MDPI

Article

A Blockwise Bootstrap-Based Two-Sample Test for High-Dimensional Time Series

Lin Yang

Joint Laboratory of Data Science and Business Intelligence, Southwestern University of Finance and Economics, Chengdu 611130, China; yanglin@smail.swufe.edu.cn

Abstract: We propose a two-sample testing procedure for high-dimensional time series. To obtain the asymptotic distribution of our ℓ_∞-type test statistic under the null hypothesis, we establish high-dimensional central limit theorems (HCLTs) for an α-mixing sequence. Specifically, we derive two HCLTs for the maximum of a sum of high-dimensional α-mixing random vectors under the assumptions of bounded finite moments and exponential tails, respectively. The proposed HCLT for α-mixing sequence under bounded finite moments assumption is novel, and in comparison with existing results, we improve the convergence rate of the HCLT under the exponential tails assumption. To compute the critical value, we employ the blockwise bootstrap method. Importantly, our approach does not require the independence of the two samples, making it applicable for detecting change points in high-dimensional time series. Numerical results emphasize the effectiveness and advantages of our method.

Keywords: two-sample testing; high-dimensional time series; α-mixing; Gaussian approximation; blockwise bootstrap

1. Introduction

A fundamental testing problem in multivariate analysis involves assessing the equality of two mean vectors, denoted as $\boldsymbol{\mu}_X$ and $\boldsymbol{\mu}_Y$. Since its inception by [1], the Hotelling T^2 test has proven to be a valuable tool in multivariate analyses. Subsequently, numerous studies have addressed the testing of $\boldsymbol{\mu}_X = \boldsymbol{\mu}_Y$, within various contexts and under distinct assumptions. See refs. [2,3], along with their respective references.

Consider two sets of observations, $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$, where $\mathbf{X}_t = (X_{t,1}, \ldots, X_{t,p})^\mathsf{T}$ and $\mathbf{Y}_t = (Y_{t,1}, \ldots, Y_{t,p})^\mathsf{T}$. These observations are drawn from two populations with means $\boldsymbol{\mu}_X$ and $\boldsymbol{\mu}_Y$, respectively. The classical test aims to test the hypotheses:

$$H_0 : \boldsymbol{\mu}_X = \boldsymbol{\mu}_Y \quad \text{versus} \quad H_1 : \boldsymbol{\mu}_X \neq \boldsymbol{\mu}_Y. \tag{1}$$

When $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ are two independent sequences and independent with each other, a considerable body of literature focuses on testing Hypothesis (1). The ℓ_2-type test statistic corresponding to (1) is of the form $(\bar{\mathbf{X}} - \bar{\mathbf{Y}})^\mathsf{T} \mathbf{S}^{-1} (\bar{\mathbf{X}} - \bar{\mathbf{Y}})$, where $\bar{\mathbf{X}} = n_1^{-1} \sum_{t=1}^{n_1} \mathbf{X}_t$, $\bar{\mathbf{Y}} = n_2^{-1} \sum_{t=1}^{n_2} \mathbf{Y}_t$ and $\mathbf{S}^{-1}$ is the weight matrix. A straightforward choice for $\mathbf{S}^{-1}$ is the identity matrix $\mathbf{I}_p$ [4,5], implying equal weighting for each dimension. Several classical asymptotic theories have been developed based on this selection of $\mathbf{S}^{-1}$. However, this choice disregards the variability in each dimension and the correlations between them, resulting in suboptimal performance, particularly in the presence of heterogeneity or the existence of correlations between dimensions. In recent decades, numerous researchers have investigated various choices for $\mathbf{S}^{-1}$ along with the corresponding asymptotic theories. See refs. [6,7]. In addition, some researchers have developed a framework centered on ℓ_∞-type test statistics, represented as $\max_{j \in [p]} |(\mathbf{S}^{-1/2}(\bar{\mathbf{X}} - \bar{\mathbf{Y}}))_j|$ [8–10]. Extreme value theory plays a pivotal role in deriving the asymptotic behaviors of these test statistics.

However, when $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ are two weakly dependent sequences and are not independent of each other, the above methods may not work well. In this paper, we introduce an ℓ_∞-type test statistic $T_n := (n_1 n_2)^{1/2}(n_1 + n_2)^{-1/2}|\bar{\mathbf{X}} - \bar{\mathbf{Y}}|_\infty$ for testing H_0 under two dependent sequences. Based on $\mathbf{\Sigma}$, which represents the variance of $(n_1 n_2)^{1/2}(n_1 + n_2)^{-1/2}(\bar{\mathbf{X}} - \bar{\mathbf{Y}})$, we construct a Gaussian maxima, denoted as T_n^{G}, to approximate T_n under the null hypothesis. When $n_1 = n_2 = n$, T_n can be written as $|S_n|_\infty$, the maximum of a sum of high-dimensional weakly dependent random vectors, where $S_n = n^{-1/2}\sum_{t=1}^{n}(\mathbf{X}_t - \mathbf{Y}_t)$. Let $T_n^{\mathrm{G}} = |G|_\infty$ with $G = (G_1, \ldots, G_p)^{\mathrm{T}} \sim \mathcal{N}\{\mathbf{0}, \mathrm{var}(S_n)\}$ and $\mathcal{A}$ be a class of Borel subsets in $\mathbb{R}^p$. Define

$$\rho_n(\mathcal{A}) = \sup_{A \in \mathcal{A}} |\mathbb{P}(S_n \in A) - \mathbb{P}(G \in A)|.$$

Paticularly, let $\mathcal{A}^{\mathrm{max}}$ consists of all sets A^{max} of the form $A^{\mathrm{max}} = \{(a_1, \ldots, a_p)^{\mathrm{T}} \in \mathbb{R}^p : \max_{j \in [p]} |a_j| \le x\}$ with some $x \in \mathbb{R}$. Then we have

$$\rho_n(\mathcal{A}^{\mathrm{max}}) = \sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \le x) - \mathbb{P}(T_n^{\mathrm{G}} \le x)|.$$

Note that $\rho_n(\mathcal{A}^{\mathrm{max}})$ is the Kolmogorov distance between T_n and T_n^{G}.

When dimension p diverges exponentially with respect to the sample size n, several studies have focused on deriving $\rho_n(\mathcal{A}^{\mathrm{max}}) = o(1)$ under a weakly dependent assumption. Based on the coupling method for β-mixing sequence, ref. [11] obtained $\rho_n(\mathcal{A}^{\mathrm{max}}) = o(1)$ under the β-mixing condition, contributing to the understanding of such phenomena. Ref. [12] extended the scope of the investigation to the physical dependence framework introduced by [13]. Considering three distinct types of dependence—namely α-mixing, m-dependence, and physical dependence measures—ref. [14] made significant strides. They established nonasymptotic error bounds for Gaussian approximations of sums involving high-dimensional dependent random vectors. Their analysis encompassed various scenarios of $\mathcal{A}$, including hyper-rectangles, simple convex sets, and sparsely convex sets. Let $\mathcal{A}^{\mathrm{re}}$ be the class of all hyper-rectangles in $\mathbb{R}^p$. Under the α-mixing scenario and some mild regularity conditions, [14] showed

$$\rho_n(\mathcal{A}^{\mathrm{re}}) \lesssim \frac{\{\log(pn)\}^{7/6}}{n^{1/9}},$$

hence the Gaussian approximation holds if $\log(pn) = o(n^{2/21})$. In this paper, under some conditions similar to or even weaker than [14], we obtain

$$\rho_n(\mathcal{A}^{\mathrm{max}}) \lesssim \frac{\{\log(pn)\}^{3/2}}{n^{1/6}},$$

which implies the Gaussian approximation holds if $\log(pn) = o(n^{1/9})$. Refer to Remark 1 for more details on the comparison of the convergence rates. By using the Gaussian-to-Gaussian comparison and Nazarov's inequality for p-dimensional random vectors, we can easily extend our result to $\rho_n(\mathcal{A}^{\mathrm{re}}) \lesssim \{\log(pn)\}^{3/2}n^{-1/6}$. Given that our framework and numerous testing procedures rely on ℓ_∞-type test statistics, we thus propose our results under $\mathcal{A}^{\mathrm{max}}$. When p diverges polynomially with respect to n, to the best of our knowledge, there is no existing literature providing the convergence rate of $\rho_n(\mathcal{A}^{\mathrm{max}})$ for α-mixing sequences under bounded finite moments.

Based on the Gaussian approximation for high-dimensional independent random vectors [15,16], we employ the coupling method for α-mixing sequence [17] and "big-and-small" block technique to specify the convergence rate of $\rho_n(\mathcal{A}^{\mathrm{max}})$ under various divergence rates of p. For more details, refer to Theorem 1 in Section 3.1 and its corresponding proof in Appendix A. Given that $\mathbf{\Sigma}$ is typically unknown in practice, we develop a data-driven procedure based on blockwise wild bootstrap [18] to determine the critical

value for a given significance level α. The blockwise wild bootstrap method is widely used in the time series analysis. See [19,20] and references within.

The independence between $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ is not a necessary assumption in our method. We only require the pair sequence $\{(\mathbf{X}_t, \mathbf{Y}_t)\}$ is weakly dependent. Therefore, our method can be applied effectively to detect change points in high-dimensional time series. Further details on this application can be found in Section 4.

The rest of this paper is organized as follows. Section 2 introduces the test statistic and the blockwise bootstrap method. The convergence rates of Gaussian approximations for high-dimensional α-mixing sequence and the theoretical properties of the proposed test can be found in Section 3. In Section 4, an application to change point detection for high-dimensional time series is presented. The selection method for tuning parameter and a simulation study to investigate the numerical performance of the test are displayed in Section 5. We apply the proposed method to the opening price data from multiple stocks in Section 6. Section 7 provides discussions on the results and outlines our future work. The proofs of the main results in Section 3 are detailed in the Appendices A–D.

Notation: For any positive integer $p \geq 1$, we write $[p] = \{1, \ldots, p\}$. We use $|\mathbf{a}|_\infty = \max_{j \in [p]} |a_j|$ to denote the ℓ_∞-norm of the p-dimensional vector $\mathbf{a}$. Let $\lfloor x \rfloor$ and $\lceil x \rceil$ represent the greatest integer less than or equal to x and the smallest integer greater than or equal to x, respectively. For two sequences of positive numbers $\{a_n\}$ and $\{b_n\}$, we write $a_n \lesssim b_n$ or $b_n \gtrsim a_n$ if $\limsup_{n \to \infty} a_n / b_n \leq c_0$ for some positive constant c_0. Let $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$ hold simultaneously. Denote $\mathbf{0}_p = (0, \ldots, 0)^{\mathrm{T}} \in \mathbb{R}^p$. For any $m \times m$ matrix $\mathbf{A} = (a_{ij})_{m \times m}$, let $|\mathbf{A}|_\infty = \max_{i,j \in [m]} |a_{ij}|$ and $\|\mathbf{A}\|_2$ be the spectral norm of $\mathbf{A}$. Additionally, denote $\lambda_{\min}(\mathbf{A})$ as the smallest eigenvalue of $\mathbf{A}$. Let $1(\cdot)$ be the indicator function. For any $x, y \in \mathbb{R}$, denote $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$. Given $\gamma > 0$, we define the function $\psi_\gamma(x) := \exp(x^\gamma) - 1$ for any $x > 0$. For a real-valued random variable ξ, we define $\|\xi\|_{\psi_\gamma} := \inf[\lambda > 0 : \mathbb{E}\{\psi_\gamma(|\xi|/\lambda)\} \leq 1]$. Throughout the paper, we use $c, C \in (0, \infty)$ to denote two generic finite constants that do not depend on (n_1, n_2, p), and may be different in different uses.

2. Methodology

2.1. Test Statistic and Its Gaussian Analog

Consider two weakly stationary time series $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ and $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ with $\mathbf{X}_t = (X_{t,1}, \ldots, X_{t,p})^{\mathrm{T}}$ and $\mathbf{Y}_t = (Y_{t,1}, \ldots, Y_{t,p})^{\mathrm{T}}$. Let $\boldsymbol{\mu}_X = \mathbb{E}(\mathbf{X}_t)$ and $\boldsymbol{\mu}_Y = \mathbb{E}(\mathbf{Y}_t)$. The primary focus is on testing equality of mean vectors of the two populations:

$$H_0 : \boldsymbol{\mu}_X = \boldsymbol{\mu}_Y \quad \text{versus} \quad H_1 : \boldsymbol{\mu}_X \neq \boldsymbol{\mu}_Y.$$

Given the observations $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$, the estimations of $\boldsymbol{\mu}_X$ and $\boldsymbol{\mu}_Y$ are, respectively, $\hat{\boldsymbol{\mu}}_X = n_1^{-1} \sum_{t=1}^{n_1} \mathbf{X}_t$ and $\hat{\boldsymbol{\mu}}_Y = n_2^{-1} \sum_{t=1}^{n_2} \mathbf{Y}_t$. In this paper, we assume $n_1 \asymp n_2 \asymp n$. It is natural to consider the ℓ_∞-type test statistic $T_n = (n_1 n_2)^{1/2} (n_1 + n_2)^{-1/2} |\hat{\boldsymbol{\mu}}_X - \hat{\boldsymbol{\mu}}_Y|_\infty$. Write $\tilde{n} = \max\{n_1, n_2\}$. Define two new sequences $\{\tilde{\mathbf{X}}_t\}_{t=1}^{\tilde{n}}$ and $\{\tilde{\mathbf{Y}}_t\}_{t=1}^{\tilde{n}}$ with

$$\tilde{\mathbf{X}}_t = \mathbf{X}_{t \wedge n_1} 1(1 \leq t \leq n_1) \quad \text{and} \quad \tilde{\mathbf{Y}}_t = \mathbf{Y}_{t \wedge n_2} 1(1 \leq t \leq n_2).$$

For each $t \in [\tilde{n}]$, let

$$\mathbf{Z}_t = \sqrt{\frac{n_2 \tilde{n}}{n_1(n_1 + n_2)}} \tilde{\mathbf{X}}_t - \sqrt{\frac{n_1 \tilde{n}}{n_2(n_1 + n_2)}} \tilde{\mathbf{Y}}_t.$$

Then, T_n can be rewritten as

$$T_n = \left| \frac{1}{\sqrt{\tilde{n}}} \sum_{t=1}^{\tilde{n}} \mathbf{Z}_t \right|_\infty. \tag{2}$$

We reject the null hypothesis H_0 if $T_n > \text{cv}_\alpha$, where cv_α represents the critical value at the significance level $\alpha \in (0,1)$. Determining cv_α involves deriving the distribution of T_n under H_0. However, due to the divergence of p in a high-dimensional scenario, obtaining the distribution of T_n is challenging. To address this challenge, we employ the Gaussian approximation theorem [15,16]. We seek a Gaussian analog, denoted as T_n^G, satisfying the property that the Kolmogorov distance between T_n and T_n^G converges to zero under H_0. Then, we can replace cv_α by $\text{cv}_\alpha^G := \inf\{x > 0 : \mathbb{P}(T_n^G > x) \le \alpha\}$. Define a p-dimensional Gaussian vector

$$\mathbf{G} \sim \mathcal{N}(\mathbf{0}_p, \Xi_{\tilde{n}}) \quad \text{with} \quad \Xi_{\tilde{n}} = \text{var}\left(\frac{1}{\sqrt{\tilde{n}}} \sum_{t=1}^{\tilde{n}} \mathbf{Z}_t\right). \tag{3}$$

We then define the Gaussian analogue of T_n as

$$T_n^G = |\mathbf{G}|_\infty.$$

Proposition 1 below demonstrates that the null distribution of T_n can be effectively approximated by the distribution of T_n^G.

2.2. Blockwise Bootstrap

Note that the long-run covariance matrix $\Xi_{\tilde{n}}$ specified in (3) is typically unknown. As a result, determining cv_α^G through the distribution of T_n^G becomes challenging. To address this challenge, we introduce a parametric bootstrap estimator for T_n using the blockwise bootstrap method [18].

For some positive constant $\vartheta \in [1/2, 1)$, let $S \asymp \tilde{n}^{1-\vartheta}$ and $B = \lceil \tilde{n}/S \rceil$ be the size of each block and the number of blocks, respectively. Denote $\mathcal{I}_b = \{(b-1)S + 1, \ldots, bS\}$ for $b \in [B-1]$ and $\mathcal{I}_B = \{(B-1)S + 1, \ldots, \tilde{n}\}$. Let $\{\varrho_b\}_{b=1}^B$ be the sequence of i.i.d. standard normal random variables and $\varrho' = (\varrho_1', \ldots, \varrho_{\tilde{n}}')$, where $\varrho_t' = \varrho_b$ if $t \in \mathcal{I}_b$. Define the bootstrap estimator of T_n as

$$\hat{T}_n^G = \left|\frac{1}{\sqrt{\tilde{n}}} \sum_{t=1}^{\tilde{n}} (\mathbf{Z}_t - \bar{\mathbf{Z}})\varrho_t'\right|_\infty,$$

where $\bar{\mathbf{Z}} = \tilde{n}^{-1} \sum_{t=1}^{\tilde{n}} \mathbf{Z}_t$. Based on this estimator, we define the estimated critical value $\hat{\text{cv}}_\alpha$ as

$$\hat{\text{cv}}_\alpha := \inf\{x > 0 : \mathbb{P}(\hat{T}_n^G > x \mid \mathcal{E}) \le \alpha\}, \tag{4}$$

where $\mathcal{E} = \{\mathbf{X}_1, \ldots, \mathbf{X}_{n_1}, \mathbf{Y}_1, \ldots, \mathbf{Y}_{n_2}\}$. Then, we reject the null hypothesis H_0 if $T_n > \hat{\text{cv}}_\alpha$. The procedure for selecting the parameter ϑ (or block size S) is detailed in Section 5.1. In practice, we obtain $\hat{\text{cv}}_\alpha$ through the following bootstrap procedure: Generate K independent sequences $\{\varrho_{(1),t}'\}_{t=1}^{\tilde{n}}, \ldots, \{\varrho_{(K),t}'\}_{t=1}^{\tilde{n}}$, with each $\{\varrho_{(k),t}'\}_{t=1}^{\tilde{n}}$ generated as $\{\varrho_t'\}_{t=1}^{\tilde{n}}$. For each $k \in [K]$, calculate $\hat{T}_{(k),n}^G$ with $\{\varrho_{(k),t}'\}_{t=1}^{\tilde{n}}$. Then, $\hat{\text{cv}}_\alpha$ is the $(1-\alpha)K$-th largest value among $\{\hat{T}_{(1),n}^G, \ldots, \hat{T}_{(K),n}^G\}$. Here, K is the number of bootstrap replications.

3. Theoretical Results

We employ the concept of 'α-mixing' to characterize the serial dependence of $\{(\mathbf{X}_t, \mathbf{Y}_t)\}$, with the α-mixing coefficient at lag κ defined as

$$\alpha(\kappa) := \sup_r \sup_{A \in \mathcal{F}_{-\infty}^r, B \in \mathcal{F}_{r+\kappa}^\infty} |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)|, \tag{5}$$

where $\mathcal{F}_{-\infty}^r$ and $\mathcal{F}_{r+\kappa}^\infty$ are the σ-fields generated by $\{(\mathbf{X}_t, \mathbf{Y}_t) : t \le r\}$ and $\{(\mathbf{X}_t, \mathbf{Y}_t) : t \ge r + \kappa\}$, respectively. We call the sequence $\{(\mathbf{X}_t, \mathbf{Y}_t)\}$ is α-mixing if $\alpha(\kappa) \to 0$ as $\kappa \to \infty$.

3.1. Gaussian Approximation for High-Dimensional α-Mixing Sequence

To show that the Kolmogorov distance between T_n and T_n^{G} converges to zero under various divergence rates of p, we need the following central limit theorems for high-dimensional α-mixing sequence.

Theorem 1. *Let $\{\boldsymbol{\xi}_t\}_{t=1}^n$ be an α-mixing sequence of p-dimensional centered random vectors and $\{\alpha(\kappa)\}_{\kappa \geq 1}$ denote the α-mixing coefficients of $\{\boldsymbol{\xi}_t\}$, defined in the same manner as (5). Write $\mathbf{S}_n = (S_{n,1}, \ldots, S_{n,p})^{\mathsf{T}} = n^{-1/2} \sum_{t=1}^n \boldsymbol{\xi}_t$ and $\mathbf{W} = (W_1, \ldots, W_p)^{\mathsf{T}} \sim \mathcal{N}(\mathbf{0}_p, \boldsymbol{\Sigma}_n)$ with $\boldsymbol{\Sigma}_n = \mathbb{E}(\mathbf{S}_n \mathbf{S}_n^{\mathsf{T}})$. Define*

$$\rho_n = \sup_{x \in \mathbb{R}} |\mathbb{P}(|\mathbf{S}_n|_\infty \leq x) - \mathbb{P}(|\mathbf{W}|_\infty \leq x)|.$$

(i)   If $\max_{t \in [n]} \max_{j \in [p]} \mathbb{E}(|\xi_{t,j}|^m) \leq C_1^*$, $\alpha(\kappa) \leq C_2^* \kappa^{-\tau}$ and $\lambda_{\min}(\boldsymbol{\Sigma}_n) \geq C_3^*$ for some $m > 3$, $\tau > \max\{2m/(m-3), 3\}$ and constants $C_1^*, C_2^*, C_3^* > 0$, we have

$$\rho_n \lesssim \frac{p^{1/2}(\log p)^{1/4}}{n^{\tilde{\tau}}}$$

provided that $p = o(n^{2\tilde{\tau}})$, where $\tilde{\tau} = \tau/(11\tau + 12)$.

(ii) If $\max_{t \in [n]} \max_{j \in [p]} \|\xi_{t,j}\|_{\psi_{\gamma_1}} \leq M_n$, $\alpha(\kappa) \leq C_1^{**} \exp(-C_2^{**} \kappa^{\gamma_2})$ and $\min_{j \in [p]}(\boldsymbol{\Sigma}_n)_{j,j} \geq C_3^{**}$ for some $M_n \geq 1$, $\gamma_1 \in (0, 2]$, $\gamma_2 > 0$ and constants $C_1^{**}, C_2^{**}, C_3^{**} > 0$, we have

$$\rho_n \lesssim \frac{M_n \{\log(pn)\}^{\max\{(2\gamma_2+1)/2\gamma_2, 3/2\}}}{n^{1/6}}$$

provided that $\{\log(pn)\}^3 = o\{n^{\gamma_1\gamma_2/(2\gamma_1+2\gamma_2-\gamma_1\gamma_2)}\}$ and $M_n^2 \{\log(pn)\}^{1/\gamma_2} = o(n^{1/3})$.

