

Robust inference for change points in high dimension

Feiyu Jiang^a, Runmin Wang^{b,*}, Xiaofeng Shao^c

^a Department of Statistics and Data Science, School of Management, Fudan University, China

^b Department of Statistics, Texas A&M University, United States of America

^c Department of Statistics, University of Illinois at Urbana Champaign, United States of America

ARTICLE INFO

Article history:

Received 11 May 2022

Received in revised form 21 September 2022

Accepted 22 September 2022

Available online 30 September 2022

AMS 2020 subject classifications:

primary 62H15

secondary 62G10

Keywords:

Change points

High dimensional data

Segmentation

Self-normalization

Spatial sign

ABSTRACT

This paper proposes a new test for a change point in the mean of high-dimensional data based on the spatial sign and self-normalization. The test is easy to implement with no tuning parameters, robust to heavy-tailedness and theoretically justified with both fixed- n and sequential asymptotics under both null and alternatives, where n is the sample size. We demonstrate that the fixed- n asymptotics provide a better approximation to the finite sample distribution and thus should be preferred in both testing and testing-based estimation. To estimate the number and locations when multiple change-points are present, we propose to combine the p -value under the fixed- n asymptotics with the seeded binary segmentation (SBS) algorithm. Through numerical experiments, we show that the spatial sign based procedures are robust with respect to the heavy-tailedness and strong coordinate-wise dependence, whereas their non-robust counterparts proposed in Wang et al. (2022)[28] appear to under-perform. A real data example is also provided to illustrate the robustness and broad applicability of the proposed test and its corresponding estimation algorithm.

© 2022 Elsevier Inc. All rights reserved.

1. Introduction

Change-point detection has been a popular research topic in statistics, and related literature is vast due to its broad applicability in biosciences, climate sciences, economics and quality control, among many other areas. See [1,2,12,25] for recent reviews. With the development of modern data collection techniques, high-dimensional data has become more common in the foregoing areas, and the associated data analysis has also triggered the advancement of inference methods for change-points in high-dimensional data, see, e.g. [7,11,13,18,27–31]. Among the proposed tests and estimation methods, most of them require quite strong moment conditions (e.g., Gaussian or sub-Gaussian assumption, or sixth moment assumption) and some of them also require weak component-wise dependence assumption. There are only a few exceptions, such as [30], where they used anti-symmetric and nonlinear kernels in a U-statistics framework to achieve robustness. However, the limiting distribution of their test statistic is non-pivotal and their procedure requires bootstrap calibration, which could be computationally demanding. In addition, their test statistic targets the sparse alternative only. As pointed out in [17], the interest in the dense alternative can be well motivated by real data and is often the type of alternative the practitioners want to detect. For example, copy number variations in cancer cells are commonly manifested as change-points occurring at the same positions across many related data sequences corresponding to cancer samples and biologically related individuals; see [8].

* Corresponding author.

E-mail address: runminw@tamu.edu (R. Wang).

In this article, we propose a new test for a change point in the mean of high-dimensional data that works for a broad class of data generating processes. In particular, our test targets the dense alternative, is robust to heavy-tailedness, and can accommodate both weak and strong coordinate-wise dependence. Our test is built on two recent advances in high-dimensional testing: spatial sign based two sample test developed in [5] and U-statistics based change-point test developed in [28]. Spatial sign based tests have been studied in the literature of multivariate data and they are usually used to handle heavy-tailedness, see [20] for a book-length review. However, it was until recently that [5,26] discovered that spatial sign could also help relax the restrictive moment conditions in high dimensional testing problems. [28] advanced the high-dimensional two sample U-statistic pioneered by [6] to the change-point setting by adopting the self-normalization (SN) [22,24]. Their test targets dense alternative, but requires sixth moment assumption and only allows for weak coordinate-wise dependence.

Building on these two recent advances, we shall propose a spatial signed SN-based test for a change point in the mean of high-dimensional data. Our contribution to the literature is threefold. Firstly, we derive the limiting null distribution of our test statistic under the so-called fixed- n asymptotics, where the sample size n is fixed and dimension p grows to infinity. We discovered that the fixed- n asymptotics provide a better approximation to the finite sample distribution when the sample size is small or moderate. We also let n grow to infinity after we derive n -dependent asymptotic distribution, and obtain the limit under the sequential asymptotics [21]. This type of asymptotics seems new to the high-dimensional change-point literature and may be more broadly adopted in change-point testing and other high-dimensional problems. Secondly, our asymptotic theory covers both scenarios, the weak coordinate-wise dependence via ρ mixing, and strong coordinate-wise dependence under the framework of “randomly scaled ρ -mixing sequence” (RSRM) in [5]. The process convergence associated with spatial signed U-process we develop in this paper further facilitates the application of our test under sequential asymptotics where n , in addition to p , also goes to infinity. In particular, we have developed novel theory to establish the process convergence result under the RSRM framework. In general, this requires to show the finite dimensional convergence and asymptotic equicontinuity (tightness). For the tightness, we derive a bound for the eighth moment of the increment of the sample path based on a conditional argument under the sequential asymptotics, which is new to the literature. Using this new technique, we provide the unconditional limiting null distribution of the test statistic for the fixed- n and growing- p case. This is stronger than the results in [5] which is a conditional limiting null distribution. Thirdly, we extend our test to estimate multiple changes by combining the p -value based on the fixed- n asymptotics and the seeded binary segmentation (SBS) [14]. The use of fixed- n asymptotics is especially recommended due to the fact that in these popular generic segmentation algorithms such as WBS [9] and SBS, test statistics over many intervals of small/moderate lengths are calculated and the sequential asymptotics is not accurate in approximating the finite sample distribution, as compared to its fixed- n counterpart. The superiority and robustness of our estimation algorithm is corroborated in a small simulation study.

The rest of the paper is organized as follows. In Section 2, we define the spatial signed SN test. Section 3 studies the asymptotic behavior of the test under both null and local alternatives. Extensions to estimating multiple change-points are elaborated in Section 4. Numerical studies for testing are relegated to Section 5. Section 6 contains a real data example and Section 7 concludes. All proofs with auxiliary lemmas are given in the appendix. Additional simulation results are provided in the online supplementary material. Throughout the paper, we denote $\xrightarrow{P}$ as the convergence in probability, $\xrightarrow{D}$ as the convergence in distribution and $\rightsquigarrow$ as the weak convergence for stochastic processes. The notations $\mathbf{1}_d$ and $\mathbf{0}_d$ are used to represent vectors of dimension d whose entries are all ones and zeros, respectively. For $a, b \in \mathbb{R}$, denote $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$. For a vector $a \in \mathbb{R}^d$, $\|a\|$ denotes its Euclidean norm. For a matrix A , $\|A\|_F$ denotes its Frobenius norm. Denote $a_n \lesssim b_n$ if there exists $M, C > 0$ such that $a_n \leq Cb_n$ for $n > M$, and let $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$. Let $S(X) = X/\|X\| \mathbf{1}(X \neq \mathbf{0})$ denote the spatial sign of a vector X .

2. Test statistics

Let $\{X_i\}_{i=1}^n$ be a sequence of i.i.d $\mathbb{R}^p$ -valued random vectors with mean $\mathbf{0}_p$ and covariance Σ . We assume that the observed data $\{Y_i\}_{i=1}^n$ satisfies $Y_i = \mu_i + X_i$, where $\mu_i \in \mathbb{R}^p$ is the mean at time i . We are interested in the following testing problem:

$$H_0 : \mu_1 = \cdots = \mu_n, \quad \text{v.s.} \quad H_1 : \mu_1 = \cdots = \mu_{k^*} \neq \mu_{k^*+1} = \cdots = \mu_n, \quad \text{for some } 2 \leq k^* \leq n-1. \quad (1)$$

In (1), under the null, the mean vectors are constant over time while under the alternative, there is one change-point at unknown time point k^* .

