hat{\boldsymbol{\delta}}_{1:\ell}^*)$ for approximating the correlation matrix of $\hat{\boldsymbol{\delta}}_{1:\ell}$, which is a by-product of the bootstrap procedure in Algorithm 1, and then compute

$$\tau_a = \text{mean of } \left\{|c_{ij}^*|, i \neq j\right\}, \quad \tau_b = \text{median of } \left\{|c_{ij}^*|, i \neq j\right\}, \quad \tau_c = 1 - \frac{n - \lambda_1(\boldsymbol{C}^*)}{n-1},$$

where c_{ij}^*'s are entries of $\boldsymbol{C}^*$ and $\lambda_1(\boldsymbol{C}^*)$ is the leading eigenvalue of $\boldsymbol{C}^*$. In particular, the form of τ_c is inspired by the eigenstructure of a compound symmetry matrix (working model). Then, we evaluate the adjusted normalizing constants $(a_{\ell,k}^{\dagger}, b_{\ell,k}^{\dagger})$ in Corollary 2 by plugging in the values of τ_a, τ_b or τ_c estimated from data.

To study the effects of the adjustment, Experiment 2 in Section 3 is re-examined for the test statistic $M_{1:\ell}$ with the sample size $T = 300$ and $T = 1000$. Other settings are identical to those of Experiment 2 in Section 3.2. The testing power based on the test rules with and without the adjustment on the normalizing constants are displayed in Figure A1 for nine settings with combinations of $k \in \{10, 20, 40\}$ and $\ell \in \{1, 2, 3\}$. We only report the adjusted results based on τ_a in Figure A1 for comparison since the adjusted decision rule performs almost identically

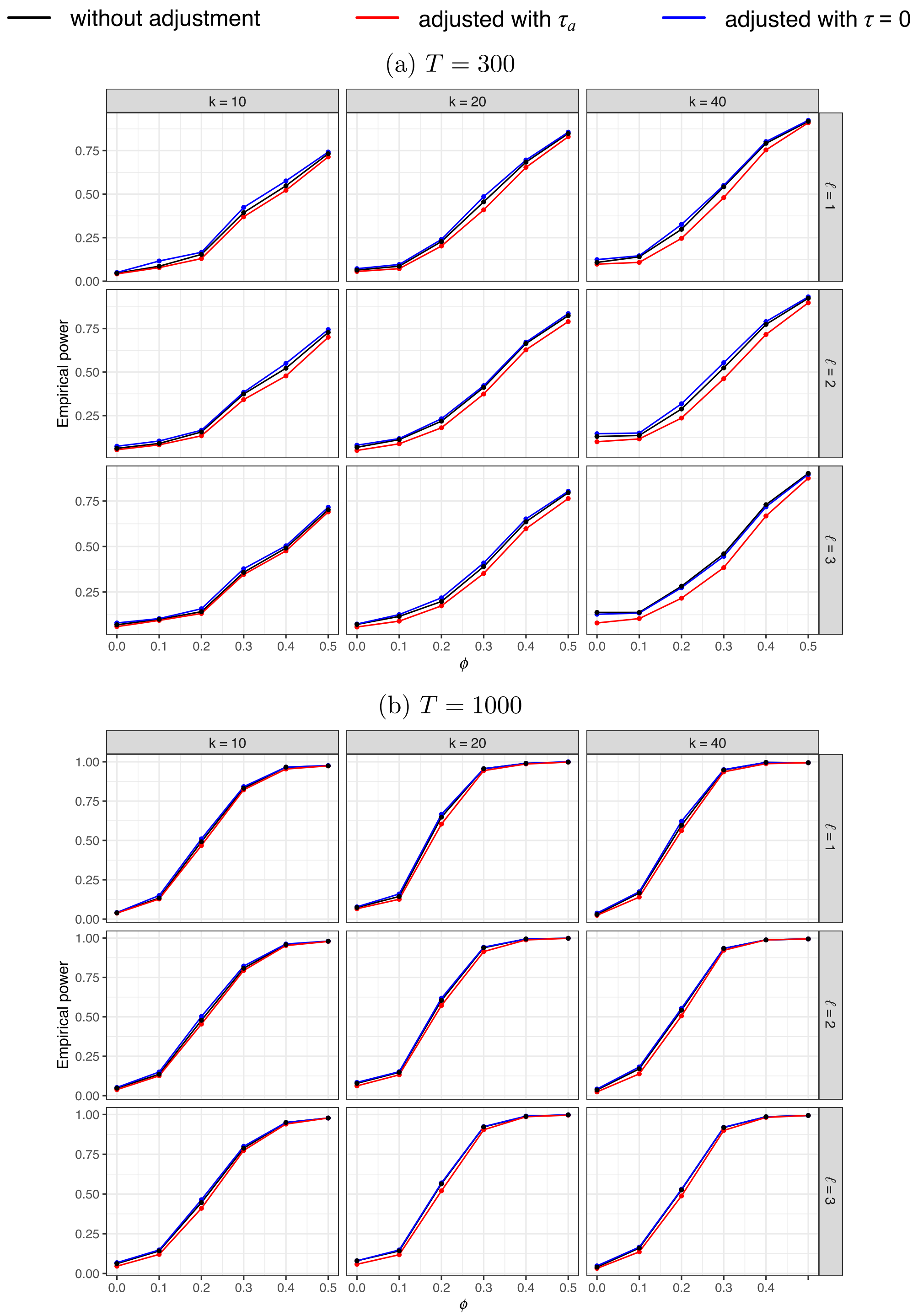


Figure A1. Empirical powers of the test statistic $M_{1:\ell}$ in Experiment 2 with or without adjustment under various settings of (k, ℓ) for data with sample sizes $T = 300$ and $T = 1000$ ($\alpha = 0.05$). The results are based on 500 simulation replicates based on any of $\{\tau_a, \tau_b, \tau_c\}$. As a reference, the adjusted test rule with $\tau = 0$ is included in the comparison, which corresponds to the uncorrelated scenario and behaves equivalently to the test rule without any adjustment. Empirical results show that, the adjusted test rules show minor differences from the test rule without any adjustment in terms of testing power. This is particularly true when the same size is $T = 1000$.

DOI: 10.1111/jtsa.12702