Remark 1. *In scenarios where the dimension p diverges polynomially with respect to n, Theorem 1(i) represents a novel contribution to the existing literature. Moreover, if $\tau \to \infty$(i.e., $\alpha(\kappa) \lesssim \exp(-C\kappa)$ for some constant $C > 0$), we have $\tilde{\tau} \to 1/11$, and thus $\rho_n = o(1)$ if $p(\log p)^{1/2} = o(n^{2/11})$. Compared with Theorem 1 in [14], which provides a Gaussian approximation result when p diverges exponentially with respect to n, Theorem 1(ii) has three improvements. Firstly, all conditions of Theorem 1(ii) are equivalent to those in Theorem 1 of [14], with the exception that we permit $\gamma_1 \in (0,1)$, thereby offering a weaker assumption that is more broadly applicable. Secondly, the convergence rate dependent on n via $n^{-1/6}$ in Theorem 1(ii) outperforms the rate of $n^{-1/9}$ demonstrated in Theorem 1 of [14]. Note that the convergence rate in Theorem 1 of [14] can be rewritten as*

$$\left[\frac{M_n\{\log(pn)\}^{(2\gamma_2+1)/2\gamma_2}}{n^{1/6}}\right]^{2/3} + \frac{M_n\{\log(pn)\}^{7/6}}{n^{1/9}}.$$

To ensure $\rho_n = o(1)$, in our result, it is necessary to allow $M_n^6\{\log(pn)\}^{(6\gamma_2+3)/\gamma_2} = o(n)$ when $\gamma_2 \leq 2/3$ and $M_n^6\{\log(pn)\}^{\max\{(6\gamma_2+3)/\gamma_2, 9\}} = o(n)$ when $\gamma_2 > 2/3$, respectively. Comparatively, the basic requirements under Theorem 1 of [14] are $M_n^6\{\log(pn)\}^{(6\gamma_2+3)/\gamma_2} = o(n)$ when $\gamma_2 \leq 2/3$ and $M_n^9\{\log(pn)\}^{21/2} = o(n)$ when $\gamma_2 > 2/3$, respectively. Due to $(6\gamma_2+3)/\gamma_2 < 21/2$ when $\gamma_2 > 2/3$, our result permits a larger or equal divergence rate of p compared with Theorem 1 in [14].

3.2. Theoretical Properties

In order to derive the theoretical properties of T_n, the following regular assumptions are needed.

Assumption 1.

(i)   For some $m > 4$, there exists a constant $C_1 > 0$ s.t. $\max_{t \in [\tilde{n}]} \max_{j \in [p]} \mathbb{E}(|Z_{t,j}|^m) \leq C_1$.
(ii) There exists a constant $C_2 > 0$ s.t. $\alpha(\kappa) \leq C_2 \kappa^{-\tau}$ for some $\tau > 3m/(m-4)$.

(iii)   *There exists a constant $C_3 > 0$ s.t. $\lambda_{\min}(\Xi_{\tilde{n}}) \geq C_3$.*

Assumption 2.

(i)   *There exists a constant $C_1' > 0$ s.t. $\max_{t \in [\tilde{n}]} \max_{j \in [p]} \|Z_{t,j}\|_{\psi_2} \leq C_1'$.*
(ii)   *There exist two constants $C_2', C_3' > 0$ s.t. $\alpha(\kappa) \leq C_2' \exp(-C_3'\kappa)$.*
(iii)   *There exists a constant $C_4' > 0$ s.t. $\min_{j \in [p]} (\Xi_{\tilde{n}})_{j,j} \geq C_4'$.*

Remark 2. *The two mild Assumptions, 1 and 2, delineate the necessary assumptions for $\{(\mathbf{X}_t, \mathbf{Y}_t)\}$ to facilitate the development of Gaussian approximation theories for the dimension p divergence, characterized by polynomial and exponential rates relative to the sample size n, respectively. Assumptions 1(i) and 1(ii) are common assumptions in multivariate time series analysis. Due to $n_1 \asymp n_2 \asymp n$, if $\max_{t \in [n_1], j \in [p]} \mathbb{E}(|X_{t,j}|^m) \leq C$ and $\max_{t \in [n_2], j \in [p]} \mathbb{E}(|Y_{t,j}|^m) \leq C$, then Assumption 1(i) holds, as verified by the triangle inequality. Additionally, Assumption 1(iii) necessitates the strong nondegeneracy of $\Xi_{\tilde{n}}$, a requirement commonly assumed in Gaussian approximation theories (see refs. [21,22], among others). Note that Assumption 2(iii) is implied by Assumption 1(iii). The latter assumption only necessitates the nondegeneracy of $\min_{j \in [p]} \text{var}(\tilde{n}^{-1/2} \sum_{t=1}^{\tilde{n}} Z_{t,j})$. We can modify Assumption 2(i) to $\max_{t \in [\tilde{n}]} \max_{j \in [p]} \|Z_{t,j}\|_{\psi_\gamma} \leq C$ for any $\gamma \in (0, 2]$, a standard assumption in the literature on ultra-high-dimensional data analysis. This assumption ensures subexponential upper bounds for the tail probabilities of the statistics in question when $p \gg n$, as discussed in [23,24]. The requirement of sub-Gaussian properties in Assumption 2(i) is made for the sake of simplicity. If $\{\mathbf{X}_t\}$ and $\{\mathbf{Y}_t\}$ share the same tail probability, Assumption 2(i) is satisfied automatically. Assumption 2(ii) necessitates that the α-mixing coefficients decay at an exponential rate.*

Write $\Delta_n := \max\{n_1, n_2\} - \min\{n_1, n_2\}$. Define two cases with respect to the distinct divergence rates of p as

- Case1: $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ satisfy Assumption 1, and the dimension p satisfies $p^2 \log p = o\{n^{4\tau/(11\tau+12)}\}$ and $\Delta_n^2 \log p = o(n)$;
- Case2: $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ satisfy Assumption 2, and the dimension p satisfies $\log(pn) = o(n^{1/9})$ and $\Delta_n^2 \log p = o(n)$.

Note that $\Delta_n^2 \log p = o(n)$ mandates the maximum difference between the two sample sizes. Proposition 1 below demonstrates that, under the aforementioned cases and H_0, the Kolmogorov distance between T_n and T_n^{G} converges to zero as the sample size approaches infinity. Proposition 1 can be directly derived from Theorem 1. Note that, in the scenario where the dimension p diverges in a polynomial rate with respect to n, obtaining Proposition 1 requires only $m > 3$ and $\tau > \max\{2m/(m-3), 3\}$, an assumption weaker than Assumption 1. The more stringent restrictions $m > 4$ and $\tau > 3m/(m-4)$ in Assumption 1 are imposed to establish the results presented in Theorems 2 and 3.

Proposition 1. *In either* Case1 *or* Case2*, it holds under the null hypothesis H_0 that*

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \mathbb{P}(T_n^{\text{G}} \leq x)| = o(1).$$

According to Proposition 1, the critical value cv_α can be substituted with $\text{cv}_\alpha^{\text{G}}$. However, in practical scenarios, the long-run covariance $\Xi_{\tilde{n}}$ defined in (3) is typically unknown. This implies that obtaining $\text{cv}_\alpha^{\text{G}}$ directly from the distribution of T_n^{G} is not feasible. We introduce a bootstrap method for obtaining the estimator $\hat{\text{cv}}_\alpha$ defined in (4). In situations where the dimension p diverges at a polynomial rate relative to the sample size n, we require an additional Assumption 3 to ensure that $\hat{\text{cv}}_\alpha$ serves as a reliable estimator for cv_α. Assumption 3 places restrictions on the cumulant function, a commonly assumed criterion in time series analysis. Refer to [25,26] for examples of such assumptions in the literature.

Assumption 3. *For each $i, j \in [p]$, define $\mathrm{cum}_{i,j}(h, t, s) = \mathrm{cov}(\mathring{Z}_{0,i}\mathring{Z}_{h,j}, \mathring{Z}_{t,i}\mathring{Z}_{s,j}) - \gamma_{t,i,i}\gamma_{s-h,j,j} - \gamma_{s,i,j}\gamma_{t-h,j,i}$, where $\gamma_{h,i,j} = \mathrm{cov}(Z_{0,i}, Z_{h,j})$ and $\mathring{Z}_{t,j} = Z_{t,j} - \mathbb{E}(Z_{t,j})$. There exists a constant $C_4 > 0$ s.t.*

$$\max_{i,j \in [p]} \sum_{h=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} |\mathrm{cum}_{i,j}(h, t, s)| < C_4.$$

Similar to Case1 and Case2, we consider two cases corresponding to different divergence rates of the dimension p, as outlined below:

- Case3: $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ satisfy Assumptions 1 and 3.
- Case4: $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ satisfy Assumption 2.

Theorem 2. *In either Case3 with $p \log p = o[n^{\min\{(1-\vartheta)/4, 2\tau/(11\tau+12)\}}]$ and $\Delta_n^2 \log p = o(n)$, or Case4 with $\log(pn) = o[n^{\min\{(1-\vartheta)/2, \vartheta/7, 1/9\}}]$ and $\Delta_n^2 \log p = o(n)$, it holds under H_0 that $\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \mathbb{P}(\hat{T}_n^{\mathrm{G}} \leq x \mid \mathcal{E})| = o_p(1)$. Moreover, it holds under H_0 that*

$$\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) \to \alpha \quad as \quad n \to \infty.$$

Theorem 3. *In either Case3 with $p = o\{n^{(1-\vartheta)/4}\}$ or Case4 with $\log(pn) = o[n^{\min\{\vartheta/3, (1-\vartheta)/2\}}]$, if $\max_{j \in [p]} |\mu_{X,j} - \mu_{Y,j}| \gg n^{-1/2}(\log p)^{1/2}$, it holds that*

$$\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) \to 1 \quad as \quad n \to \infty.$$

Remark 3. *The different requirements for the divergence rates of p follow from the fact that we do not rely on the Gaussian approximation and comparison results under certain alternative hypotheses. By Theorem 2 and Theorem 3, the optimal selections for ϑ are $1/2$ and $7/9$ in Case3 and Case4, respectively. This implies that $\lim_{n\to\infty}\mathbb{P}_{H_0}(T_n > \hat{\mathrm{cv}}_\alpha) = \alpha$ holds with $p \log p = o(n^{1/8})$ in Case3 and $\log(pn) = o(n^{1/9})$ in Case4. Under certain alternative hypotheses, $\lim_{n\to\infty}\mathbb{P}_{H_1}(T_n > \hat{\mathrm{cv}}_\alpha) = 1$ holds with $p = o(n^{1/8})$ in Case3 and $\log(pn) = o(n^{1/9})$ in Case4.*

4. Application: Change Point Detection

In this section, we elaborate that our two-sample testing procedure can be regarded as a novel method for detecting change points for high-dimensional time series. To illustrate, we provide a notation for the detection of a single change point, with the understanding that it can be easily extended to the multiple change points case.

Consider a p-dimensional time series $\{\mathbf{X}_t\}_{t=1}^n$. Let $\boldsymbol{\mu}_t = \mathbb{E}(\mathbf{X}_t)$. Consider the following hypothesis testing problem:

$$H_0' : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_n \quad \text{versus} \quad H_1' : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_{\tau_0-1} \neq \boldsymbol{\mu}_{\tau_0} = \cdots = \boldsymbol{\mu}_n.$$

Here, τ_0 is the unknown change point. Let w be a positive integer such that $w < \min\{\tau_0, n - \tau_0\}$. We define $\bar{\boldsymbol{\mu}}^t = w^{-1} \sum_{l=t-w/2+1}^{t+w/2} \boldsymbol{\mu}_l$, $\bar{\boldsymbol{\mu}}^{(1)} = w^{-1} \sum_{l=1}^{w} \boldsymbol{\mu}_l$ and $\bar{\boldsymbol{\mu}}^{(2)} = w^{-1} \sum_{l=n-w+1}^{n} \boldsymbol{\mu}_l$. Then for each $t \in [3w/2, n - 3w/2]$, define $\Delta^{t,(1)} = \bar{\boldsymbol{\mu}}^t - \bar{\boldsymbol{\mu}}^{(1)}$ and $\Delta^{t,(2)} = \bar{\boldsymbol{\mu}}^t - \bar{\boldsymbol{\mu}}^{(2)}$. Thus,

$$\Delta^{t,(1)} = \begin{cases} \mathbf{0}_p & , \text{ if } 3w/2 \leq t \leq \tau_0 - w/2, \\ (\bar{\boldsymbol{\mu}}^{(2)} - \bar{\boldsymbol{\mu}}^{(1)})\frac{t+w/2-\tau_0}{w} & , \text{ if } \tau_0 - w/2 < t \leq \tau_0 + w/2, \\ \bar{\boldsymbol{\mu}}^{(2)} - \bar{\boldsymbol{\mu}}^{(1)} & , \text{ if } \tau_0 + w/2 < t \leq n - 3w/2, \end{cases}$$

$$\Delta^{t,(2)} = \begin{cases} \bar{\boldsymbol{\mu}}^{(1)} - \bar{\boldsymbol{\mu}}^{(2)} & , \text{ if } 3w/2 \leq t \leq \tau_0 - w/2, \\ (\bar{\boldsymbol{\mu}}^{(1)} - \bar{\boldsymbol{\mu}}^{(2)})\frac{-t+w/2+\tau_0}{w} & , \text{ if } \tau_0 - w/2 < t \leq \tau_0 + w/2, \\ \mathbf{0}_p & , \text{ if } \tau_0 + w/2 < t \leq n - 3w/2. \end{cases}$$

Assume $|\bar{\boldsymbol{\mu}}^{(1)} - \bar{\boldsymbol{\mu}}^{(2)}|_\infty = O(1)$, which represents the sparse signals case. Define $t_1(\varepsilon^{t,(1)}) = \min\{t \in [3w/2, n - 3w/2] : |\Delta^{t,(1)}| > \varepsilon^{t,(1)}\}$ and $t_2(\varepsilon^{t,(2)}) = \max\{t \in [3w/2, n - 3w/2] :$

$|\Delta^{t,(2)}| > \varepsilon^{t,(2)}\}$ with two well-defined thresholds $\varepsilon^{t,(1)}, \varepsilon^{t,(2)} \geq 0$. Due to the symmetry of $|\Delta^{t,(1)}|$ and $|\Delta^{t,(2)}|$, it holds under H_1' that

$$\tau_0 = \frac{t_1(\varepsilon^{t,(1)}) + t_2(\varepsilon^{t,(2)})}{2}.$$

The sample estimators of $\bar{\mu}^t$, $\bar{\mu}^{(1)}$ and $\bar{\mu}^{(2)}$ are, respectively, $\hat{\mu}^t = w^{-1}\sum_{l=t-w/2+1}^{t+w/2}\mathbf{X}_l$, $\hat{\mu}^{(1)} = w^{-1}\sum_{l=1}^{w}\mathbf{X}_l$ and $\hat{\mu}^{(2)} = w^{-1}\sum_{l=n-w+1}^{n}\mathbf{X}_l$. Based on the method proposed in Section 2, with $n_1 = n_2 = w$, we define the following two test statistics:

$$T_w^{t,(1)} = \sqrt{w}|\hat{\mu}^t - \hat{\mu}^{(1)}|_\infty \text{ and } T_w^{t,(2)} = \sqrt{w}|\hat{\mu}^t - \hat{\mu}^{(2)}|_\infty.$$

Given a significance level $\alpha > 0$, we choose $\varepsilon^{t,(1)} = \mathrm{cv}_{1\alpha}^t$ and $\varepsilon^{t,(2)} = \mathrm{cv}_{2\alpha}^t$, where $\mathrm{cv}_{1\alpha}^t$ and $\mathrm{cv}_{2\alpha}^t$ are, respectively, the $(1-\alpha)$-quantiles of the distributions of $T_w^{t,(1)}$ and $T_w^{t,(2)}$. The estimated critical values $\hat{\mathrm{cv}}_{1\alpha}^t$ and $\hat{\mathrm{cv}}_{2\alpha}^t$ can be obtained by (4). Thus, $\hat{t}_1 = \min\{t \in [3w/2, n-3w/2] : T_w^{t,(1)} > \hat{\mathrm{cv}}_{1\alpha}^t\}$ and $\hat{t}_2 = \max\{t \in [3w/2, n-3w/2] : T_w^{t,(2)} > \hat{\mathrm{cv}}_{2\alpha}^t\}$. Hence, the estimator of τ_0 is given by

$$\hat{\tau}_0 = \frac{\hat{t}_1 + \hat{t}_2}{2}.$$

We utilize $T_w^{t,(1)}$ as an illustrative example to elucidate the applicability of our proposed method. Let w be an even integer. For any $t \in [5w/2, n-3w/2]$, we have $T_w^{t,(1)} = |w^{-1/2}\sum_{l=1}^{w}(\mathbf{X}_{t-w/2+l} - \mathbf{X}_l)|_\infty$, where the sequence $\{\mathbf{X}_{t-w/2+l} - \mathbf{X}_l\}_{l=1}^{w}$ possesses the same weakly dependence properties and similar moment/tail conditions as $\{\mathbf{X}_l\}_{l=1}^{n}$. For $t \in [3w/2, 5w/2-1]$, let $\{\tilde{\mathbf{X}}_l\}_{l=1}^{t-w/2}$ be defined as $\tilde{\mathbf{X}}_l = \mathbf{X}_l$ when $l \in [1, w]$ and $\tilde{\mathbf{X}}_l = \mathbf{0}_p$ when $l \in [w+1, t-w/2]$. Additionally, define $\{\tilde{\mathbf{Y}}_l\}_{l=t-w/2+1}^{2t-w}$ as $\tilde{\mathbf{Y}}_l = \mathbf{X}_l$ when $l \in [t-w/2+1, t+w/2]$ and $\tilde{\mathbf{Y}}_l = \mathbf{0}_p$ when $l \in [t+w/2+1, 2t-w]$. Then, $T_w^{t,(1)}$ can be expressed as $|w^{-1/2}\sum_{l=1}^{t/2-w/4}\{(\tilde{\mathbf{Y}}_{t-w/2+l} - \tilde{\mathbf{X}}_l) + (\tilde{\mathbf{Y}}_{2t-w+1-l} - \tilde{\mathbf{X}}_{t-w/2+1-l})\}|_\infty$, and $\{(\tilde{\mathbf{Y}}_{t-w/2+l} - \tilde{\mathbf{X}}_l) + (\tilde{\mathbf{Y}}_{2t-w+1-l} - \tilde{\mathbf{X}}_{t-w/2+1-l})\}_{l=1}^{t/2-w/4}$ shares the same weakly dependence properties and similar moment/tail conditions as $\{\mathbf{X}_l\}_{l=1}^{n}$. Hence, our method can be applied to change point detection.

The selections of w and α are crucial in this method. We will elaborate on the specific choices for them in future works.

5. Simulation Study

5.1. Tuning Parameter Selection

Given the observations $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$, we use the minimum volatility (MV) method proposed in [27] to select the block size S.

When the data are independent, by the multiplier bootstrap method described in [28], we set $B = \tilde{n}$ (thus $S = 1$). In this case,

$$\hat{\Xi}_{\tilde{n}} = \mathrm{var}\left(\frac{1}{\sqrt{\tilde{n}}}\sum_{t=1}^{\tilde{n}}(\mathbf{Z}_t - \bar{\mathbf{Z}})\varrho_t' \mid \mathbf{Z}_1, \ldots, \mathbf{Z}_{\tilde{n}}\right)$$

$$= \frac{1}{\tilde{n}}\sum_{b=1}^{B}\left\{\left(\sum_{t\in\mathcal{I}_b}(\mathbf{Z}_t - \bar{\mathbf{Z}})\right)\left(\sum_{t\in\mathcal{I}_b}(\mathbf{Z}_t - \bar{\mathbf{Z}})\right)^{\mathsf{T}}\right\} = \frac{1}{\tilde{n}}\sum_{t=1}^{\tilde{n}}(\mathbf{Z_t} - \bar{\mathbf{Z}})(\mathbf{Z_t} - \bar{\mathbf{Z}})^{\mathsf{T}}$$

proves to be a reliable estimator of $\Xi_{\tilde{n}}$ introduced in Section 3. When the data are weakly dependent (and thus nearly independent), we expect a small value for S and a large value for B. Therefore, we recommend exploring a narrow range of S, such as $S \in \{1, \ldots, m\}$, where m is a moderate integer. In our theoretical proof, the quality of the bootstrap approximation depends on how well the $\hat{\Xi}_{\tilde{n}}$ approximates the covariance $\Xi_{\tilde{n}}$. The idea behind the MV method is that the conditional covariance $\hat{\Xi}_{\tilde{n}}$ should exhibit stable behavior

as a function of S within an appropriate range. For more comprehensive discussions on the MV method and its applications in time series analysis, we refer readers to [27,29]. For a moderately sized integer m, let $S_1 < S_2 < \cdots < S_m$ be a sequence of equally spaced candidate block sizes, and $S_0 = 2S_1 - S_2$, $S_{m+1} = 2S_m - S_{m-1}$. For each $i \in \{0, \ldots, m+1\}$, let

$$Y_j^i = \sum_{b=1}^{B(S_i)} \left\{ \sum_{t \in \mathcal{I}_b} (Z_{t,j} - \bar{Z}_j) \right\}^2,$$

where $j \in [p]$ and $B(S) = \lceil \tilde{n}/S \rceil$. Then for each $i \in \{1, \ldots, m\}$, we compute

$$Y^i = \sum_{j=1}^{p} \mathrm{sd}\left(\{Y_j^l\}_{l=i-1}^{i+1}\right),$$

where $\mathrm{sd}(\cdot)$ is the standard deviation. Then, we select the block size S_{i*} with $i^* = \arg\min_{i \in \{1, \ldots, m\}} Y^i$.