Building on the two-sample U-statistic of [6] by Chen and Qin, it seems natural to consider the following test statistic,

$$CQ_n := \sup_{k=2, \dots, n-2} D(k; 1, n),$$

where for $1 \leq l \leq k < m \leq n$,

$$D(k; l, m) = \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} (Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4}). \quad (2)$$

However, as pointed out by [28], the limiting distribution of (properly standardized) $D(k; 1, n)$ depends on some unknown nuisance parameter, which relies heavily on the covariance (correlation) structure of Y_i and is typically unknown in practice. One may replace the unknown nuisance parameter with a consistent estimator, and this is indeed adopted in high dimensional one-sample or two-sample testing problems, see, e.g., [5,6]. Unfortunately, in the context of change-point testing, the unknown location k^* makes this method practically unreliable. To this end, [28] applied SN technique in [24] to avoid the consistent estimation of unknown nuisance parameter. SN technique was initially developed in [22,24] in the low dimensional time series setting and its main idea is to use an inconsistent variance estimator (i.e. self-normalizer) which is based on recursive subsample test statistic, so that the limiting distribution is pivotal under the null. See [23] for a recent review. In particular, the SN based test statistic in [28] is defined as

$$T_n := \sup_{k=4, \dots, n-4} \frac{(D(k; 1, n))^2}{W_n(k; 1, n)}, \quad (3)$$

where $D(k; l, m)$ is defined in (2), and

$$W_n(k; l, m) = \frac{1}{n} \sum_{t=l+1}^{k-2} D(t; l, k)^2 + \frac{1}{n} \sum_{t=k+2}^{m-2} D(t; k, m)^2.$$

The test statistic T_n is applied to the level, and its asymptotic validity requires the existence of sixth moments and weak componentwise dependence assumption; see [28]. Motivated by the success of multivariate spatial sign test [19,20], we propose to modify the test statistic (3) to accommodate for heavy-tailedness. Consider the following spatial signed SN test statistic:

$$T_n^{(s)} := \sup_{k=4, \dots, n-4} \frac{(D^{(s)}(k; 1, n))^2}{W_n^{(s)}(k; 1, n)}, \quad (4)$$

where for $1 \leq l \leq k < m \leq n$,

$$D^{(s)}(k; l, m) = \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} S(Y_{j_1} - Y_{j_2})^\top S(Y_{j_3} - Y_{j_4}), \quad (5)$$

$$W_n^{(s)}(k; l, m) = \frac{1}{n} \sum_{t=l+1}^{k-2} D^{(s)}(t; l, k)^2 + \frac{1}{n} \sum_{t=k+2}^{m-2} D^{(s)}(t; k+1, m)^2. \quad (6)$$

Here, the superscript (s) is used to highlight the role of spatial sign plays in constructing the testing statistic. Note that $S(Y_i - Y_j)$ takes value in the unit sphere of $\mathbb{R}^p$, its Euclidean norm thus possesses much higher moments than the non-spatial signed counterpart, making the test statistic (4) robust to heavy-tailedness.

Remark 1. Note if $k = k^*$, then $4^{-1} \binom{k^*}{2}^{-1} \binom{n-k^*}{2}^{-1} D^{(s)}(k; 1, n)$ is an unbiased estimator for $\|\mathbb{E}[S(Y_1 - Y_n)]\|^2$. Hence, our spatial signed test targets at $\|\mathbb{E}[S(Y_1 - Y_n)]\|$. Note $X_i - X_j$ is symmetric about $\mathbf{0}_p$, H_0 holds if and only if $\|\mathbb{E}[S(Y_i - Y_j)]\| = 0$ for all $i \neq j$, see Theorem 2.1 in [20].

3. Theoretical properties

We first introduce the concept of ρ -mixing, see e.g. [3]. Typical ρ -mixing sequences include i.i.d sequences, m -dependent sequences, stationary strong ARMA processes and many Markov chain models.

Definition 1 (ρ -mixing). A sequence of random variables $\xi_i \in \mathbb{R}$, denoted by $(\xi_1, \xi_2, \dots)$ is said to be ρ -mixing if

$$\rho(d) = \sup_{k \geq 1} \sup_{f \in \mathcal{F}_1^k, g \in \mathcal{F}_{d+k}^\infty} |\text{Corr}(f, g)| \rightarrow 0, \quad \text{as } d \rightarrow \infty.$$

where $\text{Corr}(f, g)$ denotes the correlation between f and g , and $\mathcal{F}_i^j$ is the σ -field generated by $(\xi_i, \xi_{i+1}, \dots, \xi_j)$. Here $\rho(\cdot)$ is called the ρ -mixing coefficient of $(\xi_1, \xi_2, \dots)$.

3.1. Assumptions

To analyze the asymptotic behavior of $T_n^{(s)}$, we make the following assumptions.

Assumption 1. $\{X_i\}_{i=1}^n$ are i.i.d copies of ξ , where ξ is formed by the first p observations from a sequence of strictly stationary and ρ -mixing random variables $(\xi_1, \xi_2, \dots)$ such that $E\xi_1 = 0$ and $E\xi_1^2 = \sigma^2$.

Assumption 2. The ρ -mixing coefficients of ξ satisfy $\sum_{k=1}^\infty \rho(2^k) < \infty$.

Assumptions 1 and **2** are imposed in [5] to analyze the behavior of spatial sign based two-sample test statistic for the equality of high dimensional mean. In particular, **Assumption 1** allows us to analyze the behavior of $T_n^{(s)}$ under the fixed- n scenario by letting p go to infinity alone. **Assumption 2** allows weak dependence among the p coordinates of the data, and similar assumptions are also made in, e.g. [28,31]. The strict stationary assumption can be relaxed with additional conditions and the scenario that corresponds to strong coordinate-wise dependence is provided in Section 3.4

3.2. Limiting null

We begin by deriving the limiting distribution of $T_n^{(s)}$ when n is fixed while letting $p \rightarrow \infty$, and then analyze the large sample behavior of the fixed- n limit by letting $n \rightarrow \infty$. The sequential asymptotics is fairly common in statistics and econometrics, see [21].

Theorem 1. Suppose **Assumptions 1** and **2** hold, then under H_0 : (i) for any fixed $n \geq 8$, as $p \rightarrow \infty$, we have

$$T_n^{(s)} \xrightarrow{\mathcal{D}} \mathcal{T}_n, \quad T_n \xrightarrow{\mathcal{D}} \mathcal{T}_n,$$

where

$$\mathcal{T}_n := \sup_{k=4, \dots, n-4} \frac{nG_n^2(\frac{k}{n}, \frac{1}{n}, 1)}{\sum_{t=2}^{k-2} G_n^2(\frac{t}{n}, \frac{1}{n}, \frac{k}{n}) + \sum_{t=k+2}^{n-2} G_n^2(\frac{t}{n}, \frac{k+1}{n}, 1)},$$

with

$$G_n\left(\frac{k}{n}; \frac{l}{n}, \frac{m}{n}\right) = \frac{(m-l)(m-k-1)}{n} Q_n\left(\frac{l}{n}, \frac{k}{n}\right) + \frac{(m-l)(k-l)}{n} Q_n\left(\frac{k+1}{n}, \frac{m}{n}\right) - \frac{(k-l)(m-k-1)}{n} Q_n\left(\frac{l}{n}, \frac{m}{n}\right),$$

and $Q_n(\cdot, \cdot)$ is a centered Gaussian process defined on $[0, 1]^2$ with covariance structure given by:

$$\begin{aligned} \text{Cov}(Q_n(a_1, b_1), Q_n(a_2, b_2)) \\ = n^{-2}([\lfloor nb_1 \rfloor \wedge \lfloor nb_2 \rfloor - \lfloor na_1 \rfloor \vee \lfloor na_2 \rfloor)(\lfloor nb_1 \rfloor \wedge \lfloor nb_2 \rfloor - \lfloor na_1 \rfloor \vee \lfloor na_2 \rfloor + 1)\mathbf{1}(b_1 \wedge b_2 > a_1 \vee a_2). \end{aligned}$$

(ii) Furthermore, if $n \rightarrow \infty$, then

$$\mathcal{T}_n \xrightarrow{\mathcal{D}} \mathcal{T} := \sup_{r \in (0, 1)} \frac{G(r; 0, 1)^2}{\int_0^r G(u; 0, r)^2 du + \int_r^1 G(u; r, 1)^2 du}, \quad (7)$$

with

$$G(r; a, b) = (b-a)(b-r)Q(a, r) + (r-a)(b-a)Q(r, b) - (r-a)(b-r)Q(a, b),$$

and $Q(\cdot, \cdot)$ is a centered Gaussian process defined on $[0, 1]^2$ with covariance structure given by:

$$\text{Cov}(Q(a_1, b_1), Q(a_2, b_2)) = (b_1 \wedge b_2 - a_1 \vee a_2)^2 \mathbf{1}(b_1 \wedge b_2 > a_1 \vee a_2).$$

Theorem 1(i) states that for each fixed $n \geq 8$, when $p \rightarrow \infty$, the limiting distribution $\mathcal{T}_n$ is a functional of Gaussian process, which is pivotal and can be easily simulated, see **Table 1** for tabulated quantiles with $n \in \{10, 20, 30, 40, 50, 100, 200\}$ (based on 50,000 Monte Carlo replications). **Theorem 1(ii)** indicates that $\mathcal{T}_n$ converges in distribution as n diverges, which is indeed supported by **Table 1**. In fact, $\mathcal{T}$ is exactly the same as the limiting null distribution obtained in [28] under the joint asymptotics when both p and n diverge at the same time.

Our spatial signed SN test builds on the test by [5], where an estimator $\hat{\Sigma}$ for the covariance Σ is necessary as indicated by Section 2.1 therein. However, if the sample size n is fixed, their estimator $\hat{\Sigma}$ is only unbiased but not consistent. In contrast, the SN technique adopted in this paper enables us to avoid such estimation, and thus makes the fixed n inference feasible in practice. It is worth noting that the test statistics $T_n^{(s)}$ and T_n share the same limiting null under both fixed- n asymptotics and sequential asymptotics.

Our test statistic is based on the spatial signs and only assumes finite second moment, which is much weaker than the sixth moment in [28] under joint asymptotics of p and n . The fixed- n asymptotics provides a better approximation to the finite sample distribution of $T_n^{(s)}$ and T_n when n is small or moderate. So its corresponding critical value should be preferred than the counterparts derived under the joint asymptotics. Thus, when data is heavy-tailed and data length is short, our test is more appealing.

3.3. Power analysis

Denote $\delta = \mu_n - \mu_1$ as the shift in mean under the alternative, and $\iota^2 = \lim_{p \rightarrow \infty} p^{-1} \|\delta\|^2$ as the limiting average signal. Next, we study the behavior of the test under both fixed ($\iota > 0$) and local alternatives ($\iota = 0$).

We first consider the case when the average signal is non-diminishing.

Assumption 3. (i) $\iota > 0$, (ii) $np \|\Sigma\|_F^{-1} \rightarrow \infty$ as $p \rightarrow \infty$.

Table 1

Simulated 100 γ %th quantiles of $\mathcal{T}_n$ based on 50,000 Monte Carlo replications using sequences of i.i.d. $\mathcal{N}(0, 1)$ variables.

$n \backslash \gamma$	80%	90%	95%	99%	99.5%	99.9%
10	1681.5	3080.0	5167.8	14334.1	20405.9	46201.9
20	719.0	1124.3	1624.1	3026.2	3810.6	5899.5
30	633.7	965.1	1350.5	2403.6	2988.8	4748.0
40	609.7	926.5	1283.0	2292.3	2750.0	4035.7
50	596.2	889.3	1225.0	2187.0	2624.7	3846.5
100	594.5	881.9	1200.3	2066.4	2482.5	3638.7
200	592.1	878.2	1195.3	2049.3	2456.7	3533.4

Here [Assumption 3\(ii\)](#) is quite mild and can be satisfied by many weak dependent sequences such as ARMA sequences.

Theorem 2 (Fixed Alternative). Suppose [Assumptions 1–3](#) hold, then

$$T_n^{(s)} \xrightarrow{\mathcal{P}} \infty, \quad T_n \xrightarrow{\mathcal{P}} \infty$$

[Theorem 2](#) shows that when average signal is non-diminishing, then both $T_n^{(s)}$ and T_n are consistent tests. Next, we analyze $T_n^{(s)}$ under local alternatives when $\iota = 0$.

Assumption 4. (i) $\iota = 0$, (ii) $\delta^\top \Sigma \delta = o(\|\Sigma\|_F^2)$ as $p \rightarrow \infty$.

[Assumption 4](#) regulates the behavior of the shift size, and is used to simplify the theoretical analysis of $T_n^{(s)}$ under local alternatives. Similar assumptions are also made in [5]. Clearly, when Σ is the identity matrix, [Assumption 4\(ii\)](#) automatically holds if $\iota = 0$.

Theorem 3 (Local Alternative). Suppose [Assumptions 1, 2](#) and [4](#) hold. Assume there exists a k^* such that $\mu_i = \mu$, $i \in \{1, \dots, k^*\}$ and $\mu_i = \mu + \delta$, $i \in \{k^* + 1, \dots, n\}$. Then for any fixed n , as $p \rightarrow \infty$,

- (i) if $n\|\Sigma\|_F^{-1}\|\delta\|^2 \rightarrow \infty$, then $T_n^{(s)} \xrightarrow{\mathcal{P}} \infty$ and $T_n \xrightarrow{\mathcal{P}} \infty$;
- (ii) if $n\|\Sigma\|_F^{-1}\|\delta\|^2 \rightarrow 0$, then $T_n^{(s)} \xrightarrow{\mathcal{D}} T_n$ and $T_n \xrightarrow{\mathcal{D}} T_n$;
- (iii) if $n\|\Sigma\|_F^{-1}\|\delta\|^2 \rightarrow c_n \in (0, \infty)$, then $T_n^{(s)} \xrightarrow{\mathcal{D}} \mathcal{T}_n(c_n, \Delta_n)$, and $T_n \xrightarrow{\mathcal{D}} \mathcal{T}_n(c_n, \Delta_n)$, where

$$\mathcal{T}_n(c_n, \Delta_n) = \sup_{k=4, \dots, n-4} \frac{n[\sqrt{2}G_n(\frac{k}{n}; \frac{1}{n}, 1) + c_n \Delta_n(\frac{k}{n}; \frac{1}{n}, 1)]^2}{\sum_{t=2}^{k-2} [\sqrt{2}G_n(\frac{t}{n}; \frac{1}{n}, \frac{k}{n}) + c_n \Delta_n(\frac{t}{n}; \frac{1}{n}, \frac{k}{n})]^2 + \sum_{t=k+2}^{n-2} [\sqrt{2}G_n(\frac{t}{n}; \frac{k+1}{n}, 1) + c_n \Delta_n(\frac{t}{n}; \frac{k+1}{n}, 1)]^2},$$

and

$$\Delta_n\left(\frac{k}{n}; \frac{l}{n}, \frac{m}{n}\right) = \begin{cases} \frac{4\binom{k-l+1}{2}\binom{m-k^*}{2}}{n^4}, & l < k \leq k^* < m; \\ \frac{4\binom{k^*-l+1}{2}\binom{m-k}{2}}{n^4}, & l < k^* < k < m; \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, if $\lim_{n \rightarrow \infty} c_n = c \in (0, \infty)$, then as $n \rightarrow \infty$,

$$\mathcal{T}_n(c_n, \Delta_n) \xrightarrow{\mathcal{D}} \mathcal{T}(c, \Delta) \tag{8}$$

where

$$\mathcal{T}(c, \Delta) := \sup_{r \in [0, 1]} \frac{\{\sqrt{2}G(r; 0, 1) + c\Delta(r, 0, 1)\}^2}{\int_0^r \{\sqrt{2}G(u; 0, r) + c\Delta(u, 0, r)\}^2 du + \int_r^1 \{\sqrt{2}G(u; r, 1) + c\Delta(u, r, 1)\}^2 du},$$

and for $b^* = \lim_{n \rightarrow \infty} (k^*/n)$,

$$\Delta(r, a, b) := \begin{cases} (b^* - a)^2 (b - r)^2, & a < b^* \leq r < b; \\ (r - a)^2 (b - b^*)^2, & a < r < b^* < b; \\ 0, & \text{otherwise.} \end{cases}$$

The above theorem implies that the asymptotic power of $T_n^{(s)}$ and T_n depends on the joint behavior of δ and $\|\Sigma\|_F$, holding n as fixed. If Σ is the identity matrix, then $T_n^{(s)}$ and T_n will exhibit different power behaviors according to whether $\|\delta\|/p^{1/4}$ converges to zero, infinity, or some positive constant. In addition, under the local alternative, the limiting distribution of $T_n^{(s)}$ and T_n under the sequential asymptotics coincides with that in [28] under the joint asymptotics, see Theorem 3.5 therein. In [Fig. 1](#), we plot $\mathcal{T}(c, \Delta)$ at 10%, 50% and 90% quantile levels with b^* fixed at 1/2 and it suggests that $\mathcal{T}(c, \Delta)$ is stochastically increasing with c , which further supports the consistency of both tests.