5.2. Simulation Settings

We present the results of a simulation study aimed at evaluating the performance of tests based on T_n, as defined in (2), in finite samples. To assess the finite-sample properties of the proposed test, we employed the following fundamental generating processes: $\mathbf{W} = \mathbf{HA} + \mathbf{f}(a) \in \mathbb{R}^{n \times p}$, where $\mathbf{A} \in \mathbb{R}^{p \times p}$ is the loading matrix, $\mathbf{f}(\cdot) : \mathbb{R} \to \mathbb{R}^{n \times p}$ is a constant function, the parameter a belongs to the set $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$, representing the distance between the null and alternative hypotheses. Additionally, $\mathbf{H} = (\mathbf{H}_1, \ldots, \mathbf{H}_n)^{\mathrm{T}} \in \mathbb{R}^{n \times p}$ with $\mathbf{H}_t = \rho \mathbf{H}_{t-1} + \boldsymbol{\varepsilon}_t \in \mathbb{R}^{p \times 1}$, where $\boldsymbol{\varepsilon}_t \overset{iid}{\sim} \mathcal{N}(\mathbf{0}_p, \mathrm{I}_p)$ and $\rho \in \{0, 0.1, 0.2\}$. Construct $\mathbf{f}_i(a) = (\mathbf{m}_1^{(i)}, \ldots, \mathbf{m}_n^{(i)})^{\mathrm{T}} \in \mathbb{R}^{n \times p}$ such that $\mathbf{m}_t^{(i)} = (m_{t,1}^{(i)}, \ldots, m_{t,p}^{(i)})^{\mathrm{T}}$ for $i \in \{1, 2\}$, where $m_{t,j}^{(1)} = a^j$ and $m_{t,j}^{(2)} = a(1 - j/p)$ for each $t \in [n]$ and $j \in [p]$. Then $\mathbf{f}_1(\cdot)$ and $\mathbf{f}_2(\cdot)$ represent the sparse and dense signal cases, respectively. We consider three different loading matrices for $\mathbf{A}$ as follows:

(M1).   Let $\mathbf{V} = (v_{k,l})_{1 \leq k,l \leq p}$ s.t. $v_{k,l} = 0.995^{|k-l|}$, then let $\mathbf{A} = \mathbf{V}^{1/2}$.

(M2).   Let $\mathbf{A} = (a_{k,l})_{1 \leq k,l \leq p}$ s.t. $a_{k,k} = 1$, $a_{k,l} = 0.7$ for $|k - l| = 1$ and $a_{k,l} = 0$ otherwise.

(M3).   Let $r = \lceil p/2.5 \rceil$, $\mathbf{V} = (v_{k,l})_{1 \leq k,l \leq p}$, where $v_{k,k} = 1$, $v_{k,l} = 0.9$ for $r(q-1) + 1 \leq k \neq l \leq rq$ with $q = 1, \ldots, \lfloor p/r \rfloor$, and $v_{k,l} = 0$ otherwise. Let $\mathbf{A} = \mathbf{V}^{1/2}$.

We assess the finite sample performance of our proposed test (denoted by Yang) in comparison with tests introduced by [5] (denoted by Dempster), [4] (denoted by BS), [6] (denoted by SD), and [8] (denoted by CLX). All tests in our simulations are conducted at the 5% significance level with 1000 Monte Carlo replications, and the number of bootstrap replications is set to 1000. We consider dimensions $p \in \{50, 200, 400, 800\}$ and sample size pairs $(n_1, n_2) \in \{(200, 220), (400, 420)\}$.

5.3. Simulation Results

For the testing of the null hypothesis, consider independent generations of $\{\mathbf{X}_t\}$ and $\{\mathbf{Y}_t\}$, following the same process as $\mathbf{W}$, with identical values for ρ and $\mathbf{f}(a) = \mathbf{0}$. The choice of $\mathbf{f}(a) = \mathbf{0}$ here is made for the sake of simplicity. We exclusively present the simulation results for (M1) in the main body of the paper. The results obtained for (M2) and (M3) are analogous to those of (M1) and are detailed in the Appendix E.

Table 1 presents the performance of various methods in controlling Type I errors based on (M1). As the dimension p or sample size (n_1, n_2) increases, the results of all methods exhibit small changes, except BS's. When ρ equals 0, indicating samples are generated from independent Gaussian distributions, both Yang's method and BS's method effectively control Type I errors at around 5%, while the control achieved by the other three methods is less optimal. It is noteworthy that, with an increase in ρ, the data generated by the AR(1) model significantly influence the other methods. In contrast, Yang's method demonstrates

superior and more stable results with increasing ρ. These comparative effects are also observable in the results based on (M2) and (M3) in the Appendix E. For this reason, we exclusively compare the empirical power results by different methods with $\rho = 0$.

Figures 1 and 2 depict the empirical power results of various methods for sparse and dense signals based on (M1). Similarly, as the dimension p increases, the results of all methods show little variation, except Dempster's. However, with an increase in sample size (n_1, n_2), most methods exhibit improvement in their results. In Figure 1, it is evident that Yang's method outperforms others significantly when the signal is sparse. Methods like SD, BS, and Dempster rely on the ℓ_2-norm of the data, aggregating signals across all dimensions for testing. This makes them less effective when the signal is sparse, i.e., anomalies appear in only a few dimensions. CLX's approach, akin to Yang's, tests whether the largest signal is abnormal. Consequently, CLX performs better than the other three methods in scenarios with sparse signals but still falls short of Yang's method. On the contrary, when the signal is dense, Figure 2 shows that all methods yield favorable results, with Dempster's method proving to be the most effective. Yang's method performs at a relatively high level among these methods. In contrast, the CLX's method, which performs well in sparse signals, exhibits a relatively lower level of performance in dense signals. In conclusion, the proposed method exhibits the most stable performance across all methods and performs exceptionally well on sparse data.

Table 1. The Type I error rates, expressed as percentages, were calculated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M1). The simulations were replicated 1000 times.

(n_1, n_2)	ρ	p	Yang	Dempster	BS	SD	CLX
(200,220)	0	50	5	18.5	5.8	0.9	0.3
		200	5.9	16.5	6.6	0.4	0.4
		400	5.4	17.4	6.2	0.2	0.3
		800	4.2	13.5	6.7	0.3	0.2
	0.1	50	6.5	22.8	9.3	2	1
		200	6.6	22.6	9.6	1.2	0.8
		400	7.4	22.9	10.4	1	0.8
		800	5.8	22.5	12.4	1	1.2
	0.2	50	6.8	30.2	13.8	3.1	2.5
		200	7.7	29.9	14.3	2.2	2.7
		400	9.3	30.5	18.2	2.2	2.4
		800	7.9	33.3	21.3	3	3.2
(400,420)	0	50	5.2	17.6	6.8	1	0.5
		200	5.3	17.2	6.8	0.5	0.1
		400	4.6	15.1	5.7	0.3	0
		800	5.2	14.2	6.3	0.3	0.4
	0.1	50	5.6	22.4	9.6	1.4	1
		200	6.3	22.5	9.6	1.3	0.8
		400	6.1	21.4	9.7	0.8	0.8
		800	6.5	23.6	12.1	0.7	1.2
	0.2	50	6.7	26.9	12.8	2.5	1.9
		200	7.6	29.2	14.9	2.3	2.4
		400	7.6	29.4	15.1	1.5	2.9
		800	8.3	36.3	21.9	2.5	3.8
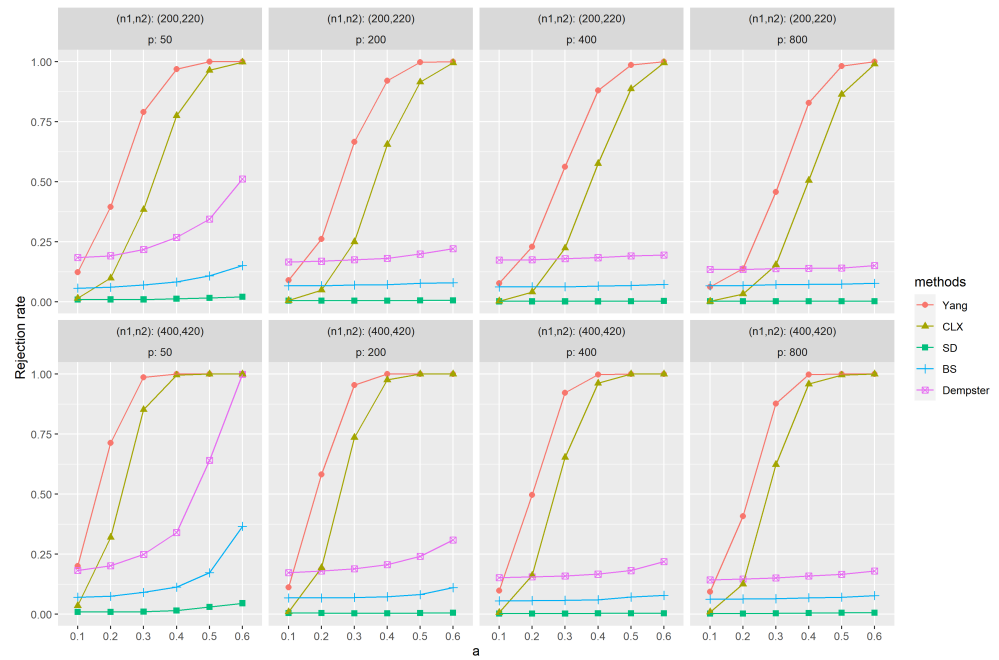
Figure 1. The empirical powers with sparse signals were evaluated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ based on (M1), $\mathbf{f}(\cdot) = \mathbf{0}$ and $\rho = 0$, and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M1), $\mathbf{f}(\cdot) = \mathbf{f}_1(\cdot)$ and $\rho = 0$. The parameter *a* represents the distance between the null and alternative hypotheses. The simulations were replicated 1000 times.
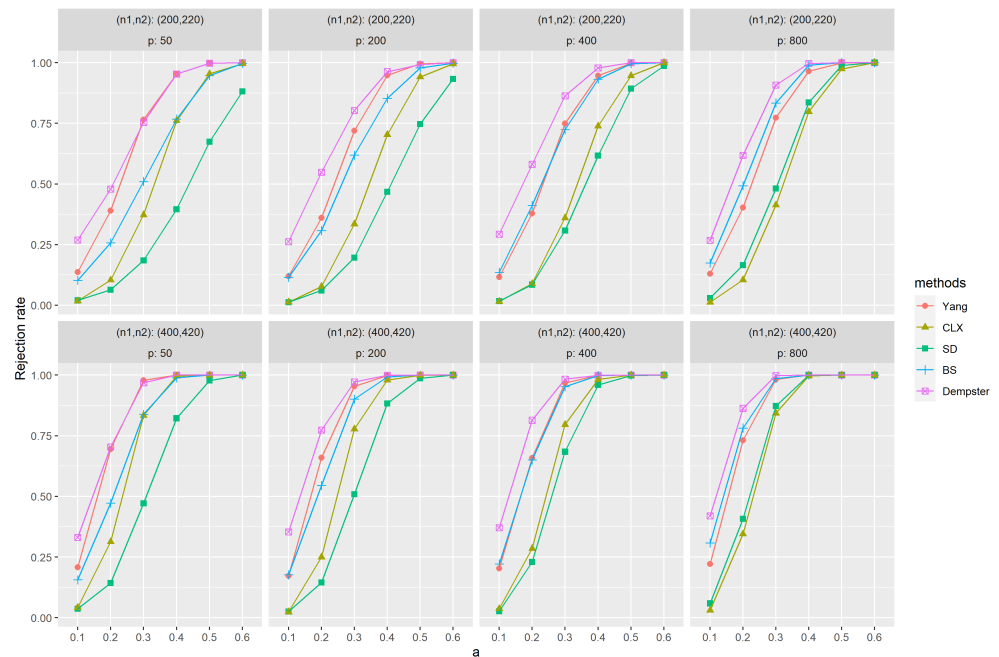
Figure 2. The empirical powers with dense signals were evaluated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ based on (M1), $\mathbf{f}(\cdot) = \mathbf{0}$ and $\rho = 0$, and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M1), $\mathbf{f}(\cdot) = \mathbf{f}_2(\cdot)$ and $\rho = 0$. The parameter *a* represents the distance between the null and alternative hypotheses. The simulations were replicated 1000 times.

6. Real Data Analysis

In this section, we apply the proposed method to a dataset comprised of stock data obtained from Bloomberg's public database. This dataset includes daily opening prices from 1 January 2018 to 31 December 2021 for 30 companies in the Consumer Discretionary

Sector (CDS) and 31 companies in the Information Technology Sector (ITS), all listed in the S&P 500. The sample sizes for the years 2018, 2019, 2020, and 2021 are 251, 250, 253, and 252, respectively. The findings are presented in Table 2. Regarding the data for the Consumer Discretionary (CD) and Information Technology (IT) sectors, all p-values from the tests between two consecutive years are 0. This suggests a significant variation in the average annual opening prices across different years for both CDs and ITs.

For data visualization, Figure 3 displays the average annual opening prices of 30 companies in the CDS (left subgragh) and 31 companies in the ITS (right subgragh) in 2018, 2019, 2020, and 2021. The two subgraghs both exhibit a pattern of annual growth in the opening prices of nearly every stock. These results are well in line with the conclusion of Table 2.

Table 2. The *p*-values for testing the equality of average annual opening prices across two consecutive years in the Consumer Discretionary Sector and Information Technology Sector, respectively.

Sector of S&P 500	2018–2019	2019–2020	2020–2021
Consumer Discretionary	0	0	0
Information Technology	0	0	0
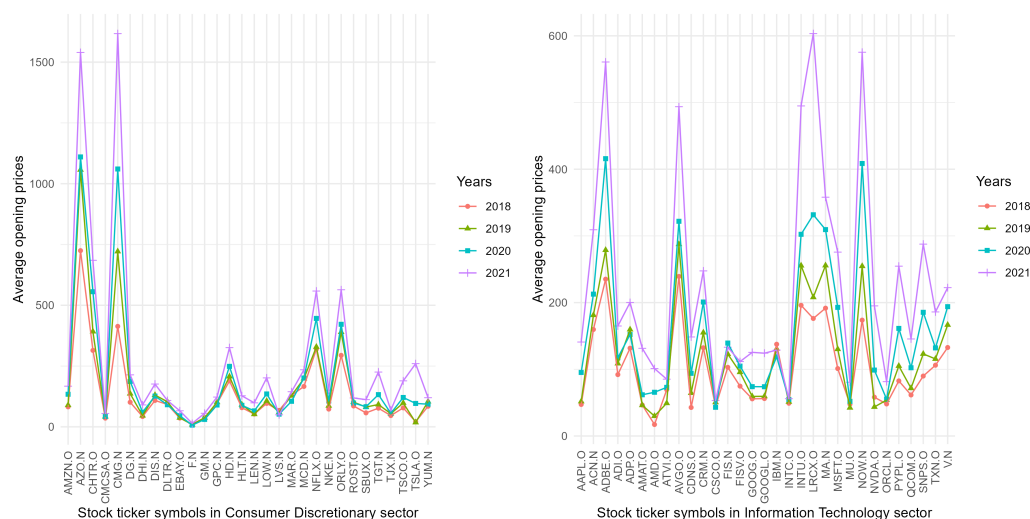
Figure 3. The average annual opening prices of 30 Consumer Discretionary corporations and 31 Information Technology corporations in 2018, 2019, 2020, and 2021.

7. Discussion

In this paper, we propose a two-sample test for high-dimensional time series based on blockwise bootstrap. Our ℓ_∞-type test statistic is designed to detect the largest abnormal signal among dimensions. Unlike some frameworks, we do not necessarily require independence within each observation or between the two sets of observations. Instead, we rely on the weak dependence property of the pair sequence $\{(\mathbf{X}_t, \mathbf{Y}_t)\}$ to ensure the asymptotic properties of our proposed method. We derive two Gaussian approximation results for two cases in which the dimension p diverges, one at a polynomial rate relative to the sample size n and the other at an exponential rate relative to the sample size n. In the bootstrap procedure, the block size serves as the tuning parameter, and we employ the minimum volatility method, as proposed by [27], for block size selection.

Our test statistic is designed to pinpoint the maximum value among dimensions, facilitating the detection of significant differences in certain dimensions. In cases where differences in each dimension are minimal, it is more appropriate to consider the ℓ_2-type

test statistic rather than the ℓ_∞-type one. Consequently, in the absence of prior information, the utilization of test statistics that combine both types proves advantageous. However, deriving theoretical results from such a combined approach is a significant challenge. As discussed in Section 4, our two-sample testing procedure can be applied to change point detection in high-dimensional time series. The choices of w, the size of each subsample mean, and the significance level α play crucial roles in this change point detection procedure. We leave these considerations for future research.

Data Availability Statement: The data used to support the findings of this study are included within the article.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Proof of Theorem 1

Appendix A.1. Proof of Theorem 1(i)

Proof. We first show that, for any $\tau > (q-1)m/(m-q)$ with some $q \in [2, \lfloor m \rfloor]$,

$$\max_{j \in [p]} \mathbb{E}\left(\left| \sum_{t=1}^{n} \xi_{t,j} \right|^q \right) \lesssim n^{q/2}. \tag{A1}$$

If $q = 2$, due to $\sum_{\kappa=1}^{\infty} \alpha^{\frac{m-2}{m}}(\kappa) \lesssim \sum_{\kappa=1}^{\infty} \kappa^{-\frac{(m-2)\tau}{m}} < \infty$, Equation (1.12b) (Davydov's inequality) of [30] yields

$$
\begin{aligned}
\mathbb{E}\left\{ \left(\sum_{t=1}^{n} \xi_{t,j} \right)^2 \right\} &= \sum_{t=1}^{n} \mathbb{E}(|\xi_{t,j}|^2) + \sum_{t_1 \neq t_2} \operatorname{cov}(\xi_{t_1,j}, \xi_{t_2,j}) \\
&\lesssim n + \sum_{t_1 \neq t_2} \{\mathbb{E}(|\xi_{t_1,j}|^m)\}^{\frac{1}{m}} \{\mathbb{E}(|\xi_{t_2,j}|^m)\}^{\frac{1}{m}} \alpha^{\frac{m-2}{m}}(|t_1 - t_2|) \\
&\lesssim n + n \sum_{\kappa=1}^{n} \alpha^{\frac{m-2}{m}}(\kappa) \lesssim n
\end{aligned}
\tag{A2}
$$

for any $j \in [p]$. For $q > 2$ and $j \in [p]$, Theorem 6.3 of [30] yields

$$\mathbb{E}\left(\left| \sum_{t=1}^{n} \xi_{t,j} \right|^q \right) \leq a_q s_{n,j}^q + n b_q \int_0^1 [\alpha^{-1}(u) \wedge n]^{q-1} \left\{ \sup_{t \in [n]} Q_{t,j}(u) \right\}^q du,$$

where $a_q, b_q > 0$ are two constants depending only on q, $s_{n,j}^2 = \sum_{t_1, t_2 = 1}^{n} |\operatorname{Cov}(\xi_{t_1,j}, \xi_{t_2,j})|$, $\alpha^{-1}(u) = \sum_{\kappa \geq 0} 1(u \leq \alpha(\kappa))$ and $Q_{t,j}(u) = \inf\{x : \mathbb{P}(|\xi_{t,j}| > x) \leq u\}$. By (A2), it holds that $s_{n,j}^q = (s_{n,j}^2)^{q/2} \lesssim n^{q/2}$. Due to $\max_{t \in [n]} \max_{j \in [p]} \mathbb{E}(|\xi_{t,j}|^m) \leq C$, we have $\max_{t \in [n]} \max_{j \in [p]} Q_{t,j}(u) \lesssim u^{-\frac{1}{m}}$. By the denifition of $\alpha^{-1}(\cdot)$, we know that $\alpha^{-1}(u) \lesssim u^{-\frac{1}{\tau}}$. Thus

$$\int_0^1 [\alpha^{-1}(u) \wedge n]^{q-1} \left\{ \sup_{t \in [n]} Q_{t,j}(u) \right\}^q du \lesssim \int_0^1 u^{-\frac{q-1}{\tau} - \frac{q}{m}} du \leq C,$$

where the last inequality follows from $\tau > (q-1)m/(m-q)$. Hence, we have

$$\mathbb{E}\left(\left| \sum_{t=1}^{n} \xi_{t,j} \right|^q \right) \lesssim n^{q/2}$$

for any $j \in [p]$. By combining above results, we complete the proof of (A1).

Now, we begin to prove Theorem 1(i). Define

$$\omega_n = \sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} S_{n,j} \le x \right) - \mathbb{P}\left(\max_{j\in[p]} W_j \le x \right) \right|.$$

Let $\check{\mathbf{S}}_n = (\check{S}_{n,1}, \ldots, \check{S}_{n,2p})^{\mathsf{T}} = (S_{n,1}, -S_{n,1}, \ldots, S_{n,p}, -S_{n,p})^{\mathsf{T}}$ and $\check{\mathbf{W}} = (\check{W}_1, \ldots, \check{W}_{2p})^{\mathsf{T}} = (W_1, -W_1, \ldots, W_p, -W_p)^{\mathsf{T}}$. Then, we have $\max_{j\in[p]} |S_{n,j}| = \max_{j\in[2p]} \check{S}_{n,j}$ and $\max_{j\in[p]} |W_j| = \max_{j\in[2p]} \check{W}_j$. Then, to obtain Theorem 1(i), without loss of generality, it suffices to specify the convergence rate of ω_n.