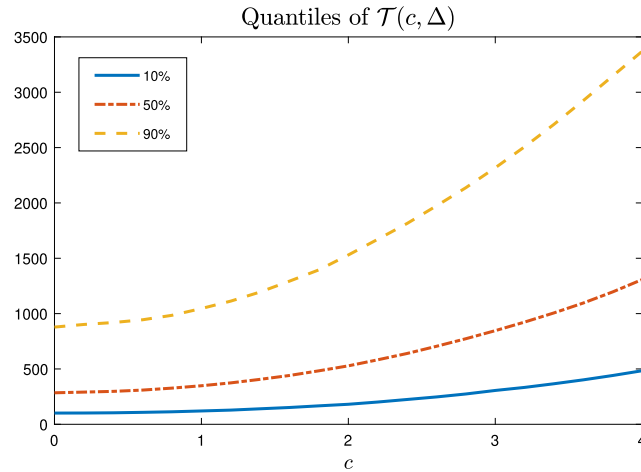


Fig. 1. $T(c, \Delta)$ (see details below (8)) at 10%, 50% and 90% quantile levels with b^* fixed at $1/2$, based on 50,000 Monte Carlo simulations.

Remark 2. Let $\rho^2 = \|\delta\|^2 k^*(n - k^*)/n$ be the signal strength parameter. In [16], the minimax rate for detection of a fully dense change-point under i.i.d. Gaussian assumption is shown to be $\rho^2 \asymp \|\Sigma\|_F \sqrt{\log \log(8n)} \vee \|\Sigma\|_{op} \log \log(8n)$, where $\|\Sigma\|_{op}$ denotes the operator norm of Σ , see Theorem 7 therein. Therefore, in view of Theorem 3, if n is fixed, both T_n and $T_n^{(s)}$ indeed achieve the minimax rate. When $k^*/n = \tau \in (0, 1)$ as n diverges, in [31], it is shown that T_n is rate-optimal (up to a logarithmic factor of n) under the joint asymptotics when both $n, p \rightarrow \infty$.

3.4. Analysis under stronger dependence structure

In this section, we focus on a special class of probability models for high dimensional data termed “randomly scaled ρ -mixing (RSRM)” sequence.

Definition 2 (RSRM, [5]). A sequence $(\eta_1, \eta_2, \dots)$ is a randomly scaled ρ -mixing sequence if there exist a zero mean ρ -mixing sequence $(\xi_1, \xi_2, \dots)$ and an independent positive non-degenerate random variable R such that $\eta_i = \xi_i/R$, $i \in \{1, 2, \dots\}$.

RSRM sequences introduce stronger dependence structure among the coordinates than ρ -mixing sequences, and many models fall into this category, see, e.g. non-Gaussian sequences in [4] and elliptically symmetric models in [26].

Assumption 5. Suppose $Y_i = X_i/R_i + \mu_i$, where $\{X_i\}_{i=1}^n$ satisfies Assumptions 1 and 2, and $\{R_i\}_{i=1}^n$ are i.i.d. copies of a positive random variable R .

Clearly, when R is degenerate (i.e., a positive constant), Assumption 5 reduces to the model assumed in previous sections. However, when R is non-degenerate, Assumption 5 imposes stronger dependence structure on coordinates of Y_i than ρ -mixing sequences, and hence result in additional theoretical difficulties. We refer to [5] for more discussions of RSRM sequences.

Theorem 4. Suppose Assumption 5 holds, then under H_0 ,

(i) let $\mathcal{R}_n = \{R_i\}_{i=1}^n$, for any fixed $n \geq 8$, if $\mathbb{E}(R_i^2) < \infty$ and $\mathbb{E}(R_i^{-2}) < \infty$, as $p \rightarrow \infty$, there exist two random variables $\mathcal{T}_n^{(\mathcal{R}_n, s)}$ and $\mathcal{T}_n^{(\mathcal{R}_n)}$ dependent on $\mathcal{R}_n$ such that,

$$T_n^{(s)} \xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n, s)}, \quad T_n \xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n)}.$$

(ii) Furthermore, if we further assume $\mathbb{E}(R_i^4) < \infty$ and $\mathbb{E}(R_i^{-4}) < \infty$, then as $n \rightarrow \infty$, we have

$$\mathcal{T}_n^{(\mathcal{R}_n, s)} \xrightarrow{\mathcal{D}} \mathcal{T}, \quad \mathcal{T}_n^{(\mathcal{R}_n)} \xrightarrow{\mathcal{D}} \mathcal{T},$$

where $\mathcal{T}$ is defined in (7).

In general, if the sample size n is small and Y_i is generated from an RSRM sequence, the unconditional limiting distributions of $T_n^{(s)}$ and T_n as $p \rightarrow \infty$ are no longer pivotal due to the randomness in R_i . Nevertheless, using the pivotal limiting distribution $\mathcal{T}_n$ in hypothesis testing can still deliver relatively good performance for $T_n^{(s)}$ in both size and power, see Section 5 for numerical evidence. If n is also diverging, the same pivotal limiting distribution as presented in Theorem 1(ii) and in Theorem 3.4 of [28] can still be reached.

Let Σ_Y be the covariance of Y_i (or equivalently X_i/R_i), the next theorem provides with the asymptotic behavior under local alternative for the RSRM model.

Theorem 5. Suppose [Assumptions 4 and 5](#) hold, then under the local alternative such that $n\|\Sigma_Y\|_F^{-1}\|\delta\|^2 \rightarrow c_n \in (0, \infty)$, (i) let $\mathcal{R}_n = \{R_i\}_{i=1}^n$, for any fixed $n \geq 8$, if $\mathbb{E}(R_i^2) < \infty$ and $\mathbb{E}(R_i^{-2}) < \infty$, as $p \rightarrow \infty$, there exist two random variables $\mathcal{T}_n^{(\mathcal{R}_n, s)}(\Delta_n^{(\mathcal{R}_n, s)})$ and $\mathcal{T}_n^{(\mathcal{R}_n)}(\Delta_n)$ dependent on $\mathcal{R}_n$ such that,

$$T_n^{(s)} \xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n, s)}(c_n, \Delta_n^{(\mathcal{R}_n, s)}), \quad T_n \xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n)}(c_n, \Delta_n).$$

(ii) Furthermore, if we assume $\mathbb{E}(R_i^4) < \infty$ and $\mathbb{E}(R_i^{-4}) < \infty$, and $\lim_{n \rightarrow \infty} c_n = c \in (0, \infty)$, then as $n \rightarrow \infty$, we have

$$\mathcal{T}_n^{(\mathcal{R}_n, s)}(c_n, \Delta_n^{(\mathcal{R}_n, s)}) \xrightarrow{\mathcal{D}} \mathcal{T}(Kc, \Delta), \quad \mathcal{T}_n^{(\mathcal{R}_n)}(c_n, \Delta_n^{(\mathcal{R}_n)}) \xrightarrow{\mathcal{D}} \mathcal{T}(c, \Delta),$$

where $\mathcal{T}(c, \Delta)$ is defined in [\(8\)](#), and

$$K = \mathbb{E}^{-1} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \mathbb{E}(R_1^{-2}) \mathbb{E}^2 \left[\frac{R_1 R_2}{\sqrt{R_1^2 + R_2^2}} \right] > 1$$

is a constant.

For the RSRM model, similar to [Theorem 4\(i\)](#), the fixed- n limiting distributions of $T_n^{(s)}$ and T_n are non-pivotal under local alternatives. However, the distribution of $T_n^{(s)}$ under sequential limit is pivotal $\mathcal{T}(Kc, \Delta)$ while that of T_n is $\mathcal{T}(c, \Delta)$. The multiplicative constant $K > 1$ suggests that for the RSRM model, using $T_n^{(s)}$ could be more powerful as $\mathcal{T}(c, \Delta)$ is expected to be monotone in c , see [Fig. 1](#). This finding coincides with [\[5\]](#) where they showed that using spatial sign based U-statistics for testing the equality of two high dimensional means could be more powerful than the conventional mean-based ones in [\[6\]](#). Thus, when strong coordinate-wise dependence is exhibited in the data, $T_n^{(s)}$ is more preferable.

4. Multiple change-point estimation

In real applications, in addition to change-point testing, another important task is to estimate the number and locations of these change-points. In this section, we assume there are $m \geq 1$ change-points and are denoted by $\mathbf{k} = (k_1, \dots, k_m) \subset \{1, \dots, n\}$. A commonly used algorithm for many practitioners would be binary segmentation (BS), where the data segments are recursively split at the maximal points of the test statistics until the null of no change-points is not rejected for each segment. However, as criticized by many researchers, BS tends to miss potential change-points when non-monotonic change patterns are exhibited. Hence, many algorithms have been proposed to overcome this drawback. Among them, wild binary segmentation (WBS) by [\[9\]](#) and its variants have become increasingly popular because of their easy-to-implement procedures. The main idea of WBS is to perform BS on randomly generated sub-intervals so that some sub-intervals can localize at most one change-point (with high probability). As pointed out by [\[14\]](#), WBS relies on randomly generated sub-intervals and different researchers may obtain different estimates. Hence, [\[14\]](#) propose seeded binary segmentation (SBS) algorithm based on deterministic construction of these sub-intervals with relatively cheaper computational costs so that results are replicable. To this end, we combine the spatial signed SN test with SBS to achieve the task of multiple change-point estimation, and we call it SBS-SN^(s). We first introduce the concept of seeded sub-intervals.

Definition 3 (Seeded Sub-Intervals, [\[14\]](#)). Let $\alpha \in [1/2, 1)$ denote a given decay parameter. For $1 \leq k \leq \lfloor \log_{1/\alpha}(n) \rfloor$ (i.e. logarithm with base $1/\alpha$) define the k th layer as the collection of n_k intervals of initial length l_k that are evenly shifted by the deterministic shift s_k as follows:

$$\mathcal{I}_k = \bigcup_{i=1}^{n_k} \{(\lfloor (i-1)s_k \rfloor, \lceil (i-1)s_k + l_k \rceil)\}$$

where $n_k = 2 \lceil (1/\alpha)^{k-1} \rceil - 1$, $l_k = 10 \lceil n\alpha^{k-1}/10 \rceil$ and $s_k = (n - l_k) / (n_k - 1)$. The overall collection of seeded intervals is

$$\mathcal{I}_\alpha(n) = \bigcup_{k=1}^{\lceil \log_{1/\alpha}(n) \rceil} \mathcal{I}_k.$$

Let $\alpha \in [1/2, 1)$ be a decay parameter, denote $\mathcal{I}_\alpha(n)$ as the set of seeded intervals based on [Definition 3](#). For each sub-interval $(a, b) \in \mathcal{I}_\alpha(n)$, we calculate the spatial signed SN test

$$T^{(s)}(a, b) = \max_{k \in \{a+3, \dots, b-4\}} \frac{(D^{(s)}(k; a, b))^2}{W_{b-a+1}^{(s)}(k; a, b)}, \quad b - a \geq 7,$$

where $D^{(s)}(k; a, b)$ and $W_{b-a+1}^{(s)}(k; a, b)$ are defined in [\(5\)](#) and [\(6\)](#). We obtain the p -value of the sub-interval test statistic $T^{(s)}(a, b)$ based on the fixed- n asymptotic distribution $\mathcal{T}_{b-a+1}$. SBS-SN^(s) then finds the smallest p -value evaluated at

all sub-intervals and compare it with a predetermined threshold level ζ_p . If the smallest p -value is also smaller than ζ_p , denote the corresponding sub-interval where the smallest p -value is achieved as (a^*, b^*) and estimate the change-point by $\hat{k} = \arg \max_{k \in \{a^*+3, \dots, b^*-4\}} \{(D^{(s)}(k; a^*, b^*))^2 / W_{b^*-a^*+1}^{(s)}(k; a^*, b^*)\}$. Once a change-point is identified, SBS-SN^(s) then divides the data sample into two subsamples accordingly and apply the same procedure to each of them. The process is implemented recursively until no change-point is detected. Details are provided in Algorithm 1.

Algorithm 1: SBS-SN^(s)

Input: Data $\{Y_t\}_{t=1}^n$, threshold p -value $\zeta_p \in (0, 1)$, SBS intervals $\mathcal{I}_\alpha(n)$.
Output: Estimated number of change-points $\hat{m}$ and estimated change-points set $\hat{k}$
Initialization: SBS-SN^(s) (1, n , ζ_p)
Procedure: SBS-SN^(s) (a , b , ζ_p)

```

1 if  $b - a + 1 < 8$  then
2   | Stop
3 else
4   |  $\mathcal{M}_{(a,b)} := \{i : [a_i, b_i] \in \mathcal{I}_\alpha(n), [a_i, b_i] \subset [a, b], b_i - a_i + 1 \geq 8\}$ ;
5   | for each  $i \in \mathcal{M}_{(a,b)}$ , find the  $p$ -value  $p_i$  of  $T^{(s)}(a_i, b_i)$  based on  $\mathcal{T}_{b_i-a_i+1}$ ;
6   |  $i^* = \arg \min_{i \in \mathcal{M}_{(a,b)}} p_i$ ;
7   | if  $p_{i^*} < \zeta_p$  then
8     |  $k^* = \arg \max_{k \in \{a_{i^*}+3, \dots, b_{i^*}-4\}} \frac{(D^{(s)}(k; a_{i^*}, b_{i^*}))^2}{W_{b_{i^*}-a_{i^*}+1}^{(s)}(k; a_{i^*}, b_{i^*})}$ ;
9     |  $\hat{k} = \hat{k} \cup k^*$ ,  $\hat{m} = \hat{m} + 1$ ;
10    | SBS-SN(s) ( $a$ ,  $k^*$ ,  $\zeta_p$ );
11    | SBS-SN(s) ( $k^* + 1$ ,  $b$ ,  $\zeta_p$ );
12  | else
13  | | Stop
14  | end
15 end

```

Our SBS-SN^(s) algorithm differs from WBS-SN algorithm in [28,31] in two aspects. First, WBS-SN is built on WBS, which relies on randomly generated intervals while SBS relies on deterministic intervals. As documented in [14], WBS is computationally more demanding than SBS. Second, the threshold used in WBS-SN is universal for each sub-interval, depends on the sample size n and dimension p and needs to be simulated via extensive Monte Carlo simulations. Generally speaking, WBS-SN requires simulating a new threshold each time for a new dataset. By contrast, our estimation procedure is based on p -values under the fixed- n asymptotics, which takes into account the interval length $b - a + 1$ for each sub-interval (a, b) . When implementing either WBS or SBS, inevitably, there will be intervals of small lengths. Hence, the universal threshold may not be suitable as it does not take into account the effect of different interval lengths. In order to alleviate the problem of multiple testing, we may set a small threshold number for ζ_p , such as 0.001 or 0.005. Furthermore, the WBS-SN requires to specify a minimal interval length which can affect the finite sample performance. In this work, when generating seed sub-intervals as in Definition 3, the lengths of these intervals are set as integer values times 10 to reduce the computational cost for simulating fixed- n asymptotic distribution $\mathcal{T}_n$. Therefore, we only require the knowledge of $\{\mathcal{T}_n\}_{n=10,20,\dots}$ for SBS-SN^(s) to work, which can be simulated once for good and do not change with a new dataset.

5. Numerical experiments

This section assesses the performance of $T_n^{(s)}$ with respect to various covariance structure of the data. In the Supplementary Material, we provide more simulation results including comparisons with [13,18,30], and performances of our multiple change-point estimation algorithm SBS-SN^(s).

Consider the following data generating process with $p = 100$ and $n \in \{10, 20, 50, 100, 200\}$:

$$Y_i = \delta \mathbf{1}(i > 0.5n) + X_i,$$

where δ represents the mean shift vector, and $\{X_i\}_{i=1}^n$ are i.i.d copies of X based on the following specifications:

- (i) $X \sim \mathcal{N}(\mathbf{0}, I_p)$;
- (ii) $X \sim t_5(I_p)$;
- (iii) $X \sim t_3(I_p)$;
- (iv) $X = (X^{(1)}, \dots, X^{(p)})^\top$, $X^{(t)} = \rho X^{(t-1)} + \epsilon_t$, $t \in \{1, \dots, p\}$, where $\epsilon_t \sim \mathcal{N}(0, 1)/2$ are i.i.d random variables;
- (v) $X = (X^{(1)}, \dots, X^{(p)})^\top$, $X^{(t)} = \rho X^{(t-1)} + \epsilon_t$, $t \in \{1, \dots, p\}$, where $\epsilon_t \sim t_5/2$ are i.i.d random variables;
- (vi) $X = R/U$, $R = (R^{(1)}, \dots, R^{(p)})^\top$, $R^{(t)} = \rho R^{(t-1)} + \epsilon_t$, $t \in \{1, \dots, p\}$, where $\epsilon_t \sim \mathcal{N}(0, 1)/2$ are i.i.d random variables, and $U \sim \text{Exp}(1)$ is independently generated;

Table 2Size and power comparison of T_n and $T_n^{(s)}$ for Case (i)–(vii) at 5% significance level based on 1000 replications.