For some constant $\varsigma \in (0,1)$, let $B_n = \lfloor n^\varsigma \rfloor$ and $K_n = \lceil n/B_n \rceil$ be the number of blocks and the size of each block, respectively. For simplicity, we assume $B_n \asymp n^\varsigma$ and $K_n = n/B_n \asymp n^{1-\varsigma}$. We first decompose the sequence $\{1, \ldots, n\}$ into B_n blocks: $\mathcal{G}_b = \{(b-1)K_n + 1, \ldots, bK_n\}$ for $b \in [B_n]$. Let $g_n \gg k_n$ be two non-negative integers such that $K_n = g_n + k_n$. We then decompose each $\mathcal{G}_b$ ($b \in [B_n]$) to a "large" block $\mathcal{I}_b$ with length g_n and a "small" block $\mathcal{J}_b$ with length k_n: $\mathcal{I}_b = \{(b-1)K_n + 1, \ldots, bK_n - k_n\}$ and $\mathcal{J}_b = \{bK_n - k_n + 1, \ldots, bK_n\}$. Let $\mathbf{H}_b = (H_{b,1}, \ldots, H_{b,p})^{\mathsf{T}} = K_n^{-1/2} \sum_{t\in\mathcal{I}_b} \boldsymbol{\xi}_t$. For each $b \in [B_n]$ and some $D_n \to \infty$, define $\mathbf{H}_b^+ = (H_{b,1}^+, \ldots, H_{b,p}^+)^{\mathsf{T}}$ with $H_{b,j}^+ = H_{b,j}\mathbf{1}(|H_{b,j}| \le D_n) - \mathbb{E}\{H_{b,j}\mathbf{1}(|H_{b,j}| \le D_n)\}$ and $\mathbf{H}_b^- = (H_{b,1}^-, \ldots, H_{b,p}^-)^{\mathsf{T}}$ with $H_{b,j}^- = H_{b,j}\mathbf{1}(|H_{b,j}| > D_n) - \mathbb{E}\{H_{b,j}\mathbf{1}(|H_{b,j}| > D_n)\}$. For each $j \in [p]$, by Theorem 2 of [17], there exists an independent sequence $\{\tilde{H}_{b,j}\}_{b=1}^{B_n}$ such that $\tilde{H}_{b,j}$ has the same distribution as $H_{b,j}^+$ and

$$\mathbb{E}(|\tilde{H}_{b,j} - H_{b,j}^+|) \lesssim \int_0^{\alpha(k_n)} \inf\{x \in \mathbb{R} : \mathbb{P}(|H_{b,j}^+| > x) \le u\} \mathrm{d}u.$$

Due to $|H_{b,j}^+| \le 2D_n$, we have $\inf\{x \in \mathbb{R} : \mathbb{P}(|H_{b,j}^+| > x) \le u\} \lesssim D_n$ for any $u \ge 0$, which implies

$$\mathbb{E}(|\tilde{H}_{b,j} - H_{b,j}^+|) \lesssim D_n \alpha(k_n). \tag{A3}$$

Define $\tilde{\mathbf{S}}_n = (\tilde{S}_{n,1}, \ldots, \tilde{S}_{n,p})^{\mathsf{T}} = B_n^{-1/2} \sum_{b=1}^{B_n} \tilde{\mathbf{H}}_b$ with $\tilde{S}_{n,j} = B_n^{-1/2} \sum_{b=1}^{B_n} \tilde{H}_{b,j}$ and

$$\tilde{\omega}_n = \sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} \tilde{S}_{n,j} \le x \right) - \mathbb{P}\left(\max_{j\in[p]} W_j \le x \right) \right|. \tag{A4}$$

For any $\epsilon_1 > 0$, triangle inequality implies

$$\mathbb{P}\left(\max_{j\in[p]} S_{n,j} \le x \right) \le \mathbb{P}\left(\max_{j\in[p]} \tilde{S}_{n,j} \le x + \epsilon_1 \right) + \mathbb{P}\left(\left| \max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j} \right| > \epsilon_1 \right)$$

$$\le \mathbb{P}\left(\max_{j\in[p]} W_j \le x + \epsilon_1 \right) + \tilde{\omega}_n + \mathbb{P}\left(\left| \max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j} \right| > \epsilon_1 \right)$$

$$\le \mathbb{P}\left(\max_{j\in[p]} W_j \le x \right) + \mathbb{P}\left(x - \epsilon_1 < \max_{j\in[p]} W_j \le x + \epsilon_1 \right) + \tilde{\omega}_n + \mathbb{P}\left(\left| \max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j} \right| > \epsilon_1 \right)$$

for any $x > 0$, then $\mathbb{P}(\max_{j\in[p]} S_{n,j} \le x) - \mathbb{P}(\max_{j\in[p]} W_j \le x) \le \mathbb{P}(x - \epsilon_1 < \max_{j\in[p]} W_j \le x + \epsilon_1) + \tilde{\omega}_n + \mathbb{P}(|\max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j}| > \epsilon_1)$. Likewise, $\mathbb{P}(\max_{j\in[p]} S_{n,j} \le x) - \mathbb{P}(\max_{j\in[p]} W_j \le x) \ge -\mathbb{P}(x - \epsilon_1 < \max_{j\in[p]} W_j \le x + \epsilon_1) - \tilde{\omega}_n - \mathbb{P}(|\max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j}| > \epsilon_1)$. Due to $\min_{j\in[p]} (\Sigma_n)_{j,j} \ge \lambda_{\min}(\Sigma_n) \ge c$, Lemma A.1 of [31] yields

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left(x - \epsilon_1 < \max_{j\in[p]} W_j \le x + \epsilon_1 \right) \lesssim \epsilon_1 (\log p)^{1/2}$$

for any $\epsilon_1 > 0$. Thus, we can conclude that

$$\omega_n \lesssim \tilde{\omega}_n + \mathbb{P}\left(\left|\max_{j \in [p]} S_{n,j} - \max_{j \in [p]} \tilde{S}_{n,j}\right| > \epsilon_1\right) + \epsilon_1 (\log p)^{1/2}. \tag{A5}$$

Define $\mathbf{S}_n^+ = (S_{n,1}^+, \ldots, S_{n,p}^+)^\mathsf{T} = B_n^{-1/2} \sum_{b=1}^{B_n} \mathbf{H}_b^+$. By triangle inequality,

$$\left|\max_{j \in [p]} S_{n,j} - \max_{j \in [p]} S_{n,j}^+\right| \leq \max_{j \in [p]} |S_{n,j} - S_{n,j}^+| \leq \max_{j \in [p]} \left|\frac{1}{n^{1/2}} \sum_{b=1}^{B_n} \sum_{t \in \mathcal{J}_b} \xi_{t,j}\right| + \max_{j \in [p]} \left|\frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} H_{b,j}^-\right|.$$

By (A1), we have $\mathbb{E}(|H_{b,j}|^3) \leq C$. Thus $\mathbb{E}(|H_{b,j}^-|^3) \lesssim \mathbb{E}(|H_{b,j}|^3) \leq C$, and

$$\mathbb{E}(|H_{b,j}^-|^2) \lesssim \mathbb{E}\{|H_{b,j}|^2 \mathbb{1}(|H_{b,j}| > D_n)\} \leq \mathbb{E}(|H_{b,j}|^3) D_n^{-1} \lesssim D_n^{-1}. \tag{A6}$$

Similar to (A2), we have $\mathbb{E}(|\sum_{b=1}^{B_n} \sum_{t \in \mathcal{J}_b} \xi_{t,j}|) \lesssim B_n^{1/2} k_n^{1/2}$ for any $j \in [p]$, and

$$\mathbb{E}\left\{\left(\sum_{b=1}^{B_n} H_{b,j}^-\right)^2\right\} = \sum_{b=1}^{B_n} \mathbb{E}(|H_{b,j}^-|^2) + \sum_{b_1 \neq b_2} \mathrm{cov}(H_{b_1,j}^-, H_{b_2,j}^-)$$

$$\lesssim B_n D_n^{-1} + \sum_{b_1 \neq b_2} \alpha^{\frac{1}{3}}\{k_n \mathbb{1}(|b_1 - b_2| = 1) + |b_2 - b_1 - 1| K_n \mathbb{1}(|b_1 - b_2| > 1)\}$$

$$\lesssim B_n D_n^{-1} + \sum_{|b_1 - b_2| = 1} \alpha^{\frac{1}{3}}(k_n) + \sum_{|b_1 - b_2| > 1} \alpha^{\frac{1}{3}}(|b_1 - b_2 - 1| K_n) \lesssim B_n D_n^{-1} + B_n k_n^{-\frac{\tau}{3}}, \tag{A7}$$

where the last inequality follows from $\tau > 3$. Thus, $\mathbb{E}(|\sum_{b=1}^{B_n} H_{b,j}^-|) \lesssim B_n^{1/2}(D_n^{-1/2} + k_n^{-\tau/6})$ and

$$\mathbb{E}\left(\left|\max_{j \in [p]} S_{n,j} - \max_{j \in [p]} S_{n,j}^+\right|\right) \lesssim \frac{pk_n^{1/2}}{K_n^{1/2}} + \frac{p}{D_n^{1/2}} + \frac{p}{k_n^{\tau/6}}. \tag{A8}$$

Due to $\tilde{H}_{b,j}$ having the same distribution as $H_{b,j}^+$ and $|H_{b,j}^+| \leq 2D_n$, by (A3), we have $\mathbb{E}(|\tilde{H}_{b,j} - H_{b,j}^+|^s) \lesssim D_n^s k_n^{-\tau}$ for $s \in \{2, 3\}$. Thus, following the same arguments as in the proof of (A7), it holds that

$$\mathbb{E}\left[\left\{\sum_{b=1}^{B_n} (\tilde{H}_{b,j} - H_{b,j}^+)\right\}^2\right]$$

$$\lesssim \sum_{b=1}^{B_n} D_n^2 k_n^{-\tau} + D_n^2 k_n^{-2\tau/3}\left\{\sum_{|b_1 - b_2| = 1} \alpha^{\frac{1}{3}}(k_n) + \sum_{|b_1 - b_2| > 1} \alpha^{\frac{1}{3}}(|b_1 - b_2 - 1| K_n)\right\}$$

$$\lesssim B_n D_n^2 k_n^{-\tau}.$$

Thus, $\mathbb{E}(|\sum_{b=1}^{B_n} (\tilde{H}_{b,j} - H_{b,j}^+)|) \lesssim B_n^{1/2} D_n k_n^{-\tau/2}$ and

$$\mathbb{E}\left(\left|\max_{j \in [p]} \tilde{S}_{n,j} - \max_{j \in [p]} S_{n,j}^+\right|\right) \leq \mathbb{E}\left(\max_{j \in [p]} |\tilde{S}_{n,j} - S_{n,j}^+|\right) \lesssim \frac{pD_n}{k_n^{\tau/2}}.$$

Together with (A8), we have

$$\mathbb{E}\left(\left|\max_{j \in [p]} S_{n,j} - \max_{j \in [p]} \tilde{S}_{n,j}\right|\right) \lesssim \frac{pk_n^{1/2}}{K_n^{1/2}} + \frac{p}{D_n^{1/2}} + \frac{p}{k_n^{\tau/6}} + \frac{pD_n}{k_n^{\tau/2}}.$$

Let $\epsilon_1 = p^{1/2}(\log p)^{-1/4}(k_n^{1/2}K_n^{-1/2} + D_n^{-1/2} + k_n^{-\tau/6} + D_n k_n^{-\tau/2})^{1/2}$. It holds by (A5) and Markov inequality that

$$\omega_n \lesssim \tilde{\omega}_n + p^{1/2}(\log p)^{1/4}\left(\frac{k_n^{1/2}}{K_n^{1/2}} + \frac{1}{D_n^{1/2}} + \frac{1}{k_n^{\tau/6}} + \frac{D_n}{k_n^{\tau/2}}\right)^{1/2}. \tag{A9}$$

Define $\tilde{\mathbf{\Sigma}}_{\mathrm{G}} = B_n^{-1}\sum_{b=1}^{B_n}\mathrm{var}(\tilde{\mathbf{H}}_b)$ and $\mathbf{\Delta} = |\mathbf{\Sigma}_n - \tilde{\mathbf{\Sigma}}_{\mathrm{G}}|$, where $\mathbf{\Sigma}_n = \mathbb{E}(\mathbf{S}_n\mathbf{S}_n^{\mathsf{T}})$. Note that

$$\mathbf{\Delta} = \left|\frac{1}{B_n}\sum_{b=1}^{B_n}\{\mathrm{var}(\tilde{\mathbf{H}}_b) - \mathrm{var}(\mathbf{H}_b^+)\} + \frac{1}{B_n}\sum_{b=1}^{B_n}\{\mathrm{var}(\mathbf{H}_b^+) - \mathrm{var}(\mathbf{H}_b)\} + \frac{1}{B_n}\sum_{b=1}^{B_n}\mathrm{var}(\mathbf{H}_b) - \mathbf{\Sigma}_n\right|$$

$$\leq \underbrace{\frac{1}{B_n}\sum_{b=1}^{B_n}|\mathrm{var}(\tilde{\mathbf{H}}_b) - \mathrm{var}(\mathbf{H}_b^+)|}_{\mathbf{\Delta}_1} + \underbrace{\frac{1}{B_n}\sum_{b=1}^{B_n}|\mathrm{var}(\mathbf{H}_b^+) - \mathrm{var}(\mathbf{H}_b)|}_{\mathbf{\Delta}_2} + \underbrace{\left|\frac{1}{B_n}\sum_{b=1}^{B_n}\mathrm{var}(\mathbf{H}_b) - \mathbf{\Sigma}_n\right|}_{\mathbf{\Delta}_3}. \tag{A10}$$

In this sequel, we specify the convergence rates of $|\mathbf{\Delta}_1|_\infty$, $|\mathbf{\Delta}_2|_\infty$, and $|\mathbf{\Delta}_3|_\infty$, respectively. Note that the (i,j)-th element of $\mathrm{var}(\tilde{\mathbf{H}}_b) - \mathrm{var}(\mathbf{H}_b^+)$ is $\mathbb{E}(\tilde{H}_{b,i}\tilde{H}_{b,j} - H_{b,i}^+ H_{b,j}^+)$. Due to $\tilde{H}_{b,j}$ having the same distribution as $H_{b,j}^+$ and $|H_{b,j}^+| \lesssim D_n$ for any $b \in [B_n]$ and $j \in [p]$, it holds by (A3) that

$$|\mathbb{E}(\tilde{H}_{b,i}\tilde{H}_{b,j} - H_{b,i}^+ H_{b,j}^+)| \leq |\mathbb{E}\{(\tilde{H}_{b,i} - H_{b,i}^+)\tilde{H}_{b,j}\}| + |\mathbb{E}\{(\tilde{H}_{b,j} - H_{b,j}^+)\tilde{H}_{b,i}^+\}| \lesssim D_n^2 k_n^{-\tau}$$

for any $b \in [B_n]$ and $i,j \in [p]$. Thus, we can conclude that $|\mathbf{\Delta}_1|_\infty \lesssim D_n^2 k_n^{-\tau}$. The (i,j)-th element of $\mathrm{var}(\mathbf{H}_b^+) - \mathrm{var}(\mathbf{H}_b)$ is $\mathbb{E}(H_{b,i}^+ H_{b,j}^+ - H_{b,i}H_{b,j})$. Note that $\mathbb{E}(|H_{b,j}^-|) \lesssim \mathbb{E}\{|H_{b,j}|\mathbf{1}(|H_{b,j}| > D_n)\} \leq \mathbb{E}(|H_{b,j}|^3)D_n^{-2} \lesssim D_n^{-2}$. Due to $H_{b,j} = H_{b,j}^+ + H_{b,j}^-$, it holds by (A6) that

$$|\mathbb{E}(H_{b,i}^+ H_{b,j}^+ - H_{b,i}H_{b,j})| = |\mathbb{E}\{H_{b,i}^+ H_{b,j}^+ - (H_{b,i}^+ + H_{b,i}^-)(H_{b,j}^+ + H_{b,j}^-)\}|$$

$$\leq |\mathbb{E}(H_{b,i}^+ H_{b,j}^-)| + |\mathbb{E}(H_{b,j}^+ H_{b,i}^-)| + |\mathbb{E}(H_{b,i}^- H_{b,j}^-)| \lesssim D_n^{-1}$$

for any $b \in [B_n]$ and $i,j \in [p]$. Thus, we can conclude that $|\mathbf{\Delta}_2|_\infty \lesssim D_n^{-1}$. The (i,j)-th element of $\mathbf{\Sigma}_n - B_n^{-1}\sum_{b=1}^{B_n}\mathrm{var}(\mathbf{H}_b)$ is $n^{-1}\sum_{t_1,t_2=1}^n\mathbb{E}(\xi_{t_1,i}\xi_{t_2,j}) - n^{-1}\sum_{b=1}^{B_n}\sum_{t_1,t_2\in\mathcal{I}_b}\mathbb{E}(\xi_{t_1,i}\xi_{t_2,j})$, and

$$\left|\frac{1}{n}\sum_{t_1,t_2=1}^n\mathbb{E}(\xi_{t_1,i}\xi_{t_2,j}) - \frac{1}{n}\sum_{b=1}^{B_n}\sum_{t_1,t_2\in\mathcal{I}_b}\mathbb{E}(\xi_{t_1,i}\xi_{t_2,j})\right|$$

$$= \frac{1}{n}\left|\sum_{b_1\neq b_2}\mathbb{E}\left\{\left(\sum_{t\in\mathcal{G}_{b_1}}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{G}_{b_2}}\xi_{t,j}\right)\right\}\right.$$

$$\left.+ \sum_{b=1}^{B_n}\mathbb{E}\left\{\left(\sum_{t\in\mathcal{I}_b}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{J}_b}\xi_{t,j}\right) + \left(\sum_{t\in\mathcal{J}_b}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{G}_b}\xi_{t,j}\right)\right\}\right|. \tag{A11}$$

Similar to the proof of (A2), we have

$$\left|\mathbb{E}\left\{\left(\sum_{t\in\mathcal{J}_b}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{G}_b}\xi_{t,j}\right)\right\}\right|$$

$$= \left|\sum_{t\in\mathcal{J}_b}\mathrm{cov}(\xi_{t,i},\xi_{t,j}) + \sum_{t_1\neq t_2:t_1,t_2\in\mathcal{J}_b}\mathrm{cov}(\xi_{t_1,i},\xi_{t_2,j}) + \sum_{t_1\in\mathcal{J}_b}\sum_{t_2\in\mathcal{I}_b}\mathrm{cov}(\xi_{t_1,i},\xi_{t_2,j})\right|$$

$$\lesssim k_n + \sum_{t_1\neq t_2:t_1,t_2\in\mathcal{J}_b}\{\mathbb{E}(|\xi_{t_1,i}|^3)\}^{\frac{1}{3}}\{\mathbb{E}(|\xi_{t_2,j}|^3)\}^{\frac{1}{3}}\alpha^{\frac{1}{3}}(|t_1 - t_2|)$$

$$+ \sum_{t_1\in\mathcal{J}_b}\sum_{t_2\in\mathcal{I}_b}\{\mathbb{E}(|\xi_{t_1,i}|^3)\}^{\frac{1}{3}}\{\mathbb{E}(|\xi_{t_2,j}|^3)\}^{\frac{1}{3}}\alpha^{\frac{1}{3}}(|t_1 - t_2|) \lesssim k_n.$$

Similarly, we can also obtain

$$\left| \mathbb{E}\left\{ \left(\sum_{t\in\mathcal{I}_b} \xi_{t,i} \right) \left(\sum_{t\in\mathcal{J}_b} \xi_{t,j} \right) \right\} \right| \lesssim k_n.$$

Thus,

$$\left| \sum_{b=1}^{B_n} \mathbb{E}\left\{ \left(\sum_{t\in\mathcal{I}_b} \xi_{t,i} \right) \left(\sum_{t\in\mathcal{J}_b} \xi_{t,j} \right) + \left(\sum_{t\in\mathcal{J}_b} \xi_{t,i} \right) \left(\sum_{t\in\mathcal{G}_b} \xi_{t,j} \right) \right\} \right| \lesssim k_n B_n.$$

Analogously to the proof of (A2), if $b_1 < b_2$, due to $\tau > 2m/(m-3)$,

$$\left| \mathbb{E}\left\{ \left(\sum_{t\in\mathcal{G}_{b_1}} \xi_{t,i} \right) \left(\sum_{t\in\mathcal{G}_{b_2}} \xi_{t,j} \right) \right\} \right| \lesssim \sum_{t_1\in\mathcal{G}_{b_1}} \sum_{t_2\in\mathcal{G}_{b_2}} \{\mathbb{E}(|\xi_{t_1,i}|^m)\}^{\frac{1}{m}} \{\mathbb{E}(|\xi_{t_2,i}|^m)\}^{\frac{1}{m}} \alpha^{\frac{m-2}{m}}(|t_1-t_2|)$$

$$\lesssim \sum_{\delta=1}^{K_n} \delta\alpha^{\frac{m-2}{m}} \{(b_2-b_1-1)K_n + \delta\}$$

$$\lesssim 1(b_2-b_1=1) + K_n^2 \alpha^{\frac{m-2}{m}} \{(b_2-b_1-1)K_n\} 1(b_2-b_1 > 1).$$

Then,

$$\sum_{b_1 < b_2} \left| \mathbb{E}\left\{ \left(\sum_{t\in\mathcal{G}_{b_1}} \xi_{t,i} \right) \left(\sum_{t\in\mathcal{G}_{b_2}} \xi_{t,j} \right) \right\} \right| \lesssim B_n + B_n K_n^{\frac{2m-(m-2)\tau}{m}} \sum_{\delta=1}^{B_n} \delta^{\frac{-(m-2)\tau}{m}} \lesssim B_n.$$

The same result still holds for $b_1 > b_2$. Thus, we can conclude that

$$\sum_{b_1 \neq b_2} \left| \mathbb{E}\left\{ \left(\sum_{t\in\mathcal{G}_{b_1}} \xi_{t,i} \right) \left(\sum_{t\in\mathcal{G}_{b_2}} \xi_{t,j} \right) \right\} \right| \lesssim B_n.$$

Then, by (A11), it holds that

$$\left| \frac{1}{n} \sum_{t_1,t_2=1}^{n} \mathbb{E}(\xi_{t_1,i}\xi_{t_2,j}) - \frac{1}{n} \sum_{b=1}^{B_n} \sum_{t_1,t_2\in\mathcal{I}_b} \mathbb{E}(\xi_{t_1,i}\xi_{t_2,j}) \right| \lesssim \frac{k_n}{K_n}$$

for any $i,j \in [p]$. Thus, $|\mathbf{\Delta}_3|_\infty \lesssim k_n K_n^{-1}$. By (A10), we can conclude that

$$|\mathbf{\Delta}|_\infty \leq |\mathbf{\Delta}_1|_\infty + |\mathbf{\Delta}_2|_\infty + |\mathbf{\Delta}_3|_\infty \lesssim \frac{D_n^2}{k_n^\tau} + \frac{1}{D_n} + \frac{k_n}{K_n}.$$