Case	Test	Limit	Under H_0					Under H_a^1					Under H_a^2							
			n	10	20	50	100	200	n	10	20	50	100	200	n	10	20	50	100	200
(i)	T_n	$\mathcal{T}_n$	5.6	4.8	6.9	4.0	6.4	6.3	6.5	15.2	34.7	77.1	7.3	11.0	34.1	78.0	99.9			
		$\mathcal{T}$	27.4	9.0	7.4	4.1	6.7	29.9	11.2	16.5	35.2	77.8	33.7	18.1	34.9	78.1	99.9			
	$T_n^{(s)}$	$\mathcal{T}_n$	5.5	4.8	6.2	4.3	6.6	6.2	5.9	15.0	33.4	76.7	7.2	10.4	33.3	77.7	99.8			
		$\mathcal{T}$	28.5	8.7	6.8	4.4	7.1	29.8	10.8	15.7	34.6	77.6	33.5	17.3	34.6	78.6	99.8			
(ii)	T_n	$\mathcal{T}_n$	6.9	6.4	6.8	4.3	6.0	7.2	7.2	11.7	18.5	41.4	8.0	8.8	22.0	47.2	87.3			
		$\mathcal{T}$	31.8	12.6	7.6	4.3	6.2	31.8	12.4	12.8	19.0	42.5	33.9	15.3	22.9	47.5	87.4			
	$T_n^{(s)}$	$\mathcal{T}_n$	5.3	5.3	6.2	4.1	5.6	5.7	5.5	11.8	26.1	59.6	6.4	8.1	26.7	62.9	96.8			
		$\mathcal{T}$	28.2	9.7	6.7	4.2	5.7	28.5	10.0	12.6	27.0	60.4	30.8	14.4	28.0	63.2	96.8			
(iii)	T_n	$\mathcal{T}_n$	9.0	9.5	9.2	6.7	7.9	10.0	10.4	11.8	14.2	25.7	9.8	12.3	17.9	27.9	57.5			
		$\mathcal{T}$	35.8	16.1	9.6	6.9	8.5	35.6	16.1	12.6	14.7	26.0	36.8	18.7	18.8	28.7	58.2			
	$T_n^{(s)}$	$\mathcal{T}_n$	5.6	5.0	6.4	4.8	6.4	5.7	4.9	10.5	21.2	50.2	6.2	6.9	21.9	53.0	93.4			
		$\mathcal{T}$	27.5	9.6	7.0	4.9	6.8	29.2	9.3	11.4	21.7	50.8	28.6	12.8	23.0	53.9	93.7			
(iv)	T_n	$\mathcal{T}_n$	5.9	4.8	6.1	6.8	5.4	6.5	8.6	23.5	46.4	78.8	9.3	14.7	43.2	83.8	99.6			
		$\mathcal{T}$	28.1	9.2	6.7	6.9	5.7	30.9	13.8	24.6	47.4	79.1	33.9	20.1	44.4	84.0	99.6			
	$T_n^{(s)}$	$\mathcal{T}_n$	4.8	3.9	6.0	6.3	5.4	5.5	7.1	22.5	44.2	77.9	6.9	13.0	41.5	84.1	99.8			
		$\mathcal{T}$	27.6	8.2	6.5	6.6	5.4	30.6	12.2	23.5	45.3	78.1	33.1	18.8	43.0	84.6	99.8			
(v)	T_n	$\mathcal{T}_n$	7.0	7.6	6.0	6.8	6.1	8.6	11.1	17.5	30.1	54.2	9.4	11.3	26.7	56.7	94.0			
		$\mathcal{T}$	33.5	12.9	6.6	7.2	6.1	33.6	16.9	17.9	30.4	54.6	34.4	18.3	27.9	57.1	94.2			
	$T_n^{(s)}$	$\mathcal{T}_n$	5.3	4.4	5.0	6.9	5.2	6.1	7.5	18.5	37.1	65.2	5.6	9.4	35.2	73.7	98.7			
		$\mathcal{T}$	29.3	8.5	5.3	7.5	5.5	30.5	11.8	19.0	37.7	65.6	30.6	14.1	35.8	74.2	98.8			
(vi)	T_n	$\mathcal{T}_n$	34.7	39.7	39.2	34.6	33.6	34.6	40.7	39.4	35.6	34.2	35.0	39.6	40.1	34.3	33.8			
		$\mathcal{T}$	60.2	46.7	40.5	34.9	34.1	62.5	47.5	40.3	36.1	34.8	60.6	46.9	41.0	34.4	34.1			
	$T_n^{(s)}$	$\mathcal{T}_n$	5.0	4.2	5.3	5.9	5.9	6.0	4.8	11.3	20.1	35.3	5.6	7.1	16.8	37.2	73.5			
		$\mathcal{T}$	27.9	8.6	5.7	6.2	6.1	28.1	10.0	12.0	20.3	35.4	28.2	11.9	17.6	38.0	74.0			
(vii)	T_n	$\mathcal{T}_n$	33.7	40.6	37.9	36.5	36.6	34.3	40.3	37.9	37.0	36.9	33.5	40.6	38.3	36.9	36.8			
		$\mathcal{T}$	61.9	47.3	38.6	37.2	36.8	62.2	46.5	39.1	37.4	37.1	61.5	47.7	39.8	37.7	36.9			
	$T_n^{(s)}$	$\mathcal{T}_n$	4.3	4.4	5.2	6.4	6.0	5.1	6.2	9.5	17.5	32.9	5.1	5.8	14.1	28.5	62.5			
		$\mathcal{T}$	30.2	8.4	5.5	6.7	6.5	30.6	10.1	10.2	17.7	33.5	30.4	9.2	15.3	29.1	63.0			

(vii) $X = R/U$, $R = (R^{(1)}, \dots, R^{(p)})^\top$, $R^{(t)} = \rho R^{(t-1)} + \epsilon_t$, $t \in \{1, \dots, p\}$, where $\epsilon_t \sim t_5/2$ are i.i.d random variables, and $U \sim \text{Exp}(1)$ is independently generated;

where $t_\nu(I_p)$ is the multivariate t distribution with degree of freedom ν and covariance I_p ; $\text{Exp}(1)$ is the exponential distribution with mean 1.

Case (i) assumes that coordinates of X are independent and light-tailed; Cases (ii) and (iii) consider the scenario of heavy-tailedness of X ; Cases (iv) and (v) assume the coordinates of X are consecutive random observations from a stationary AR(1) model with autoregressive coefficient $\rho = 0.7$; and Cases (vi) and (vii) assume the coordinates of X are generated from an RSRM with $\rho = 0.7$.

Table 2 shows the empirical rejection rate of T_n and $T_n^{(s)}$ in percentage based on 1000 replications under the null with $H_0 : \delta = \mathbf{0}$; dense alternative $H_a^1 : \delta = 1/\sqrt{p}\mathbf{1}_p$; and sparse alternative $H_a^2 : \delta = (\mathbf{1}_2^\top, \mathbf{0}_{p-2}^\top)^\top$. We compare the approximation using the limiting null distribution of fixed- n asymptotics $\mathcal{T}_n$ and sequential asymptotics $\mathcal{T}$ at 5% level.

We summarize the findings of Table 2 as follows: (1) both T_n and $T_n^{(s)}$ suffer from severe size distortion using sequential asymptotics $\mathcal{T}$ if n is small (i.e., $n \in \{10, 20, 50\}$); (2) both fixed- n asymptotics $\mathcal{T}_n$ and large- n asymptotics $\mathcal{T}$ work well for T_n and $T_n^{(s)}$ when n is large under weak dependence in coordinates (cases (i)–(v)); (3) T_n and $T_n^{(s)}$ are both accurate in size and comparable in power performance when X_i 's are light-tailed (cases (i),(ii), (iv) and (v)) if appropriate limiting distributions are used; (4) T_n is slightly oversized compared with $T_n^{(s)}$ under heavy-tailed distributions (case (iii)); (5) when strong dependence is exhibited in coordinates (cases (vi) and (vii)), $(T_n^{(s)}, \mathcal{T}_n)$ still works for small n while other combinations of tests and asymptotics generally fail; (6) increasing the data length n enhances power under all settings while increasing dependence in coordinates generally reduces power. Overall, the spatial signed SN test using fixed- n asymptotic critical value outperforms (or is comparable with) all other tests and should be preferred due to its robustness and size accuracy.

6. Data application

In this section, we analyze the genomic micro-array (ACGH) dataset for 43 individuals with bladder tumor. The ACGH data contains log intensity ratios of these individuals measured at 2215 different loci on their genome, and copy number variations in the loci can be viewed as the change-point in the genome. Hence change-point estimation could be helpful in determining the abnormality regions, as analyzed by [27,31]. The data is denoted by $\{Y_i\}_{i=1}^{2215}$.

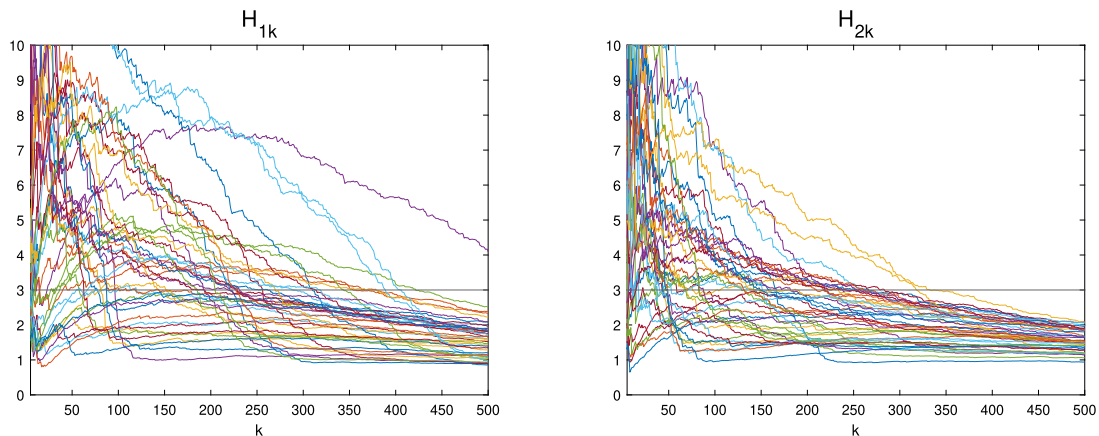


Fig. 2. Hill's estimator (9) for 43 individuals based on the ascending order statistics $Y_{(i),j}$ of the j th individual, $j \in \{1, \dots, 43\}$, $i \in \{1, \dots, 2215\}$.

To illustrate the necessity of robust estimation method proposed in this paper, we use Hill's estimator to estimate the tail index of a sequence, see [10]. Specifically, let $Y_{(i),j}$ be the ascending order statistics of the j th individual (coordinate) across 2215 observations. For $j \in \{1, 2, \dots, 43\}$, we give the left-tail and right-tail Hill estimators respectively by

$$H_{1k,j} = \left\{ \frac{1}{k} \sum_{i=1}^k \log \left(\frac{Y_{(i),j}}{Y_{(k+1),j}} \right) \right\}^{-1}, \quad H_{2k,j} = \left\{ \frac{1}{k} \sum_{i=1}^k \log \left(\frac{Y_{(n-i+1),j}}{Y_{(n-k),j}} \right) \right\}^{-1}, \quad (9)$$

and they are plotted in Fig. 2. From the plot, we see that most of the right-tail and the left-tail indices are below 3, suggesting the data is very likely heavy-tailed.

We take the first 200 loci for our SBS-SN^(s) change-point estimation following the practice in [31], where the decay rate for generation of seeded interval in SBS is $2^{-1/4}$. We also compare the results obtained for Adaptive WBS-SN in [31] and 20 most significant points detected by INSPECT in [27]. For this dataset, INSPECT is more like a screening method as it delivers a total of 67 change-points. In contrast to Adaptive WBS-SN and INSPECT where the thresholds for change-point estimation are simulated, the threshold used in SBS-SN^(s) can be pre-specified, and it reflects a researcher's confidence in detecting the change-points. We set the p -value threshold ζ_p as 0.001, 0.005 and 0.01 and the results are as follows:

Adaptive WBS-SN	15, 32, 38, 44, 59, 74, 91, 97, 102, 116, 134, 158, 173, 186, 191
INSPECT	15, 26, 28, 33, 36, 40, 56, 73, 91, 97, 102, 119, 131, 134, 135, 146, 155, 174, 180, 191
SBS-SN ^(s) , $\zeta_p = 0.001$	30, 41, 72, 89, 130, 136, 174
SBS-SN ^(s) , $\zeta_p = 0.005$	30, 41, 56, 72, 89, 97, 116, 130, 136, 155, 174, 191
SBS-SN ^(s) , $\zeta_p = 0.01$	30, 41, 56, 72, 89, 97, 111, 116, 130, 136, 155, 174, 191

As we see, increasing the p -value threshold ζ_p leads to more estimated change-points, and the set of estimated change-points by using larger ζ_p contain those by smaller ζ_p as subsets. In addition, increasing ζ_p from 0.005 to 0.01 only brings in one more estimated change-point, suggesting $\zeta_p = 0.005$ may be a reasonable choice for the ACGH dataset.

All of our detected change-points at $\zeta_p = 0.005$ are also detected by INSPECT, i.e., 30(28), 41(40), 56, 72(73), 89(91), 97, 116, 130(131), 136 (134,135), 155, 174, 191. Although most of these points also coincide with Adaptive WBS-SN, there are non-overlapping ones. For example, 41, 56, 130 in SBS-SN^(s) seem to be missed in Adaptive WBS-SN while 102 is missed by our SBS-SN^(s) as it is detected by both Adaptive WBS-SN and INSPECT. These results are not really in conflict as Adaptive WBS-SN targets both sparse and dense alternatives, whereas our procedure aims to detect dense change with robustness properties.

7. Conclusion

In this paper, we propose a new method for testing and estimation of change-points in high dimensional independent data. Our test statistic builds on two recent advances in high-dimensional testing problem: spatial sign used in two-sample testing in [5] and self-normalized U-statistics in [28], and inherits many advantages therein such as robustness to heavy-tailedness and tuning-free. The test is theoretically justified under both fixed- n asymptotics and sequential asymptotics, and under both null and alternatives. When data exhibits stronger dependence in coordinates, we further enhance the analysis by focusing on RSRM models, and discover that using spatial sign leads to power improvement compared with mean based tests in [28]. As for multiple change-point estimation, we propose to combine p -values under the fixed- n

asymptotics with the SBS algorithm. Numerical simulations demonstrate that our fixed- n asymptotics for spatial sign based test provides a better approximation to the finite sample distribution, and the estimation algorithm outperforms the mean-based ones when data is heavy-tailed and when coordinates are strongly dependent.

To conclude, we mention a few interesting topics for future research. Our method builds on spatial sign and targets dense signals by constructing unbiased estimators for $\|\mathbb{E}S(Y_1 - Y_n)\|$. As pointed out by [18], many real data exhibit both sparse and dense changes, and it would be interesting to combine with the adaptive SN based test in [31] to achieve both robustness and adaptiveness. In addition, the independence assumption imposed in this paper may limit its applicability to high dimensional time series where temporal dependence cannot be neglected. Methodologically, we believe the time series extension based on our proposed test and a trimming technique should be possible, following the non-spatial signed time series extension considered in [28]. It would also be interesting to develop robust methods for detecting change-points in other quantities beyond mean, such as quantiles, covariance matrices and parameter vectors in high dimensional linear models.

CRedit authorship contribution statement

Feiyu Jiang: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Runmin Wang:** Formal analysis, Writing – original draft, Writing – review & editing. **Xiaofeng Shao:** Conceptualization, Supervision, Project administration, Writing – original draft, Writing – review & editing.

Acknowledgments

We thank Bin Liu and Mengjia Yu for providing the code in [18,30], respectively. We are also grateful to Editor-in-Chief Professor von Rosen, Associate Editor and three anonymous referees for their insightful comments and suggestions, which led to substantial improvements. Jiang is supported by Shanghai Sailing Program No. 22YF1402400, NSFC 12201124 (China); Wang is supported in part by NSF DMS-2210007 (United States of America); Shao is supported in part by National Science Foundation grants NSF DMS-1807032, NSF DMS-2014018 (United States of America).