Let $\{\tilde{\mathbf{H}}_b^{\mathrm{G}}\}_{b=1}^{B_n}$ be a sequence of an independent Gaussian vector such that $\tilde{\mathbf{H}}_b^{\mathrm{G}} = (\tilde{H}_{b,1}^{\mathrm{G}}, \ldots, \tilde{H}_{b,p}^{\mathrm{G}})^{\mathrm{T}}$ $\sim \mathcal{N}\{\mathbf{0}_p, \mathrm{var}(\tilde{\mathbf{H}}_b)\}$ for each $b \in [B_n]$, where $\tilde{\mathbf{H}}_b = (\tilde{H}_{b,1}, \ldots, \tilde{H}_{b,p})^{\mathrm{T}}$. By Theorem 1.1 of [15], Cauchy–Schwarz inequality and Jensen's inequality,

$$\sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j} \leq x \right) - \mathbb{P}\left(\max_{j\in[p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j}^{\mathrm{G}} \leq x \right) \right|$$

$$\lesssim p^{1/4} \cdot \left\{ \sum_{b=1}^{B_n} \mathbb{E}(|\tilde{\mathbf{\Sigma}}_{\mathrm{G}}^{-1/2} B_n^{-1/2} \tilde{\mathbf{H}}_b|_2^3) \right\}$$

$$\lesssim \frac{p^{1/4}}{B_n^{3/2}} \|\tilde{\mathbf{\Sigma}}_{\mathrm{G}}^{-1/2}\|_2^3 \cdot \left[\sum_{b=1}^{B_n} \mathbb{E}\left\{ \left(\sum_{j=1}^{p} \tilde{H}_{b,j}^2 \right)^{3/2} \right\} \right]$$

$$\lesssim \frac{p^{7/4}}{B_n^{3/2}} \|\tilde{\mathbf{\Sigma}}_{\mathrm{G}}^{-1/2}\|_2^3 \cdot \left\{ \sum_{b=1}^{B_n} \max_{j\in[p]} \mathbb{E}(|\tilde{H}_{b,j}|^3) \right\},$$

where $\tilde{\boldsymbol{\Sigma}}_{\mathrm{G}} = B_n^{-1} \sum_{b=1}^{B_n} \mathrm{var}(\tilde{\mathbf{H}}_b)$. Note that

$$|\lambda_{\min}(\tilde{\boldsymbol{\Sigma}}_{\mathrm{G}}) - \lambda_{\min}(\boldsymbol{\Sigma}_n)| \leq \|\boldsymbol{\Delta}\|_2 \leq p|\boldsymbol{\Delta}|_\infty.$$

Due to $\lambda_{\min}(\boldsymbol{\Sigma}_n) \geq c$, we have $\lambda_{\min}(\tilde{\boldsymbol{\Sigma}}_{\mathrm{G}}) \geq c$ as long as $p|\boldsymbol{\Delta}|_\infty = o(1)$. Thus, if $p|\boldsymbol{\Delta}|_\infty = o(1)$, we have $\|\tilde{\boldsymbol{\Sigma}}_{\mathrm{G}}^{-1/2}\|_2 \leq C$. Since $H_{b,j} = K_n^{-1/2} \sum_{t \in \mathcal{I}_b} \xi_{t,j}$, (A1) yields $\mathbb{E}(|\tilde{H}_{b,j}|^3) = \mathbb{E}(|H_{b,j}^+|^3) \lesssim \mathbb{E}(|H_{b,j}|^3) \leq C$ for any $b \in [B_n]$ and $j \in [p]$, which implies

$$\sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j} \leq x \right) - \mathbb{P}\left(\max_{j\in[p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j}^{\mathrm{G}} \leq x \right) \right| \lesssim \frac{p^{7/4}}{B_n^{1/2}} \qquad \text{(A12)}$$

provided that $p|\boldsymbol{\Delta}|_\infty = o(1)$. By Proposition 2.1 of [16], we have

$$\sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j}^{\mathrm{G}} \leq x \right) - \mathbb{P}\left(\max_{j\in[p]} W_j \leq x \right) \right| \lesssim |\boldsymbol{\Delta}|_\infty^{1/2} \log p. \qquad \text{(A13)}$$

Then, by (A4), (A12), and (A13), we have

$$\tilde{\omega}_n \lesssim \frac{p^{7/4}}{B_n^{1/2}} + |\boldsymbol{\Delta}|_\infty^{1/2} \log p$$

provided that $p|\boldsymbol{\Delta}|_\infty = o(1)$. Together with (A9),

$$\omega_n \lesssim \frac{p^{7/4}}{B_n^{1/2}} + |\boldsymbol{\Delta}|_\infty^{1/2} \log p + p^{1/2} (\log p)^{1/4} \left(\frac{k_n^{1/2}}{K_n^{1/2}} + \frac{1}{D_n^{1/2}} + \frac{1}{k_n^{\tau/6}} + \frac{D_n}{k_n^{\tau/2}} \right)^{1/2}$$

provided that $p|\boldsymbol{\Delta}|_\infty = o(1)$. Select $D_n \asymp n^{4\tau/(11\tau+12)}$, $k_n \asymp n^{12/(11\tau+12)}$, and $\varsigma = 7\tau/(11\tau+12)$. Then, if $p = o\{n^{2\tau/(11\tau+12)}\}$, we have

$$\omega_n \lesssim \frac{p^{7/4}}{n^{7\tau/(22\tau+24)}} + \frac{\log p}{n^{2\tau/(11\tau+12)}} + \frac{p^{1/2} (\log p)^{1/4}}{n^{\tau/(11\tau+12)}} \lesssim \frac{p^{1/2} (\log p)^{1/4}}{n^{\tau/(11\tau+12)}}.$$

Hence, we complete the proof of Theorem 1(i). $\square$

Appendix A.2. Proof of Theorem 1(ii)

Proof. Define $\{(\mathcal{G}_b, \mathcal{I}_b, \mathcal{J}_b)\}_{b=1}^{B_n}$, $\{\mathbf{H}_b^+\}_{b=1}^{B_n}$, and $\{\mathbf{H}_b^-\}_{b=1}^{B_n}$ in the same manner as in the proof of Theorem 1(i) with $B_n \asymp n^\varsigma$, $K_n \asymp n^{1-\varsigma}$, $k_n \ll n^{1-\varsigma}$ and $D_n \to \infty$, where $\varsigma \in (0,1)$. Let

$$\omega_n = \sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} S_{n,j} \leq x \right) - \mathbb{P}\left(\max_{j\in[p]} W_j \leq x \right) \right|.$$

Analogously to (A5), due to $\min_{j\in[p]} (\boldsymbol{\Sigma}_n)_{j,j} > c$, we have

$$\omega_n \lesssim \tilde{\omega}_n + \mathbb{P}\left(\left| \max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j} \right| > \epsilon_2 \right) + \epsilon_2 (\log p)^{1/2} \qquad \text{(A14)}$$

for some $\epsilon_2 > 0$, where $\tilde{S}_{n,j} = B_n^{-1/2} \sum_{b=1}^{B_n} \tilde{H}_{b,j}$ with $\{\tilde{H}_{b,j}\}$ specified in the same manner as in the proof of Theorem 1(i), and

$$\tilde{\omega}_n = \sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} \tilde{S}_{n,j} \leq x \right) - \mathbb{P}\left(\max_{j\in[p]} W_j \leq x \right) \right|.$$

Define $\mathbf{S}_n^+ = (S_{n,1}^+, \ldots, S_{n,p}^+)^\mathsf{T} = B_n^{-1/2} \sum_{b=1}^{B_n} \mathbf{H}_b^+$. By triangle inequality,

$$\left| \max_{j\in[p]} S_{n,j} - \max_{j\in[p]} S_{n,j}^+ \right| \le \max_{j\in[p]} |S_{n,j} - S_{n,j}^+| \le \max_{j\in[p]} \left| \frac{1}{n^{1/2}} \sum_{b=1}^{B_n} \sum_{t\in\mathcal{J}_b} \xi_{t,j} \right| + \max_{j\in[p]} \left| \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} H_{b,j}^- \right|.$$

Note that $\mathbb{P}(|\xi_{t,j} M_n^{-1}| > x) \lesssim \exp(-Cx^{\gamma_1})$ for any $x > 0$. Let $\tilde{\gamma} = (1/\gamma_1 + 1/\gamma_2)^{-1}$. By Theorem 1 of [32] and Bonferroni inequality, we have

$$\mathbb{P}\left(\max_{j\in[p]} \left| \frac{1}{n^{1/2}} \sum_{b=1}^{B_n} \sum_{t\in\mathcal{J}_b} \xi_{t,j} \right| > x \right) \lesssim p B_n k_n \exp\left(-\frac{Cn^{\tilde{\gamma}/2} x^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right) + p \exp\left(-\frac{Cnx^2}{M_n^2 B_n k_n} \right) \quad \text{(A15)}$$

for any $x \gg M_n n^{-1/2}$. Similarly, by Theorem 1 of [32] again, for any $x \gg M_n K_n^{-1/2}$,

$$\mathbb{P}(|H_{b,j}| > x) = \mathbb{P}\left(\left| \frac{1}{K_n^{1/2}} \sum_{t\in\mathcal{I}_b} \xi_{t,j} \right| > x \right) \lesssim K_n \exp\left(-\frac{CK_n^{\tilde{\gamma}/2} x^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right) + \exp\left(-\frac{Cx^2}{M_n^2} \right).$$

Then, if $D_n > M_n$,

$$\mathbb{E}\{H_{b,j}^2 \mathbf{1}(|H_{b,j}| > D_n)\} = 2\int_0^{D_n} x\mathbb{P}(|H_{b,j}| > D_n)\mathrm{d}x + 2\int_{D_n}^\infty x\mathbb{P}(|H_{b,j}| > x)\mathrm{d}x$$

$$\lesssim D_n^2 \left\{ K_n \exp\left(-\frac{CK_n^{\tilde{\gamma}/2} D_n^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right) + \exp\left(-\frac{CD_n^2}{M_n^2} \right) \right\}$$

$$+ K_n \int_{D_n}^\infty x\exp\left(-\frac{CK_n^{\tilde{\gamma}/2} x^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right)\mathrm{d}x + \int_{D_n}^\infty x\exp\left(-\frac{Cx^2}{M_n^2} \right)\mathrm{d}x$$

$$\lesssim D_n^2 \left\{ K_n \left(-\frac{CK_n^{\tilde{\gamma}/2} D_n^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right) + \exp\left(-\frac{CD_n^2}{M_n^2} \right) \right\}.$$

Thus, for any $b \in [B_n]$ and $j \in [p]$,

$$\mathbb{E}(|H_{b,j}^-|^2) \lesssim \mathbb{E}\{H_{b,j}^2 \mathbf{1}(|H_{b,j}| > D_n)\} \lesssim D_n^2 \left\{ K_n \left(-\frac{CK_n^{\tilde{\gamma}/2} D_n^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right) + \exp\left(-\frac{CD_n^2}{M_n^2} \right) \right\}. \quad \text{(A16)}$$

Select $D_n = C^* M_n \{\log(pn)\}^{1/2}$ for some sufficiently large constant $C^* > 0$. Thus, for any $x \ge 0$,

$$\mathbb{P}\left(\max_{j\in[p]} \left| \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} H_{b,j}^- \right| > x \right) \le \frac{p B_n^{1/2} \max_{j\in[p]} \max_{b\in[B_n]} \mathbb{E}(|H_{b,j}^-|)}{x} \lesssim \frac{(pn)^{-1}}{x}$$

provided that $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Then, by (A15), we can conclude that for any $x \gg M_n n^{-1/2}$,

$$\mathbb{P}\left(\left| \max_{j\in[p]} S_{n,j} - \max_{j\in[p]} S_{n,j}^+ \right| > x \right) \lesssim p B_n k_n \exp\left(-\frac{Cn^{\tilde{\gamma}/2} x^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}} \right) + p \exp\left(-\frac{Cnx^2}{M_n^2 B_n k_n} \right) + \frac{(pn)^{-1}}{x} \quad \text{(A17)}$$

provided that $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Similar to (A3), we have

$$\mathbb{E}(|\tilde{H}_{b,j} - H_{b,j}^+|) \lesssim D_n \alpha(k_n) \lesssim D_n \exp(-Ck_n^{\gamma_2}). \quad \text{(A18)}$$

Select $k_n = C^{**} \{\log(pn)\}^{1/\gamma_2}$ for some sufficiently large constant $C^{**} > 0$. By (A18) and triangle inequality,

$$\mathbb{P}\left(\left| \max_{j\in[p]} \tilde{S}_{n,j} - \max_{j\in[p]} S_{n,j}^+ \right| > x \right) \le \frac{p B_n^{1/2} \max_{b\in[B_n]} \max_{j\in[p]} \mathbb{E}(|\tilde{H}_{b,j} - H_{b,j}^+|)}{x} \lesssim \frac{(pn)^{-1}}{x}$$

for any $x \geq 0$. Thus, by (A17), for any $x \gg M_n n^{-1/2}$,

$$\mathbb{P}\left(\left|\max_{j\in[p]} S_{n,j} - \max_{j\in[p]} \tilde{S}_{n,j}\right| > x\right) \lesssim pB_n k_n \exp\left(-\frac{Cn^{\tilde{\gamma}/2}x^{\tilde{\gamma}}}{M_n^{\tilde{\gamma}}}\right) + p\exp\left(-\frac{Cnx^2}{M_n^2 B_n k_n}\right) + \frac{(pn)^{-1}}{x} \quad \text{(A19)}$$

provided that $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Let $\epsilon_2 = C^{***}M_n k_n^{1/2} K_n^{-1/2}\{\log(pn)\}^{1/2}$ for some sufficient large constant $C^{***} > 0$. It holds by (A14) that

$$\omega_n \lesssim \tilde{\omega}_n + \frac{M_n\{\log(pn)\}^{(2\gamma_2+1)/2\gamma_2}}{K_n^{1/2}} \quad \text{(A20)}$$

provided that $\log(pn) = o\{k_n^{\tilde{\gamma}/(2-\tilde{\gamma})} B_n^{\tilde{\gamma}/(2-\tilde{\gamma})} \wedge K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Define $\tilde{\mathbf{\Sigma}}_G = B_n^{-1}\sum_{b=1}^{B_n} \text{var}(\tilde{\mathbf{H}}_b)$ and $\mathbf{\Delta} = |\mathbf{\Sigma}_n - \tilde{\mathbf{\Sigma}}_G|$, where $\mathbf{\Sigma}_n = \mathbb{E}(\mathbf{S}_n \mathbf{S}_n^{\mathsf{T}})$. Note that

$$\mathbf{\Delta} = \left|\frac{1}{B_n}\sum_{b=1}^{B_n}\{\text{var}(\tilde{\mathbf{H}}_b) - \text{var}(\mathbf{H}_b^+)\} + \frac{1}{B_n}\sum_{b=1}^{B_n}\{\text{var}(\mathbf{H}_b^+) - \text{var}(\mathbf{H}_b)\} + \frac{1}{B_n}\sum_{b=1}^{B_n}\text{var}(\mathbf{H}_b) - \mathbf{\Sigma}_n\right|$$

$$\leq \underbrace{\frac{1}{B_n}\sum_{b=1}^{B_n}|\text{var}(\tilde{\mathbf{H}}_b) - \text{var}(\mathbf{H}_b^+)|}_{\mathbf{\Delta}_1} + \underbrace{\frac{1}{B_n}\sum_{b=1}^{B_n}|\text{var}(\mathbf{H}_b^+) - \text{var}(\mathbf{H}_b)|}_{\mathbf{\Delta}_2} + \underbrace{\left|\frac{1}{B_n}\sum_{b=1}^{B_n}\text{var}(\mathbf{H}_b) - \mathbf{\Sigma}_n\right|}_{\mathbf{\Delta}_3}. \quad \text{(A21)}$$

In this sequel, we will specify the convergence rates of $|\mathbf{\Delta}_1|_\infty$, $|\mathbf{\Delta}_2|_\infty$ and $|\mathbf{\Delta}_3|_\infty$, respectively. Note that the (i,j)-th element of $\text{var}(\tilde{\mathbf{H}}_b) - \text{var}(\mathbf{H}_b^+)$ is $\mathbb{E}(\tilde{H}_{b,i}\tilde{H}_{b,j} - H_{b,i}^+ H_{b,j}^+)$. Due to $\tilde{H}_{b,j}$ has the same distribution as $H_{b,j}^+$ and $|H_{b,j}^+| \lesssim D_n$ for any $b \in [B_n]$ and $j \in [p]$, it holds by (A18) that

$$|\mathbb{E}(\tilde{H}_{b,i}\tilde{H}_{b,j} - H_{b,i}^+ H_{b,j}^+)| \leq |\mathbb{E}\{(\tilde{H}_{b,i} - H_{b,i}^+)\tilde{H}_{b,j}\}| + |\mathbb{E}\{(\tilde{H}_{b,j} - H_{b,j}^+)\tilde{H}_{b,i}^+\}| \lesssim (pn)^{-1}$$

for any $b \in [B_n]$ and $i,j \in [p]$. Thus, we can conclude that $|\mathbf{\Delta}_1|_\infty \lesssim (pn)^{-1}$. The (i,j)-th element of $\text{var}(\mathbf{H}_b^+) - \text{var}(\mathbf{H}_b)$ is $\mathbb{E}(H_{b,i}^+ H_{b,j}^+ - H_{b,i}H_{b,j})$. Due to $H_{b,j} = H_{b,j}^+ + H_{b,j}^-$, then it holds by (A16) that

$$|\mathbb{E}(H_{b,i}^+ H_{b,j}^+ - H_{b,i}H_{b,j})| = |\mathbb{E}\{H_{b,i}^+ H_{b,j}^+ - (H_{b,i}^+ + H_{b,i}^-)(H_{b,j}^+ + H_{b,j}^-)\}|$$

$$\leq |\mathbb{E}(H_{b,i}^+ H_{b,j}^-)| + |\mathbb{E}(H_{b,j}^+ H_{b,i}^-)| + |\mathbb{E}(H_{b,i}^- H_{b,j}^-)| \lesssim (pn)^{-1}$$

for any $b \in [B_n]$ and $i,j \in [p]$ provided that $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Thus, we can conclude that $|\mathbf{\Delta}_2|_\infty \lesssim (pn)^{-1}$ provided that $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. The (i,j)-th element of $\mathbf{\Sigma}_n - B_n^{-1}\sum_{b=1}^{B_n} \text{var}(\mathbf{H}_b)$ is $n^{-1}\sum_{t_1,t_2=1}^n \mathbb{E}(\xi_{t_1,i}\xi_{t_2,j}) - n^{-1}\sum_{b=1}^{B_n}\sum_{t_1,t_2\in\mathcal{I}_b}\mathbb{E}(\xi_{t_1,i}\xi_{t_2,j})$, and

$$\left|\frac{1}{n}\sum_{t_1,t_2=1}^n \mathbb{E}(\xi_{t_1,i}\xi_{t_2,j}) - \frac{1}{n}\sum_{b=1}^{B_n}\sum_{t_1,t_2\in\mathcal{I}_b}\mathbb{E}(\xi_{t_1,i}\xi_{t_2,j})\right|$$

$$= \frac{1}{n}\left|\sum_{b_1\neq b_2}\mathbb{E}\left\{\left(\sum_{t\in\mathcal{G}_{b_1}}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{G}_{b_2}}\xi_{t,j}\right)\right\}\right.$$

$$\left.+ \sum_{b=1}^{B_n}\mathbb{E}\left\{\left(\sum_{t\in\mathcal{I}_b}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{J}_b}\xi_{t,j}\right) + \left(\sum_{t\in\mathcal{J}_b}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{G}_b}\xi_{t,j}\right)\right\}\right|. \quad \text{(A22)}$$

Note that $\mathbb{E}(|\xi_{t,j}|^r) \lesssim M_n^r$ for any constant integer $r > 0$. Equation (1.12b) of [30] yields

$$\left|\mathbb{E}\left\{\left(\sum_{t\in\mathcal{J}_b}\xi_{t,i}\right)\left(\sum_{t\in\mathcal{G}_b}\xi_{t,j}\right)\right\}\right|$$

$$= \left|\sum_{t\in\mathcal{J}_b}\text{cov}(\xi_{t,i},\xi_{t,j}) + \sum_{t_1\neq t_2:t_1,t_2\in\mathcal{J}_b}\text{cov}(\xi_{t_1,i},\xi_{t_2,j}) + \sum_{t_1\in\mathcal{J}_b}\sum_{t_2\in\mathcal{I}_b}\text{cov}(\xi_{t_1,i},\xi_{t_2,j})\right|$$

$$\lesssim M_n^2 k_n + \sum_{t_1\neq t_2:t_1,t_2\in\mathcal{J}_b}\{\mathbb{E}(|\xi_{t_1,i}|^3)\}^{\frac{1}{3}}\{\mathbb{E}(|\xi_{t_2,j}|^3)\}^{\frac{1}{3}}\alpha^{\frac{1}{3}}(|t_1-t_2|)$$

$$+ \sum_{t_1 \in \mathcal{J}_b} \sum_{t_2 \in \mathcal{I}_b} \{\mathbb{E}(|\xi_{t_1,i}|^3)\}^{\frac{1}{3}} \{\mathbb{E}(|\xi_{t_2,j}|^3)\}^{\frac{1}{3}} \alpha^{\frac{1}{3}} (|t_1 - t_2|) \lesssim M_n^2 k_n. \tag{A23}$$

Similarly, we can also obtain

$$\left| \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{I}_b} \xi_{t,i} \right) \left(\sum_{t \in \mathcal{J}_b} \xi_{t,j} \right) \right\} \right| \lesssim M_n^2 k_n.$$

Thus,

$$\left| \sum_{b=1}^{B_n} \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{I}_b} \xi_{t,i} \right) \left(\sum_{t \in \mathcal{J}_b} \xi_{t,j} \right) + \left(\sum_{t \in \mathcal{J}_b} \xi_{t,i} \right) \left(\sum_{t \in \mathcal{G}_b} \xi_{t,j} \right) \right\} \right| \lesssim M_n^2 k_n B_n.$$

By Equation (1.12b) of [30], if $b_1 < b_2$,

$$\left| \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{G}_{b_1}} \xi_{t,i} \right) \left(\sum_{t \in \mathcal{G}_{b_2}} \xi_{t,j} \right) \right\} \right|$$

$$\lesssim \sum_{t_1 \in \mathcal{G}_{b_1}} \sum_{t_2 \in \mathcal{G}_{b_2}} \{\mathbb{E}(|\xi_{t_1,i}|^3)\}^{\frac{1}{3}} \{\mathbb{E}(|\xi_{t_2,j}|^3)\}^{\frac{1}{3}} \alpha^{\frac{1}{3}} (|t_1 - t_2|)$$

$$\lesssim M_n^2 \sum_{\delta=1}^{K_n} \delta \exp[-C\{(b_2 - b_1 - 1)K_n + \delta\}^{\gamma_2}]$$

$$\lesssim M_n^2 \mathbb{1}(b_2 - b_1 = 1) + M_n^2 K_n^2 \exp\{-C(b_2 - b_1 - 1)^{\gamma_2} K_n{}^{\gamma_2}\} \mathbb{1}(b_2 - b_1 > 1).$$

Thus,

$$\sum_{b_1 < b_2} \left| \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{G}_{b_1}} \xi_{t,i} \right) \left(\sum_{t \in \mathcal{G}_{b_2}} \xi_{t,j} \right) \right\} \right|$$

$$\lesssim M_n^2 B_n + M_n^2 K_n^2 \sum_{b_2 - b_1 = 2}^{B_n - 1} \exp\{-C(b_2 - b_1 - 1)^{\gamma_2} K_n{}^{\gamma_2}\} \lesssim M_n^2 B_n.$$

Same result holds for $b_1 > b_2$. Thus we can conclude that

$$\sum_{b_1 \neq b_2} \left| \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{G}_{b_1}} \xi_{t,i} \right) \left(\sum_{t \in \mathcal{G}_{b_2}} \xi_{t,j} \right) \right\} \right| \lesssim M_n^2 B_n.$$

Note that the above upper bounds do not depend on (i, j). Then by (A22), it holds that $|\mathbf{\Delta}_3|_\infty \lesssim M_n^2 k_n K_n^{-1}$. By (A21), we can conclude that

$$|\mathbf{\Delta}|_\infty \lesssim \frac{M_n^2 k_n}{K_n} \tag{A24}$$

provided that $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Let $\{\tilde{\mathbf{H}}_b^{\mathrm{G}}\}_{b=1}^{B_n}$ be a sequence of independent Gaussian vector such that $\tilde{\mathbf{H}}_b^{\mathrm{G}} = (\tilde{H}_{b,1}^{\mathrm{G}}, \ldots, \tilde{H}_{b,p}^{\mathrm{G}})^{\mathsf{T}} \sim \mathcal{N}\{\mathbf{0}_p, \mathrm{var}(\tilde{\mathbf{H}}_b)\}$ for any $b \in [B_n]$, where $\tilde{\mathbf{H}}_b = (\tilde{H}_{b,1}, \ldots, \tilde{H}_{b,p})^{\mathsf{T}}$. Due to $k_n \asymp \{\log(pn)\}^{1/\gamma_2}$, we know that $\min_{j \in [p]} (\tilde{\mathbf{\Sigma}}_{\mathrm{G}})_{j,j} > c$ provided that $M_n^2 \{\log(pn)\}^{1/\gamma_2} = o(K_n)$ and $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. Due to $\tilde{H}_{b,j} \leq 2D_n \lesssim M_n\{\log(pn)\}^{1/2}$, it holds that $\mathbb{E}(\tilde{H}_{b,j}^4) \lesssim D_n^2 \mathbb{E}(\tilde{H}_{b,j}^2) \lesssim M_n^4 \log(pn)$ for any $b \in [B_n]$ and $j \in [p]$, where the last inequality follows from $\mathbb{E}(\tilde{H}_{b,j}^2) = \mathbb{E}(|H_{b,j}^+|^2) \lesssim \mathbb{E}(H_{b,j}^2)$ and the similar arguments as in the proof of (A23). By Theorem 2.1 of [16], we have

$$\sup_{x > 0} \left| \mathbb{P} \left(\max_{j \in [p]} \tilde{S}_{n,j} \leq x \right) - \mathbb{P} \left(\max_{j \in [p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j}^{\mathrm{G}} \leq x \right) \right| \lesssim \frac{M_n \{\log(pn)\}^{3/2}}{B_n^{1/4}}. \tag{A25}$$

provided that $M_n^2\{\log(pn)\}^{1/\gamma_2} = o(K_n)$ and $\log(pn) = o\{K_n^{\tilde{\gamma}/(2-\tilde{\gamma})}\}$. By Proposition 2.1 of [16] and (A24), we have

$$
\sup_{x>0} \left| \mathbb{P}\left(\max_{j\in[p]} \frac{1}{B_n^{1/2}} \sum_{b=1}^{B_n} \tilde{H}_{b,j}^{\mathrm{G}} \le x \right) - \mathbb{P}\left(\max_{j\in[p]} W_j \le x \right) \right|
$$

$$
\lesssim |\boldsymbol{\Delta}|_\infty^{1/2} \log p \lesssim \frac{M_n\{\log(pn)\}^{(2\gamma_2+1)/2\gamma_2}}{K_n^{1/2}}. \tag{A26}
$$

By (A20), (A25) and (A26), due to $\tilde{\gamma} = (1/\gamma_1 + 1/\gamma_2)^{-1}$, we have

$$
\omega_n \lesssim \frac{M_n\{\log(pn)\}^{3/2}}{B_n^{1/4}} + \frac{M_n\{\log(pn)\}^{(2\gamma_2+1)/2\gamma_2}}{K_n^{1/2}}
$$

provided that $\log(pn) = o\{B_n^{\gamma_1\gamma_2/(\gamma_1+2\gamma_2-\gamma_1\gamma_2)} \wedge K_n^{\gamma_1\gamma_2/(2\gamma_1+2\gamma_2-\gamma_1\gamma_2)}\}$ and $M_n^2\{\log(pn)\}^{1/\gamma_2} = o(K_n)$. Select $\varsigma = 2/3$. Then $B_n \asymp n^{2/3}$, $K_n \asymp n^{1/3}$ and

$$
\omega_n \lesssim \frac{M_n\{\log(pn)\}^{\max\{(2\gamma_2+1)/2\gamma_2, 3/2\}}}{n^{1/6}}
$$

provided that $M_n^2\{\log(pn)\}^{1/\gamma_2} = o(n^{1/3})$ and $\{\log(pn)\}^3 = o\{n^{\gamma_1\gamma_2/(2\gamma_1+2\gamma_2-\gamma_1\gamma_2)}\}$. Thus we complete the proof of Theorem 1(ii). $\square$

Appendix B. Proof of Proposition 1

Proof. Define

$$
\mathring{T}_n = \left| \frac{1}{\sqrt{\tilde{n}}} \sum_{t=1}^{\tilde{n}} \mathring{\mathbf{Z}}_t \right|_\infty,
$$

where $\mathring{\mathbf{Z}}_t = \mathbf{Z}_t - \mathbb{E}(\mathbf{Z}_t)$. Under H_0, we know that $\boldsymbol{\mu}_X = \boldsymbol{\mu}_Y =: \boldsymbol{\mu}$. Recall $n_1 \asymp n_2 \asymp n$ and $\Delta_n = n_1 \vee n_2 - n_1 \wedge n_2$. Without loss of generality, we assume $n_1 \le n_2$. By triangle inequality, for any $j \in [p]$,

$$
\left| \sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} - \sum_{t=1}^{\tilde{n}} Z_{t,j} \right| \lesssim \sum_{t=1}^{n_1} \left| \sqrt{\frac{n_2^2}{n_1(n_1+n_2)}} - \sqrt{\frac{n_1}{(n_1+n_2)}} \right| + \sum_{t=n_1+1}^{n_2} \left| \sqrt{\frac{n_1}{(n_1+n_2)}} \right| = O(\Delta_n).
$$

Thus $|T_n - \mathring{T}_n| = O(\Delta_n n^{-1/2})$. Write $\delta_n = \Delta_n n^{-1/2}\pi_n$, where $\pi_n > 0$ diverges at a sufficiently slow rate. Thus, we have

$$
\mathbb{P}(T_n \le x) \le \mathbb{P}(\mathring{T}_n \le x + \delta_n) + \mathbb{P}(|T_n - \mathring{T}_n| > \delta_n)
$$

$$
\le \mathbb{P}(T_n^{\mathrm{G}} \le x + \delta_n) + \sup_{x\in\mathbb{R}} |\mathbb{P}(\mathring{T}_n \le x) - \mathbb{P}(T_n^{\mathrm{G}} \le x)| + o(1)
$$

$$
\le \mathbb{P}(T_n^{\mathrm{G}} \le x) + \sup_{x\in\mathbb{R}} \mathbb{P}(x - \delta_n \le T_n^{\mathrm{G}} \le x + \delta_n) + \sup_{x\in\mathbb{R}} |\mathbb{P}(\mathring{T}_n \le x) - \mathbb{P}(T_n^{\mathrm{G}} \le x)| + o(1).
$$

Analogously, we can also obtain that $\mathbb{P}(T_n \le x) \ge \mathbb{P}(T_n^{\mathrm{G}} \le x) - \sup_{x\in\mathbb{R}} \mathbb{P}(x - \delta_n \le T_n^{\mathrm{G}} \le x + \delta_n) - \sup_{x\in\mathbb{R}} |\mathbb{P}(\mathring{T}_n \le x) - \mathbb{P}(T_n^{\mathrm{G}} \le x)| - o(1)$. Thus,

$$
\sup_{x\in\mathbb{R}} |\mathbb{P}(T_n \le x) - \mathbb{P}(T_n^{\mathrm{G}} \le x)|
$$

$$
\le \sup_{x\in\mathbb{R}} \mathbb{P}(x - \delta_n \le T_n^{\mathrm{G}} \le x + \delta_n) + \sup_{x\in\mathbb{R}} |\mathbb{P}(\mathring{T}_n \le x) - \mathbb{P}(T_n^{\mathrm{G}} \le x)| + o(1).
$$

In Case1, by Assumption 1(iii), we have $\min_{j\in[p]}(\Xi_{\tilde{n}})_{j,j} > c$. Then by Lemma A.1 of [31], due to $\Delta_n^2 \log p = o(n)$, we have $\sup_{x\in\mathbb{R}} \mathbb{P}(x - \delta_n \le T_n^{\mathrm{G}} \le x + \delta_n) \lesssim \Delta_n n^{-1/2}\pi_n$

$(\log p)^{1/2} = o(1)$. By Assumption 1(i), we have $\max_{t \in [\tilde{n}]} \max_{j \in [p]} \mathbb{E}(|\mathring{Z}_{t,j}|^m) \leq C$. Note that $\Xi_{\tilde{n}} = \mathbb{E}(\tilde{n}^{-1/2} \sum_{t=1}^{\tilde{n}} \mathring{Z}_t, \tilde{n}^{-1/2} \sum_{t=1}^{\tilde{n}} \mathring{Z}_t^{\mathrm{T}})$. Then by Assumption 1 and Theorem 1(i), due to $3m/(m-4) > \max\{2m/(m-3), 3\}$, we have $\sup_{x \in \mathbb{R}} |\mathbb{P}(\mathring{T}_n \leq x) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| = o(1)$ provided that $p^2 \log p = o\{n^{4\tau/(11\tau+12)}\}$. Thus, if $p^2 \log p = o\{n^{4\tau/(11\tau+12)}\}$,

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| = o(1).$$

Similarly, in Case2, by Assumption 2 and Theorem 1(ii) with $(M_n, \gamma_1, \gamma_2) = (C, 2, 1)$, we have $\sup_{x \in \mathbb{R}} \mathbb{P}(x - \delta_n \leq T_n^{\mathrm{G}} \leq x + \delta_n) \lesssim \Delta_n n^{-1/2} \pi_n (\log p)^{1/2} = o(1)$ and $\sup_{x \in \mathbb{R}} |\mathbb{P}(\mathring{T}_n \leq x) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| = o(1)$ provided that $\log(pn) = o(n^{1/9})$. Thus, if $\log(pn) = o(n^{1/9})$,

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| = o(1).$$

We complete the proof of Proposition 1. □

Appendix C. Proof of Theorem 2

Appendix C.1. Proof of Theorem 2 under Case3

Proof. By Proposition 1 under Case1, it suffices to show

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(\hat{T}_n^{\mathrm{G}} \leq x \,|\, \mathcal{E}) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| = o_{\mathrm{p}}(1).$$

Recall $T_n^{\mathrm{G}} = |\mathbf{G}|_\infty$ with $\mathbf{G} \sim \mathcal{N}(\mathbf{0}, \Xi_{\tilde{n}})$ and $\hat{T}_n^{\mathrm{G}} = |\tilde{n}^{-1/2} \sum_{t=1}^{\tilde{n}} (\mathbf{Z}_t - \bar{\mathbf{Z}}) \varrho_t'|_\infty$, where $\Xi_{\tilde{n}} = \mathrm{var}(\tilde{n}^{-1/2} \sum_{t=1}^{\tilde{n}} \mathbf{Z}_t)$. Let

$$\hat{\Xi}_{\tilde{n}} = \frac{1}{\tilde{n}} \sum_{b=1}^{B} \left\{ \left(\sum_{t \in \mathcal{I}_b} (\mathbf{Z}_t - \bar{\mathbf{Z}}) \right) \left(\sum_{t \in \mathcal{I}_b} (\mathbf{Z}_t - \bar{\mathbf{Z}}) \right)^{\mathrm{T}} \right\}. \tag{A27}$$

By Proposition 2.1 of [16], we have

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(\hat{T}_n^{\mathrm{G}} \leq x \,|\, \mathcal{E}) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| \lesssim \Gamma^{1/2} \log p, \tag{A28}$$

where

$$\Gamma = |\Xi_{\tilde{n}} - \hat{\Xi}_{\tilde{n}}|_\infty = \frac{1}{\tilde{n}} \left| \sum_{b=1}^{B} \left\{ \left(\sum_{t \in \mathcal{I}_b} (\mathbf{Z}_t - \bar{\mathbf{Z}}) \right) \left(\sum_{t \in \mathcal{I}_b} (\mathbf{Z}_t - \bar{\mathbf{Z}}) \right)^{\mathrm{T}} \right\} - \mathrm{var}\left(\sum_{t=1}^{\tilde{n}} \mathbf{Z}_t \right) \right|_\infty.$$

Let $\mathring{\mathbf{Z}}_t = (\mathring{Z}_{t,1}, \ldots, \mathring{Z}_{t,p})^{\mathrm{T}} = \mathbf{Z}_t - \mathbb{E}(\mathbf{Z}_t)$. Then, for any $i, j \in [p]$, triangle inequality yields

$$\left| \sum_{b=1}^{B} \left\{ \left(\sum_{t \in \mathcal{I}_b} (Z_{t,i} - \bar{Z}_i) \right) \left(\sum_{t \in \mathcal{I}_b} (Z_{t,j} - \bar{Z}_j) \right) \right\} - \mathbb{E} \left\{ \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i} \right) \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} \right) \right\} \right|$$

$$= \left| \sum_{b=1}^{B} \left\{ \left(\sum_{t \in \mathcal{I}_b} (\mathring{Z}_{t,i} - \bar{\mathring{Z}}_i) \right) \left(\sum_{t \in \mathcal{I}_b} (\mathring{Z}_{t,j} - \bar{\mathring{Z}}_j) \right) \right\} - \mathbb{E} \left\{ \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i} \right) \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} \right) \right\} \right|$$

$$\leq \left| \sum_{b=1}^{B} \left\{ \left(\sum_{t \in \mathcal{I}_b} (\mathring{Z}_{t,i} - \bar{\mathring{Z}}_i) \right) \left(\sum_{t \in \mathcal{I}_b} (\mathring{Z}_{t,j} - \bar{\mathring{Z}}_j) \right) \right\} - \sum_{b=1}^{B} \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i} \right) \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j} \right) \right\} \right|$$

$$+ \left| \sum_{b=1}^{B} \mathbb{E} \left\{ \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i} \right) \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j} \right) \right\} - \mathbb{E} \left\{ \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i} \right) \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} \right) \right\} \right|$$

$$\leq \underbrace{\left| \sum_{b=1}^{B} \sum_{t_1, t_2 \in \mathcal{I}_b} \{\mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j} - \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j})\} \right|}_{\mathrm{I}_{1,i,j}} + \underbrace{\frac{S}{\tilde{n}} \left| \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i} \right) \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} \right) \right|}_{\mathrm{I}_{2,i,j}}$$

$$+ \left| \sum_{b=1}^{B} \mathbb{E}\left\{ \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i} \right) \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j} \right) \right\} - \mathbb{E}\left\{ \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i} \right) \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} \right) \right\} \right|.$$
$$\underbrace{\phantom{\sum_{b=1}^{B}}}_{\mathrm{I}_{3,i,j}}$$

In this sequel, we will specify the upper bounds of $\mathrm{I}_{1,i,j}$, $\mathrm{I}_{2,i,j}$ and $\mathrm{I}_{3,i,j}$, respectively. Without loss of generality, we assume $\tilde{n} = BS$ with $B \asymp n^\vartheta$ and $S \asymp n^{1-\vartheta}$. By Assumption 1(i), it holds that $\max_{t \in [\tilde{n}]} \max_{j \in [p]} \mathbb{E}(|\mathring{Z}_{t,j}|^m) \leq C$ for some $m > 4$. Then, due to $\tau > 3m/(m-4)$, (A1) yields

$$\mathbb{E}\left(\left[\sum_{t_1, t_2 \in \mathcal{I}_b} \{ \mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j} - \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j}) \} \right]^2 \right) \leq \mathbb{E}\left\{ \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i} \right)^2 \left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j} \right)^2 \right\} \lesssim S^2.$$

By triangle inequality,

$$\mathbb{E}\left(\left[\sum_{b=1}^{B} \sum_{t_1, t_2 \in \mathcal{I}_b} \{ \mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j} - \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j}) \} \right]^2 \right)$$

$$\lesssim BS^2 + \sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left| \sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathrm{cov}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j}, \mathring{Z}_{t_3,i} \mathring{Z}_{t_4,j}) \right|$$

$$\leq BS^2 + \sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left| \sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathrm{cum}_{i,j}(t_2 - t_1, t_3 - t_1, t_4 - t_1) \right|$$

$$+ \sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left| \sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_3,i}) \mathbb{E}(\mathring{Z}_{t_2,j} \mathring{Z}_{t_4,j}) \right|$$

$$+ \sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left| \sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_4,j}) \mathbb{E}(\mathring{Z}_{t_3,i} \mathring{Z}_{t_2,j}) \} \right|. \tag{A29}$$

By Assumption 3, $\sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathrm{cum}_{i,j}(t_2 - t_1, t_3 - t_1, t_4 - t_1) \lesssim S$, which implies

$$\sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left| \sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathrm{cum}_{i,j}(t_2 - t_1, t_3 - t_1, t_4 - t_1) \right| \lesssim B^2 S. \tag{A30}$$

For any $b \in [B-1]$ and $s \in [B-b]$, due to $\tau > 3m/(m-4)$, Equation (1.12b) of [30] yields

$$\left| \sum_{t_1 \in \mathcal{I}_b} \sum_{t_3 \in \mathcal{I}_{b+s}} \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_3,i}) \right| \leq \sum_{t_1 \in \mathcal{I}_b} \sum_{t_3 \in \mathcal{I}_{b+s}} |\mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_3,i})|$$

$$\lesssim \sum_{t_1 \in \mathcal{I}_b} \sum_{t_3 \in \mathcal{I}_{b+s}} \{ \mathbb{E}(|Z_{t_1,i}|^m) \}^{\frac{1}{m}} \{ \mathbb{E}(|Z_{t_3,i}|^m) \}^{\frac{1}{m}} \alpha^{\frac{m-2}{m}}(t_3 - t_1) \lesssim \sum_{t_1 \in \mathcal{I}_b} \sum_{t_3 \in \mathcal{I}_{b+s}} \alpha^{\frac{m-2}{m}}(t_3 - t_1)$$

$$\lesssim \sum_{h=1}^{S} h \alpha^{\frac{m-2}{m}} \{ h + (s-1)S \} \lesssim 1(s=1) + S^{\frac{2m-(m-2)\tau}{m}}(s-1)^{-\frac{(m-2)\tau}{m}} 1(s>1).$$

Similarly, we also have

$$\left| \sum_{t_2 \in \mathcal{I}_b} \sum_{t_4 \in \mathcal{I}_{b+s}} \mathbb{E}(\mathring{Z}_{t_2,j} \mathring{Z}_{t_4,j}) \right| \lesssim 1(s=1) + S^{\frac{2m-(m-2)\tau}{m}}(s-1)^{-\frac{(m-2)\tau}{m}} 1(s>1).$$

Thus,

$$\sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left| \sum_{t_1, t_2 \in \mathcal{I}_b} \sum_{t_3, t_4 \in \mathcal{I}_{b+s}} \mathbb{E}(\mathring{Z}_{t_1,i} \mathring{Z}_{t_3,i}) \mathbb{E}(\mathring{Z}_{t_2,j} \mathring{Z}_{t_4,j}) \right|$$

$$\lesssim \sum_{b=1}^{B-1} \sum_{s=1}^{B-b} \left\{ 1(s=1) + S^{\frac{4m-2(m-2)\tau}{m}}(s-1)^{-\frac{2(m-2)\tau}{m}} 1(s>1) \right\} \lesssim B. \tag{A31}$$

Analogously, we also have $\sum_{b=1}^{B-1} \sum_{s=1}^{B-b} |\sum_{t_1,t_2 \in \mathcal{I}_b} \sum_{t_3,t_4 \in \mathcal{I}_{b+s}} \mathbb{E}(\mathring{Z}_{t_1,i}\mathring{Z}_{t_4,j})\mathbb{E}(\mathring{Z}_{t_3,i}\mathring{Z}_{t_2,j})\}| \lesssim B$. Combining this with (A29)–(A31), due to $B \geq S$,

$$\mathbb{E}\left(\left[\sum_{b=1}^{B} \sum_{t_1,t_2 \in \mathcal{I}_b} \{\mathring{Z}_{t_1,i}\mathring{Z}_{t_2,j} - \mathbb{E}(\mathring{Z}_{t_1,i}\mathring{Z}_{t_2,j})\}\right]^2\right) \lesssim B^2 S.$$

Then it holds that

$$\mathrm{I}_{1,i,j} = O_{\mathrm{p}}(BS^{1/2}). \tag{A32}$$

Similar to (A1), we have $|(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i})(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j})| = O_{\mathrm{p}}(n)$. Thus, we know that

$$\mathrm{I}_{2,i,j} = O_{\mathrm{p}}(S). \tag{A33}$$

Note that

$$\mathrm{I}_{3,i,j} \leq \sum_{b_1 \neq b_2} \left| \mathbb{E}\left\{ \left(\sum_{t \in \mathcal{I}_{b_1}} \mathring{Z}_{t,i} \right) \left(\sum_{t \in \mathcal{I}_{b_2}} \mathring{Z}_{t,j} \right) \right\} \right|.$$

For $b_1 < b_2$, due to $\tau > 3m/(m-4)$, Equation (1.12b) of [30] yields

$$\left| \mathbb{E}\left\{ \left(\sum_{t \in \mathcal{I}_{b_1}} \mathring{Z}_{t,i} \right) \left(\sum_{t \in \mathcal{I}_{b_2}} \mathring{Z}_{t,j} \right) \right\} \right|$$

$$\lesssim \sum_{s=1}^{S} s\{\mathbb{E}(|Z_{t,i}|^m)\}^{\frac{1}{m}} \{\mathbb{E}(|Z_{t+s,j}|^m)\}^{\frac{1}{m}} \alpha^{\frac{m-2}{m}} \{s + (b_2 - b_1 - 1)S\}$$

$$\lesssim 1(b_2 - b_1 = 1) + S^{\frac{2m-(m-2)\tau}{m}}(b_2 - b_1 - 1)^{-\frac{(m-2)\tau}{m}} 1(b_2 - b_1 > 1).$$

Thus,

$$\sum_{b_1 < b_2} \left| \mathbb{E}\left\{ \left(\sum_{t \in \mathcal{I}_{b_1}} \mathring{Z}_{t,i} \right) \left(\sum_{t \in \mathcal{I}_{b_2}} \mathring{Z}_{t,j} \right) \right\} \right| \lesssim B + S^{\frac{2m-(m-2)\tau}{m}} \sum_{b_2 - b_1 > 1} (b_2 - b_1 - 1)^{-\frac{(m-2)\tau}{m}} \lesssim B.$$

Similarly, we can also obtain $\sum_{b_1 > b_2} |\mathbb{E}\{(\sum_{t \in \mathcal{I}_{b_1}} \mathring{Z}_{t,i})(\sum_{t \in \mathcal{I}_{b_2}} \mathring{Z}_{t,j})\}| \lesssim B$, which implies $\mathrm{I}_{3,i,j} \lesssim B$. Then by (A32) and (A33), it holds that

$$\frac{1}{\tilde{n}} \left| \sum_{b=1}^{B} \left(\sum_{t \in \mathcal{I}_b} (Z_{t,i} - \bar{Z}_i) \right) \left(\sum_{t \in \mathcal{I}_b} (Z_{t,j} - \bar{Z}_j) \right) - \mathbb{E}\left\{ \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i} \right) \left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j} \right) \right\} \right| = O_{\mathrm{p}}(S^{-1/2}).$$

Then, by Markov's inequality,

$$\Gamma = |\Xi_{\tilde{n}} - \hat{\Xi}_{\tilde{n}}|_\infty = O_{\mathrm{p}}(p^2 S^{-1/2}). \tag{A34}$$

By (A28), due to $S \asymp n^{1-\vartheta}$, it holds that

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(\hat{T}_n^{\mathrm{G}} \leq x \mid \mathcal{E}) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| = o_{\mathrm{p}}(1) \tag{A35}$$

provided that $p \log p = o\{n^{(1-\vartheta)/4}\}$.

Recall $\hat{\mathrm{cv}}_\alpha = \inf\{x > 0 : \mathbb{P}(\hat{T}_n^{\mathrm{G}} > x \mid \mathcal{E}) \leq \alpha\}$. For any $\epsilon > 0$, let $\mathrm{cv}_\alpha^{(\epsilon)}$ and $\mathrm{cv}_\alpha^{(-\epsilon)}$ be two constants which satisfy $\mathbb{P}\{T_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(\epsilon)}\} = \alpha + \epsilon$ and $\mathbb{P}\{T_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(-\epsilon)}\} = \alpha - \epsilon$,

respectively. We claim that for any $\epsilon > 0$, it holds that $\mathbb{P}\{\mathrm{cv}_\alpha^{(\epsilon)} < \hat{\mathrm{cv}}_\alpha < \mathrm{cv}_\alpha^{(-\epsilon)}\} \to 1$ as $n \to \infty$. Otherwise, if $\hat{\mathrm{cv}}_\alpha \leq \mathrm{cv}_\alpha^{(\epsilon)}$, by (A35), we have

$$\alpha = \mathbb{P}(\hat{T}_n^{\mathrm{G}} > \hat{\mathrm{cv}}_\alpha \,|\, \mathcal{E}) \geq \mathbb{P}\{\hat{T}_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(\epsilon)} \,|\, \mathcal{E}\} = \mathbb{P}\{T_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(\epsilon)}\} + o_{\mathrm{p}}(1) = \alpha + \epsilon + o_{\mathrm{p}}(1),$$

which is a contradiction with probability approaching one as $n \to \infty$. Analogously, if $\hat{\mathrm{cv}}_\alpha \geq \mathrm{cv}_\alpha^{(-\epsilon)}$, by (A35), we have

$$\alpha = \mathbb{P}(\hat{T}_n^{\mathrm{G}} > \hat{\mathrm{cv}}_\alpha \,|\, \mathcal{E}) \leq \mathbb{P}\{\hat{T}_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(-\epsilon)} \,|\, \mathcal{E}\} = \mathbb{P}\{T_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(-\epsilon)}\} + o_{\mathrm{p}}(1) = \alpha - \epsilon + o_{\mathrm{p}}(1),$$

which is also a contradiction with probability approaching one as $n \to \infty$.

For any $\epsilon > 0$, define the event $\mathcal{E}_{1,\epsilon} = \{\mathrm{cv}_\alpha^{(\epsilon)} < \hat{\mathrm{cv}}_\alpha < \mathrm{cv}_\alpha^{(-\epsilon)}\}$. Then $\mathbb{P}(\mathcal{E}_{1,\epsilon}) \to 1$ as $n \to \infty$. On the one hand, by Proposition 1,

$$\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) \leq \mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha \,|\, \mathcal{E}_{1,\epsilon}) + \mathbb{P}(\mathcal{E}_{1,\epsilon}^c) \leq \mathbb{P}\{T_n > \mathrm{cv}_\alpha^{(\epsilon)}\} + o(1)$$
$$= \mathbb{P}\{T_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(\epsilon)}\} + o(1) = \alpha + \epsilon + o(1),$$

which implies that $\overline{\lim}_{n\to\infty}\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) \leq \alpha + \epsilon$. On the other hand, by Proposition 1,

$$\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) \geq \mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha \,|\, \mathcal{E}_{1,\epsilon}) \geq \mathbb{P}\{T_n > \mathrm{cv}_\alpha^{(-\epsilon)}\} - \mathbb{P}(\mathcal{E}_{1,\epsilon}^c)$$
$$\geq \mathbb{P}\{T_n^{\mathrm{G}} > \mathrm{cv}_\alpha^{(-\epsilon)}\} - o(1) = \alpha - \epsilon - o(1),$$

which implies that $\underline{\lim}_{n\to\infty}\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) \geq \alpha - \epsilon$. Since $\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha)$ does not depend on ϵ, by letting $\epsilon \to 0^+$, we have $\lim_{n\to\infty}\mathbb{P}(T_n > \hat{\mathrm{cv}}_\alpha) = \alpha$. Thus we complete the proof of Theorem 2 under Case3. □

Appendix C.2. Proof of Theorem 2 under Case4

Proof. By Proposition 1 under Case2 and the arguments in Appendix C.1, it suffices to show

$$\sup_{x\in\mathbb{R}} |\mathbb{P}(\hat{T}_n^{\mathrm{G}} \leq x \,|\, \mathcal{E}) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| \lesssim \Gamma^{1/2} \log p = o_{\mathrm{p}}(1),$$

where

$$\Gamma \leq \max_{i,j\in[p]} \tilde{n}^{-1}|\mathrm{I}_{1,i,j}| + \max_{i,j\in[p]} \tilde{n}^{-1}|\mathrm{I}_{2,i,j}| + \max_{i,j\in[p]} \tilde{n}^{-1}|\mathrm{I}_{3,i,j}|$$

with $\mathrm{I}_{1,i,j}$, $\mathrm{I}_{2,i,j}$ and $\mathrm{I}_{3,i,j}$ specified in Appendix C.1. In this sequel, we will specify the upper bounds of $\max_{i,j\in[p]} |\mathrm{I}_{1,i,j}|$, $\max_{i,j\in[p]} |\mathrm{I}_{2,i,j}|$ and $\max_{i,j\in[p]} |\mathrm{I}_{3,i,j}|$, respectively.

Without loss of generality, we assume $\tilde{n} = BS$ with $B \asymp \tilde{n}^\vartheta$ and $S \asymp \tilde{n}^{1-\vartheta}$ for some $\vartheta \in [1/2, 1)$. Let $W_{b,i,j} = \sum_{t_1,t_2\in\mathcal{I}_b} \mathring{Z}_{t_1,i}\mathring{Z}_{t_2,j} - \mathbb{E}(\sum_{t_1,t_2\in\mathcal{I}_b} \mathring{Z}_{t_1,i}\mathring{Z}_{t_2,j})$. For $R_n > C^*S$ with some sufficiently large constant $C^* > 0$, denote $W_{b,i,j}^+ = W_{b,i,j}\mathbf{1}(|W_{b,i,j}| \leq R_n) - \mathbb{E}\{W_{b,i,j}\mathbf{1}(|W_{b,i,j}| \leq R_n)\}$ and $W_{b,i,j}^- = W_{b,i,j}\mathbf{1}(|W_{b,i,j}| > R_n) - \mathbb{E}\{W_{b,i,j}\mathbf{1}(|W_{b,i,j}| > R_n)\}$. Then for some $C_n > 0$, it holds by Bonferroni inequality that

$$\mathbb{P}\left(\max_{i,j\in[p]} \left|\sum_{b=1}^B W_{b,i,j}\right| > \tilde{n}x\right) \leq p^2 \max_{i,j\in[p]} \mathbb{P}\left(\left|\sum_{b=1}^B W_{b,i,j}^+\right| + \left|\sum_{b=1}^B W_{b,i,j}^-\right| > \tilde{n}x\right)$$

$$\leq p^2 \max_{i,j\in[p]} \left\{\mathbb{P}\left(\left|\sum_{b=1}^B W_{b,i,j}^+\right| > \tilde{n}x - C_n\right) + \mathbb{P}\left(\left|\sum_{b=1}^B W_{b,i,j}^-\right| > C_n\right)\right\}$$

for all $x > C_n\tilde{n}^{-1}$. Note that

$$\mathbb{E}\{W_{b,i,j}^2\mathbf{1}(|W_{b,i,j}| > R_n)\} = 2\int_0^{R_n} x\mathbb{P}(|W_{b,i,j}| > R_n)\mathrm{d}x + 2\int_{R_n}^\infty x\mathbb{P}(|W_{b,i,j}| > x)\mathrm{d}x$$

By Assumptions 2(i)–(ii) and Cauchy–Schwarz inequality, $\mathbb{E}(\sum_{t_1,t_2 \in \mathcal{I}_b} \mathring{Z}_{t_1,i} \mathring{Z}_{t_2,j}) \lesssim S$. By Assumptions 2(i)–(ii) again and Theorem 1 of [32], we know that

$$\mathbb{P}\left(\left|\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i}\right| > x\right) \lesssim S \exp(-Cx^{2/3}) + \exp(-CS^{-1}x^2)$$

for any $x \to \infty$. Thus, for any $x > CS$, we have

$$\mathbb{P}(|W_{b,i,j}| > x) \leq \mathbb{P}\left(\left|\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i}\right|\left|\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j}\right| > Cx\right)$$

$$\leq \mathbb{P}\left(\left|\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i}\right| > Cx^{1/2}\right) + \mathbb{P}\left(\left|\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j}\right| > Cx^{1/2}\right) \lesssim S \exp(-Cx^{1/3}) + \exp(-CS^{-1}x).$$

Due to $R_n > CS$, we can show that

$$\mathbb{E}(|W_{b,i,j}^-|^2) \lesssim \mathbb{E}\{W_{b,i,j}^2 1(|W_{b,i,j}| > R_n)\} \lesssim R_n^2 S \exp(-CR_n^{1/3}) + R_n^2 \exp(-CS^{-1}R_n).$$

Selecting $R_n = C^{**}S \log(pn)$ for some sufficiently large constant $C^{**} > 0$, and $C_n \asymp B^{1/2}$, it holds by Markov's inequality that

$$p^2 \max_{i,j \in [p]} \mathbb{P}\left(\left|\sum_{b=1}^B W_{b,i,j}^-\right| > C_n\right) \lesssim p^2 B^{1/2} \max_{i,j \in [p]} \max_{b \in [B]} \mathbb{E}(|W_{b,i,j}^-|) = o(1)$$

provided that $\log(pn) = o(S^{1/2})$. Due to $|W_{b,i,j}^+| \leq 2R_n$, by Theorem 1 of [33],

$$\mathbb{P}\left(\left|\sum_{b=1}^B W_{b,i,j}^+\right| > \tilde{n}x - C_n\right) \lesssim \exp\left\{-\frac{C\tilde{n}^2 x^2}{BR_n^2 + R_n\tilde{n}x(\log B)(\log\log B)}\right\}$$

for any $x > CB^{-1/2}S^{-1}$. Thus, we can conclude that

$$\max_{i,j \in [p]} \tilde{n}^{-1}|\mathrm{I}_{1,i,j}| = \max_{i,j \in [p]}\left|\frac{1}{\tilde{n}}\sum_{b=1}^B W_{b,i,j}\right| = O_{\mathrm{p}}\left[\frac{\{\log(pn)\}^{3/2}}{B^{1/2}}\right]$$

provided that $\log(pn) = o[\min\{S^{1/2}, B(\log n \log\log n)^{-2}\}]$. By Bonferroni inequality and Theorem 1 of [32], we know that

$$\mathbb{P}\left(\max_{i,j \in [p]} \tilde{n}^{-1}|\mathrm{I}_{2,i,j}| > x\right) \lesssim p^2 \max_{i \in [p]} \mathbb{P}\left(\left|\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i}\right| > CS^{-1/2}nx^{1/2}\right)$$

$$\lesssim p^2 n \exp(-CS^{-1/3}n^{2/3}x^{1/3}) + p^2 \exp(-CS^{-1}nx)$$

for any $x \gg Sn^{-2}$. Then, we can conclude that $\max_{i,j \in [p]} \tilde{n}^{-1}|\mathrm{I}_{2,i,j}| = O_{\mathrm{p}}\{B^{-1}\log(pn)\}$ provided that $\log(pn) = o(n^{1/2})$. Finally, Equation (1.12b) of [30] yields

$$\left|\sum_{b=1}^B \mathbb{E}\left\{\left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,i}\right)\left(\sum_{t \in \mathcal{I}_b} \mathring{Z}_{t,j}\right)\right\} - \mathbb{E}\left\{\left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,i}\right)\left(\sum_{t=1}^{\tilde{n}} \mathring{Z}_{t,j}\right)\right\}\right|$$

$$\leq \sum_{b=1}^{B-1}\sum_{\kappa=1}^{B-b}\left|\sum_{t_1 \in \mathcal{I}_b}\sum_{t_2 \in \mathcal{I}_{b+\kappa}} \mathbb{E}(\mathring{Z}_{t_1,i}\mathring{Z}_{t_2,j})\right| \lesssim \sum_{b=1}^{B-1}\sum_{\kappa=1}^{B-b}\sum_{\delta=1}^B \delta\exp[-C\{\delta + (\kappa-1)S\}]$$

$$\lesssim \sum_{b=1}^{B-1}\sum_{\delta=1}^B \delta\exp(-C\delta) + \sum_{b=1}^{B-1}\sum_{\kappa=2}^{B-b} B^2\exp\{-C(\kappa-1)S\} \lesssim B + B^3\exp(-CS) \lesssim B,$$

which implies $\max_{i,j \in [p]} \tilde{n}^{-1} |\mathrm{I}_{3,i,j}| = O(S^{-1})$. Thus,

$$\Gamma = O_{\mathrm{P}} \left[\frac{\{\log(pn)\}^{3/2}}{B^{1/2}} + \frac{1}{S} \right] \tag{A36}$$

provided that $\log(pn) = o[\min\{S^{1/2}, B(\log n \log \log n)^{-2}\}]$. It holds that

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(\hat{T}_n^{\mathrm{G}} \leq x \mid \mathcal{E}) - \mathbb{P}(T_n^{\mathrm{G}} \leq x)| \lesssim \Gamma^{1/2} \log p = o_{\mathrm{p}}(1)$$

provided that $\log(pn) = o[n^{\min\{(1-\vartheta)/2, \vartheta/7\}}]$. The proof of the second result of Theorem 2 under Case4 is the same as in the proof of the second result of Theorem 2 under Case3. Thus, we complete the proof of Theorem 2. $\square$

Appendix D. Proof of Theorem 3

Proof. Let $s = C\pi_n p^2 S^{-1/2}$ in Case3 and $s = C\pi_n [B^{-1/2}\{\log(pn)\}^{3/2} + CS^{-1}]$ in Case4, where $\pi_n > 0$ diverges at a sufficiently slow rate. Then $s = o(1)$ provided that $p = o(S^{1/4})$ in Case3 and $\log(pn) = o(B^{1/3})$ in Case4. Define an event

$$\Phi(s) = \left\{ \max_{j \in [p]} \left| \frac{(\hat{\Xi}_{\tilde{n}})_{j,j}}{(\Xi_{\tilde{n}})_{j,j}} - 1 \right| \leq s \right\},$$

where $\hat{\Xi}_{\tilde{n}}$ and $\Xi_{\tilde{n}}$ are specified in (A27) and (3), respectively. By (A34) and (A36) in Appendix C, we have

$$\max_{j \in [p]} |(\hat{\Xi}_{\tilde{n}})_{j,j} - (\Xi_{\tilde{n}})_{j,j}| \leq |\hat{\Xi}_{\tilde{n}} - \Xi_{\tilde{n}}|_\infty = o_{\mathrm{p}}(s)$$

holds under Case3 and Case4 with $\log(pn) = o(S^{1/2})$. By Assumption 1(iii) and Assumption 2(iii) , we know that $\min_{j \in [p]} (\Xi_{\tilde{n}})_{j,j} > c$ holds under Case3 and Case4. Therefore,

$$\max_{j \in [p]} \left| \frac{(\hat{\Xi}_{\tilde{n}})_{j,j}}{(\Xi_{\tilde{n}})_{j,j}} - 1 \right| \leq \frac{\max_{j \in [p]} |(\hat{\Xi}_{\tilde{n}})_{j,j} - (\Xi_{\tilde{n}})_{j,j}|}{\min_{j \in [p]} (\Xi_{\tilde{n}})_{j,j}} = o_{\mathrm{p}}(s).$$

Then it holds that $\mathbb{P}\{\Phi^c(s) \mid \mathcal{E}\} = o_{\mathrm{p}}(1)$ under Case3 and Case4. Let $\varrho = \max_{j \in [p]} (\Xi_{\tilde{n}})_{j,j}$. Restricted on $\Phi(s)$, there exists a constant $C_0 > 0$ such that

$$\mathbb{E}(\hat{T}_n^{\mathrm{G}} \mid \mathcal{E}) \leq C_0 (\log p)^{1/2} \max_{j \in [p]} (\hat{\Xi}_{\tilde{n}})_{j,j}^{1/2} \leq (1+s)^{1/2} C_0 (\log p)^{1/2} \varrho^{1/2}.$$

By Borell inequality for Gaussian process,

$$\mathbb{P}\{\hat{T}_n^{\mathrm{G}} > \mathbb{E}(\hat{T}_n^{\mathrm{G}} \mid \mathcal{E}) + x \mid \mathcal{E}\} \leq 2\exp\left\{ -\frac{x^2}{2\max_{j \in [p]} (\hat{\Xi}_{\tilde{n}})_{j,j}} \right\}$$

for any $x > 0$. Let $x_0 = \varrho^{1/2}(1+s)^{1/2}[C_0(\log p)^{1/2} + \{2\log(4/\alpha)\}^{1/2}]$. Restricted on $\Phi(s)$, we have

$$x_0 \geq \mathbb{E}(\hat{T}_n^{\mathrm{G}} \mid \mathcal{E}) + (2\varrho)^{1/2}(1+s)^{1/2} \log^{1/2}\left(\frac{4}{\alpha}\right),$$

which implies

$$\mathbb{P}\{\hat{T}_n^{\mathrm{G}} > x_0, \Phi(s) \mid \mathcal{E}\} \leq 2\exp\left\{ -\frac{2\varrho(1+s)\log(4/\alpha)}{2\varrho(1+s)} \right\} = \frac{\alpha}{2}.$$

Since $\mathbb{P}\{\Phi^c(s) \mid \mathcal{E}\} = o_{\mathrm{p}}(1)$, then $\mathbb{P}\{\Phi^c(s) \mid \mathcal{E}\} \leq \alpha/4$ with probability approaching one. Hence, $\mathbb{P}(\hat{T}_n^{\mathrm{G}} > x_0 \mid \mathcal{E}) \leq \alpha$ with probability approaching one. Similar to the proof of (A1),

we know that $\varrho \leq C$ under Case3 and Case4. By the definition of $\hat{cv}_\alpha$, it holds with probability approaching one that

$$\hat{cv}_\alpha \leq \varrho^{1/2}(1+s)^{1/2}[C_0(\log p)^{1/2} + \{2\log(4/\alpha)\}^{1/2}] \lesssim (\log p)^{1/2}$$

under Case3 with $p = o(S^{1/4})$ and Case4 with $\log(pn) = o(B^{1/3} \wedge S^{1/2})$. Let $\mu_X = (\mu_{X,1}, \ldots, \mu_{X,p})^\mathsf{T} = \mathbb{E}(X_t)$, $\mu_Y = (\mu_{Y,1}, \ldots, \mu_{Y,p})^\mathsf{T} = \mathbb{E}(Y_t)$ and $j_0 = \arg\max_{j \in [p]} |\mu_{X,j} - \mu_{Y,j}|$, then

$$
\begin{aligned}
T_n &= \sqrt{\frac{n_1 n_2}{n_1 + n_2}} |\hat{\mu}_X - \hat{\mu}_Y|_\infty \geq \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left| \frac{1}{n_1} \sum_{t=1}^{n_1} X_{t,j_0} - \frac{1}{n_2} \sum_{t=1}^{n_2} Y_{t,j_0} \right| \\
&= \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left| \frac{1}{n_1} \sum_{t=1}^{n_1} (X_{t,j_0} - \mu_{X,j_0}) - \frac{1}{n_2} \sum_{t=1}^{n_2} (Y_{t,j_0} - \mu_{Y,j_0}) + \mu_{X,j_0} - \mu_{Y,j_0} \right| \\
&\geq \sqrt{\frac{n_1 n_2}{n_1 + n_2}} |\mu_{X,j_0} - \mu_{Y,j_0}| - \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left| \frac{1}{n_1} \sum_{t=1}^{n_1} (X_{t,j_0} - \mu_{X,j_0}) - \frac{1}{n_2} \sum_{t=1}^{n_2} (Y_{t,j_0} - \mu_{Y,j_0}) \right|.
\end{aligned}
$$

Similar to the proof of (A1), we know that

$$\left| \frac{1}{n_1} \sum_{t=1}^{n_1} (X_{t,j_0} - \mu_{X,j_0}) - \frac{1}{n_2} \sum_{t=1}^{n_2} (Y_{t,j_0} - \mu_{Y,j_0}) \right| = O_{\mathrm{p}}(n^{-1/2})$$

under Case3 and Case4. If $|\mu_{X,j_0} - \mu_{Y,j_0}| \gg n^{-1/2}$, we can conclude that $\mathbb{P}(T_n \geq C^* n^{1/2} |\mu_{X,j_0} - \mu_{Y,j_0}|) \to 1$ as $n \to \infty$ for some constant $C^* > 0$. Due to $\hat{cv}_\alpha \lesssim (\log p)^{1/2} = o(n^{1/2}|\mu_{X,j_0} - \mu_{Y,j_0}|)$ under Case3 and Case4, we have that Theorem 3 holds under Case3 and Case4. □
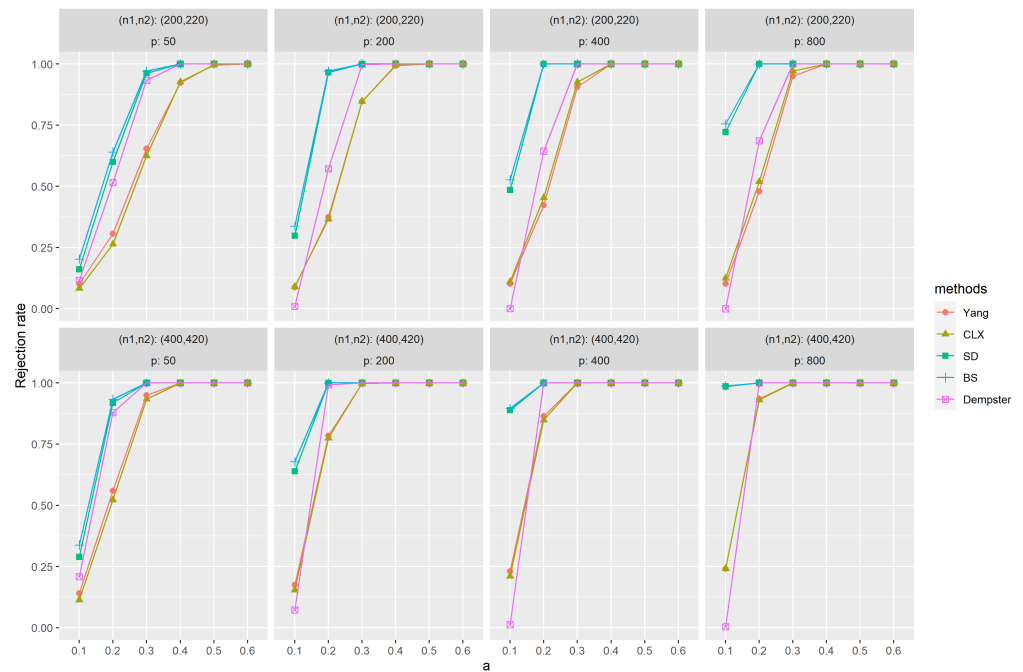
Appendix E. Additional Simulation Results

Table A1. The Type I error rates, expressed as percentages, were calculated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M2). The simulations were replicated 1000 times.

| (n_1, n_2)	ρ	p	Yang	Dempster	BS	SD	CLX
(200,220)	0	50	4.6	2.3	6.4	4.5	3
		200	3.3	0	6.7	5.8	3.4
		400	3.7	0	5	4.4	4.1
		800	3.3	0	6.4	5.2	4.2
	0.1	50	6.2	12.4	23.4	18.4	9.8
		200	5.2	1.8	43.2	39.5	12.9
		400	5.8	0.1	63	59.7	14.7
		800	5.3	0	88.7	87.3	19.4
	0.2	50	8.1	35.6	51.3	44.3	21.9
		200	7.7	23.5	87.2	85.5	37.9
		400	9.2	16.9	98.5	98.3	43
		800	9.6	9.8	100	100	54.5
(400,420)	0	50	4.9	1.8	5.4	3.6	3.5
		200	3.5	0	6.4	5.3	3.2
		400	5	0	5.7	4.8	4.4
		800	4.3	0	6	5	4.5
	0.1	50	6.9	12.2	21.7	17	9.3
		200	4.9	1.9	41.7	39	11.1
		400	7.8	0.1	63.6	61.4	17.7
		800	7.3	0	87.9	86.7	18.3
	0.2	50	8.6	33.7	46.9	40.7	20.6
		200	7.7	23.7	86.3	84.7	31
		400	9.4	17.1	99.2	99	43.6
		800	9	9.5	100	100	53.2
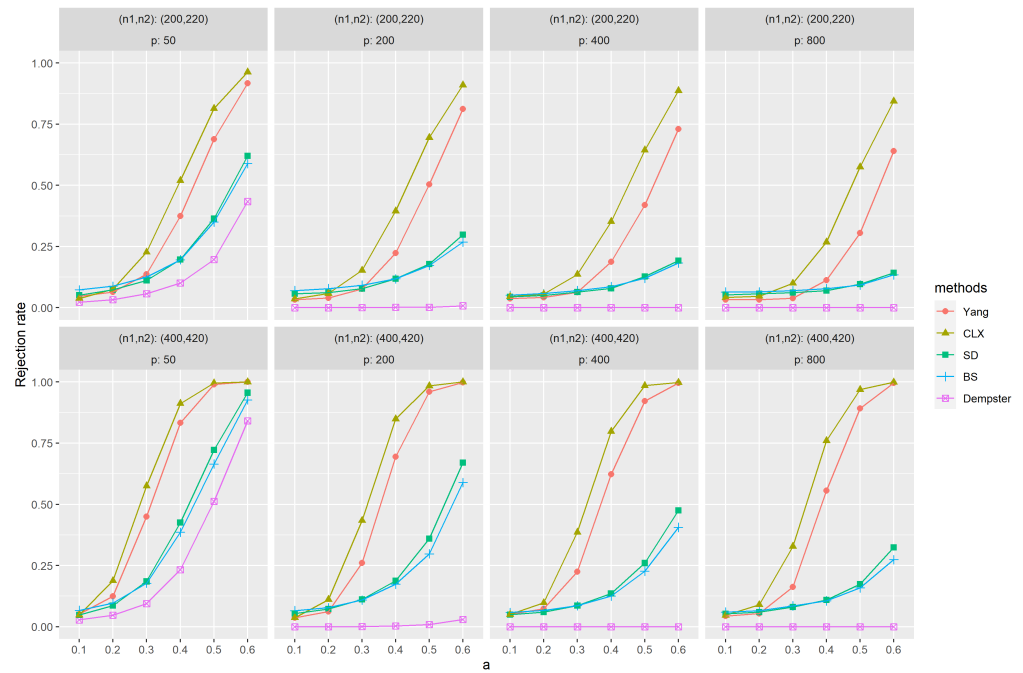
Figure A1. The empirical powers with sparse signals were evaluated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ based on (M2), $\mathbf{f}(\cdot) = \mathbf{0}$ and $\rho = 0$, and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M2), $\mathbf{f}(\cdot) = \mathbf{f}_1(\cdot)$ and $\rho = 0$. The parameter a represents the distance between the null and alternative hypotheses. The simulations were replicated 1000 times.
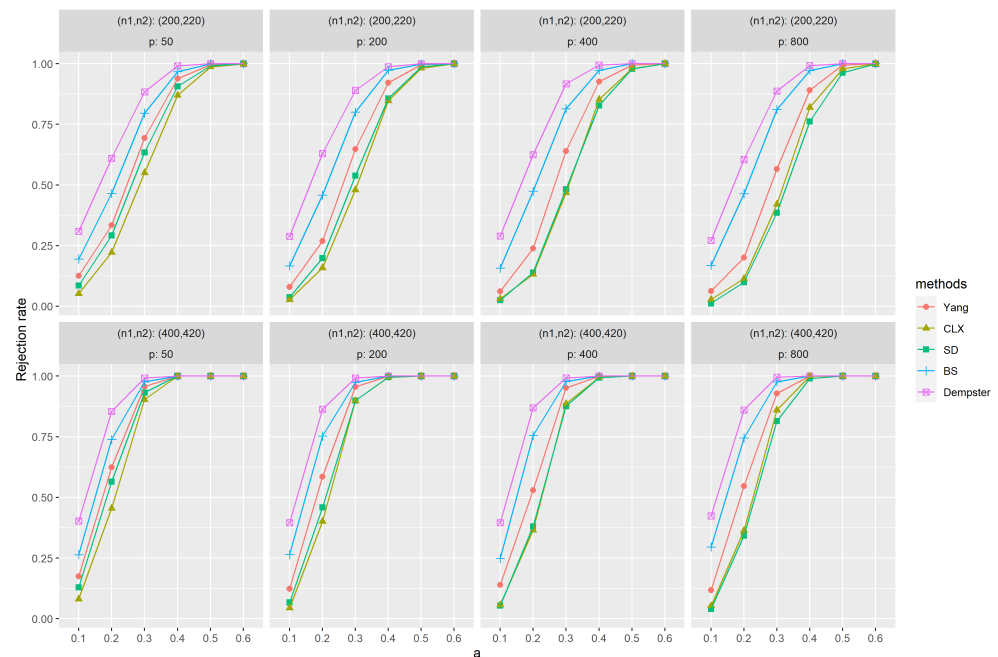

Figure A2. The empirical powers with dense signals were evaluated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ based on (M2), $\mathbf{f}(\cdot) = \mathbf{0}$ and $\rho = 0$, and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M2), $\mathbf{f}(\cdot) = \mathbf{f}_2(\cdot)$ and $\rho = 0$. The parameter a represents the distance between the null and alternative hypotheses. The simulations were replicated 1000 times.

Table A2. The Type I error rates, expressed as percentages, were calculated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M3). The simulations were replicated 1000 times.

(n_1, n_2)	ρ	p	Yang	Dempster	BS	SD	CLX
(200,220)	0	50	5.7	16.8	7.7	3	1.6
		200	4.3	14.9	6.9	0.9	1.6
		400	3.5	14.7	7.7	0.2	1.2
		800	4.2	15.4	6.9	0.2	1.7
	0.1	50	7.9	25.2	13.7	5.5	5.4
		200	6.2	23	12	2.7	5
		400	5.5	23.3	12.5	1.2	4.2
		800	6.9	24	12.9	0.7	5.5
	0.2	50	8.6	33.8	21	10.7	12.8
		200	7.5	32.5	19.7	5.8	13.9
		400	6.9	30.4	20.2	4.4	15
		800	9.3	32.3	20.7	1.7	18.4
(400,420)	0	50	5.4	13.9	6.7	1.7	1.6
		200	5.1	15.5	6.4	1	1.1
		400	5.3	14.1	7.1	0.8	1.3
		800	4	16.2	6.3	0.1	1.3
	0.1	50	6.9	21.3	10.7	4.7	5.4
		200	6.6	23.1	12.5	2.7	4.9
		400	7.3	22	11.4	1.7	5.9
		800	6.2	23.8	12.7	0.6	5.4
	0.2	50	8.2	31.8	18.2	8.6	11
		200	8.2	31	19.6	5.2	13.5
		400	8.7	32.6	18.9	3.9	14.7
		800	7.4	35.2	21.3	1.8	17.3
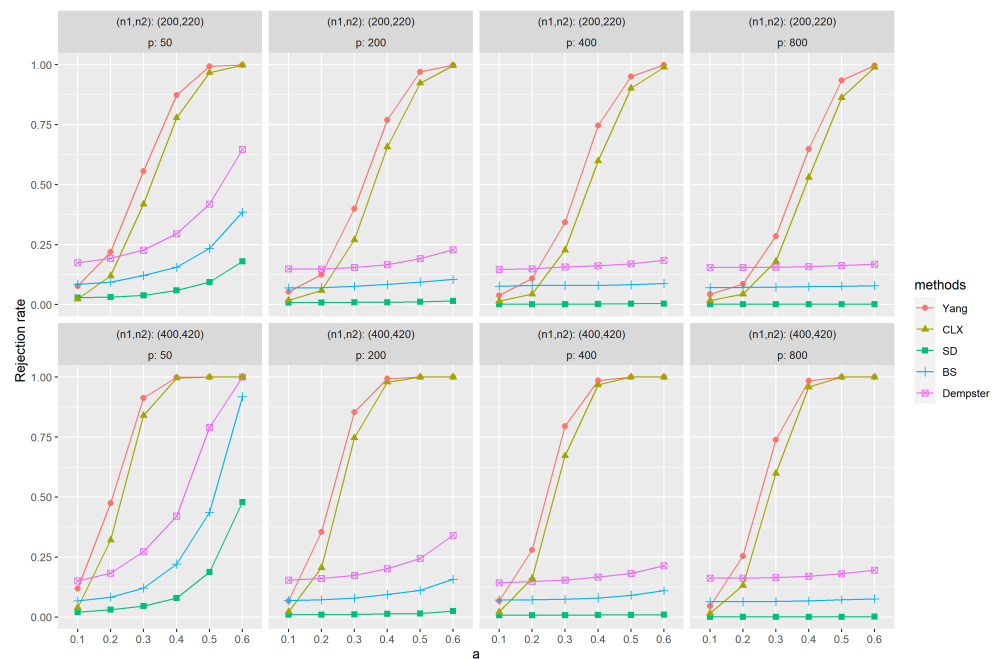
Figure A3. The empirical powers with sparse signals were evaluated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ based on (M3), $\mathbf{f}(\cdot) = \mathbf{0}$ and $\rho = 0$, and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M3), $\mathbf{f}(\cdot) = \mathbf{f}_1(\cdot)$ and $\rho = 0$. The parameter a represents the distance between the null and alternative hypotheses. The simulations were replicated 1000 times.

Figure A4. The empirical powers with dense signals were evaluated by independently generated sequences $\{\mathbf{X}_t\}_{t=1}^{n_1}$ based on (M3), $\mathbf{f}(\cdot) = \mathbf{0}$ and $\rho = 0$, and $\{\mathbf{Y}_t\}_{t=1}^{n_2}$ based on (M3), $\mathbf{f}(\cdot) = \mathbf{f}_2(\cdot)$ and $\rho = 0$. The parameter a represents the distance between the null and alternative hypotheses. The simulations were replicated 1000 times.

References

1. Hotelling, H. The generalization of student's ratio. *Ann. Math. Stat.* **1931**, *2*, 360–378. [CrossRef]
2. Hu, J.; Bai, Z. A review of 20 years of naive tests of significance for high-dimensional mean vectors and covariance matrices. *Sci. China Math.* **2016**, *59*, 2281–2300. [CrossRef]
3. Harrar, S.W.; Kong, X. Recent developments in high-dimensional inference for multivariate data: Parametric, semiparametric and nonparametric approaches. *J. Multivar. Anal.* **2022**, *188*, 104855. [CrossRef]
4. Bai, Z.; Saranadasa, H. Effect of high dimension: By an example of a two sample problem. *Stat. Sin.* **1996**, *6*, 311–329.
5. Dempster, A.P. A high dimensional two sample significance test. *Ann. Math. Stat.* **1958**, *29*, 995–1010. [CrossRef]
6. Srivastava, M.S.; Du, M. A test for the mean vector with fewer observations than the dimension. *J. Multivar. Anal.* **2008**, *99*, 386–402. [CrossRef]
7. Gregory, K.B.; Carroll, R.J.; Baladandayuthapani, V.; Lahiri, S.N. A two-sample test for equality of means in high dimension. *J. Am. Stat. Assoc.* **2015**, *110*, 837–849. [CrossRef] [PubMed]
8. Cai, T.T.; Liu, W.; Xia, Y. Two-sample test of high dimensional means under dependence. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2014**, *76*, 349–372.
9. Chang, J.; Zheng, C.; Zhou, W.X.; Zhou, W. Simulation-Based Hypothesis Testing of High Dimensional Means Under Covariance Heterogeneity. *Biometrics* **2017**, *73*, 1300–1310. [CrossRef]
10. Xu, G.; Lin, L.; Wei, P.; Pan, W. An adaptive two-sample test for high-dimensional means. *Biometrika* **2017**, *103*, 609–624. [CrossRef]
11. Chernozhukov, V.; Chetverikov, D.; Kato, K. *Testing Many Moment Inequalities*; Cemmap working paper, No. CWP42/16; Centre for Microdata Methods and Practice (cemmap): London, UK, 2016.
12. Zhang, D.; Wu, W.B. Gaussian approximation for high dimensional time series. *Ann. Stat.* **2017**, *45*, 1895–1919. [CrossRef]
13. Wu, W.B. Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14150–14154. [CrossRef]
14. Chang, J.; Chen, X.; Wu, M. Central limit theorems for high dimensional dependent data. *Bernoulli* **2024**, *30*, 712–742. [CrossRef]
15. Raič, M. A multivariate berry–esseen theorem with explicit constants. *Bernoulli* **2019**, *25*, 2824–2853.
16. Chernozhukov, V.; Chetverikov, D.; Kato, K.; Koike, Y. Improved central limit theorem and bootstrap approximations in high dimensions. *Ann. Stat.* **2022**, *50*, 2562–2586. [CrossRef]
17. Peligrad, M. Some remarks on coupling of dependent random variables. *Stat. Probab. Lett.* **2002**, *60*, 201–209. [CrossRef]
18. Künsch, H.R. The jackknife and the bootstrap for general stationary observations. *Ann. Stat.* **1989**, *17*, 1217–1241. [CrossRef]
19. Liu, R.Y. Bootstrap procedures under some non-I.I.D. models. *Ann. Stat.* **1988**, *16*, 1696–1708. [CrossRef]
20. Hill, J.B.; Li, T. A global wavelet based bootstrapped test of covariance stationarity. *arXiv* **2022**, arXiv:2210.14086.
21. Fang, X.; Koike, Y. High-dimensional central limit theorems by Stein's method. *Ann. Appl. Probab.* **2021**, *31*, 1660–1686. [CrossRef]

22. Chernozhukov, V.; Chetverikov, D.; Koike, Y. Nearly optimal central limit theorem and bootstrap approximations in high dimensions. *Ann. Appl. Probab.* **2023**, *33*, 2374–2425. [CrossRef]
23. Chang, J.; He, J.; Yang, L.; Yao, Q. Modelling matrix time series via a tensor CP-decomposition. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2023**, *85*, 127–148. [CrossRef]
24. Koike, Y. High-dimensional central limit theorems for homogeneous sums. *J. Theor. Probab.* **2023**, *36*, 1–45. [CrossRef]
25. Hörmann, S.; Kokoszka, P. Weakly dependent functional data. *Ann. Stat.* **2010**, *38*, 1845–1884. [CrossRef]
26. Zhang, X. White noise testing and model diagnostic checking for functional time series. *J. Econ.* **2016**, *194*, 76–95. [CrossRef]
27. Politis, D.N.; Romano, J.P.; Wolf, M. *Subsampling*; Springer Series in Statistics; Springer: Berlin/Heidelberg, Germany, 1999.
28. Chernozhukov, V.; Chetverikov, D.; Kato, K. Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Stat.* **2013**, *41*, 2786–2819. [CrossRef]
29. Zhou, Z. Heteroscedasticity and autocorrelation robust structural change detection. *J. Am. Stat. Assoc.* **2013**, *108*, 726–740. [CrossRef]
30. Rio, E. *Inequalities and Limit Theorems for Weakly Dependent Sequences, 3rd Cycle*; France, 2013; p. 170. Available online: https://cel.hal.science/cel-00867106v2 (accessed on 8 December 2023).
31. Chernozhukov, V.; Chetverikov, D.; Kato, K. Central limit theorems and bootstrap in high dimensions. *Ann. Probab.* **2017**, *45*, 2309–2352. [CrossRef]
32. Merlevède, F.; Peligrad, M.; Rio, E. A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probab. Theory Relat. Fields* **2011**, *151*, 435–474. [CrossRef]
33. Merlevède, F.; Peligrad, M.; Rio, E. Bernstein inequality and moderate deviations under strong mixing conditions. In *High Dimensional Probability V: The Luminy Volume*; Institute of Mathematical Statistics: Waite Hill, OH, USA, 2009; Volume 5, pp. 273–292.