Appendix A. Proofs of theorems

In what follows, let $x_{i,k}$ denote the k th coordinate of a vector x_i .

Proof of Theorem 1. First, we have that

$$\|Y_i - Y_j\|^2 = \sum_{\ell=1}^p (X_{i,\ell} - X_{j,\ell})^2 + 2(\mu_j - \mu_i)^\top (X_j - X_i) + \|\mu_i - \mu_j\|^2. \quad (10)$$

(i) Under H_0 , by Theorem 8.2.2 in [15], as $p \rightarrow \infty$, we have almost surely,

$$\frac{1}{p} \|Y_i - Y_j\|^2 = 2\sigma^2. \quad (11)$$

Then, for any fixed k, l, m , we have that

$$\begin{aligned} D^{(s)}(k; l, m) &= \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4})}{2p\sigma^2} \\ &\quad + \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4})}{2p\sigma^2} \left\{ \frac{2p\sigma^2}{\|Y_{j_1} - Y_{j_2}\| \|Y_{j_3} - Y_{j_4}\|} - 1 \right\} \\ &=: (2p\sigma^2)^{-1} [D_1(k; l, m) + D_2(k; l, m)], \end{aligned} \quad (12)$$

where clearly $D_1(k; l, m) = D(k; l, m)$, and

$$D_2(k; l, m) = \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} (Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4}) \left\{ \frac{2p\sigma^2}{\|Y_{j_1} - Y_{j_2}\| \|Y_{j_3} - Y_{j_4}\|} - 1 \right\}.$$

Then, Theorem 4.0.1 in [15] implies that

$$\frac{\Gamma^{-1/2}(k; l, m) \{D_1(k; l, m)\}}{(m-k)(m-k-1)(k-l+1)(k-l)} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

where $\Gamma(k; l, m) = \frac{2[(m-l)(m-l-1)]}{(m-k)(m-k-1)(k-l+1)(k-l)} \text{tr}(\Sigma^2)$, or equivalently

$$\frac{1}{n^3 \|\Sigma\|_F} \left\{ D_1(k; l, m) \right\} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{16 \binom{m-1}{2} \binom{k-l+1}{2} \binom{m-k}{2}}{n^6}\right). \quad (13)$$

Next, since we view n as fixed, then for all $j_1 \neq j_3, j_3 \neq j_4$, by Theorem 4.0.1 in [15], it follows that $\|\Sigma\|_F^{-1}(Y_{j_1} - Y_{j_2})^\top(Y_{j_3} - Y_{j_4}) = O_p(1)$. In addition, in view of (11) we have $\frac{2p\sigma^2}{\|Y_{j_1} - Y_{j_2}\| \|Y_{j_3} - Y_{j_4}\|} - 1 = o_p(1)$, and this implies that $n^{-3} \|\Sigma\|_F^{-1} D_2(k; l, m) = o_p(1)$.

Hence, combined with (13), we have

$$\begin{aligned} T_n^{(s)} &= \sup_{k=4, \dots, n-4} \frac{\left(2p\sigma^2 n^{-3} \|\Sigma\|_F^{-1} D^{(s)}(k; 1, n)\right)^2}{4p\sigma^4 n^{-6} \|\Sigma\|_F^{-2} W_n^{(s)}(k; 1, n)} = \sup_{k=4, \dots, n-4} \frac{n^{-6} \|\Sigma\|_F^{-2} [D_1(k; 1, n) + D_2(k; 1, n)]^2}{n^{-6} \|\Sigma\|_F^{-2} W_n(k; 1, n)} + o_p(1) \\ &= T_n + o_p(1), \end{aligned} \quad (14)$$

where the last equality holds since $n^{-3} \|\Sigma\|_F^{-1} D_2(k; l, m) = o_p(1)$ for each triplet (k, l, m) .

For $0 \leq k < m \leq n$, we let $Z(k, m) = \sum_{i=k+1}^m \sum_{j=k}^{i-1} X_i^\top X_j$, then it follows that

$$\begin{aligned} D(k; l, m) &= 2(m-k)(m-k-1)Z(l, k) + 2(k-l+1)(k-l)Z(k+1, m) \\ &\quad - 2(k-l)(m-k-1)[Z(l, m) - Z(l, k) - Z(k+1, m)]. \end{aligned} \quad (15)$$

Then, by Lemma 1, and continuous mapping theorem, we have

$$T_n \xrightarrow{\mathcal{D}} \sup_{k=4, \dots, n-4} \frac{nG_n^2(\frac{k}{n}; \frac{1}{n}, 1)}{\sum_{t=2}^{k-1} G_n^2(\frac{t}{n}; \frac{1}{n}, \frac{k}{n}) + \sum_{t=k+2}^{n-2} G_n^2(\frac{t}{n}; \frac{k+1}{n}, 1)}.$$

(ii) The proof is a simplified version of the proof of Theorem 4(ii), hence omitted here. $\square$

Proof of Theorem 2. Clearly,

$$T_n^{(s)} = \sup_{k=4, \dots, n-4} \frac{(D^{(s)}(k; 1, n))^2}{W_n^{(s)}(k; 1, n)} \geq \frac{(D^{(s)}(k^*; 1, n))^2}{W_n^{(s)}(k^*; 1, n)},$$

and

$$T_n = \sup_{k=4, \dots, n-4} \frac{(D(k; 1, n))^2}{W_n(k; 1, n)} \geq \frac{(D(k^*; 1, n))^2}{W_n(k^*; 1, n)},$$

Note that $W_n^{(s)}(k; 1, n) = \frac{1}{n} \sum_{t=2}^{k-2} D^{(s)}(t; 1, k^*)^2 + \frac{1}{n} \sum_{t=k^*+2}^{n-2} D^{(s)}(t; k^*+1, n)^2$. The construction of $D^{(s)}(t; 1, k^*)^2$ (or $D^{(s)}(t; k^*+1, n)^2$) only uses sample before (or after) the change point, so the change point has no influence on this part. The proof of Theorem 1 indicates that $4p^2 n^{-6} \|\Sigma\|_F^{-2} W_n^{(s)}(k; 1, n) = O_p(1)$ and similarly $4n^{-6} \|\Sigma\|_F^{-2} W_n(k; 1, n) = O_p(1)$. Hence, it suffices to show $pn^{-3} \|\Sigma\|_F^{-1} D^{(s)}(k^*; 1, n) \xrightarrow{\mathcal{P}} \infty$ and $n^{-3} \|\Sigma\|_F^{-1} D(k^*; 1, n) \xrightarrow{\mathcal{P}} \infty$.

Denote δ_i as the i th element of δ . By (10), for $1 \leq j_1 \neq j_3 \leq k^*$ and $k^*+1 \leq j_2 \neq j_4 \leq n$,

$$\begin{aligned} p^{-1} \|Y_{j_1} - Y_{j_2}\|^2 &= p^{-1} \|\delta\|^2 + p^{-1} \sum_{i=1}^p (X_{j_1,i} - X_{j_2,i})^2 - p^{-1} \sum_{i=1}^p 2\delta_i (X_{j_1,i} - X_{j_2,i}), \\ p^{-1} \|Y_{j_3} - Y_{j_4}\|^2 &= p^{-1} \|\delta\|^2 + p^{-1} \sum_{i=1}^p (X_{j_3,i} - X_{j_4,i})^2 - p^{-1} \sum_{i=1}^p 2\delta_i (X_{j_3,i} - X_{j_4,i}), \end{aligned}$$

and

$$\begin{aligned} p^{-1} (Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4}) &= p^{-1} \|\delta\|^2 + p^{-1} \sum_{i=1}^p (X_{j_1,i} - X_{j_2,i})(X_{j_3,i} - X_{j_4,i}) \\ &\quad - p^{-1} \sum_{i=1}^p \delta_i (X_{j_1,i} - X_{j_2,i}) - p^{-1} \sum_{i=1}^p \delta_i (X_{j_3,i} - X_{j_4,i}). \end{aligned}$$

Using Theorem 8.2.2 in [15], and the independence of X_i 's, we have

$$p^{-1} \|Y_{j_1} - Y_{j_2}\|^2 \xrightarrow{\mathcal{P}} \iota^2 + 2\sigma^2, \quad p^{-1} \|Y_{j_3} - Y_{j_4}\|^2 \xrightarrow{\mathcal{P}} \iota^2 + 2\sigma^2, \quad \text{and} \quad p^{-1} (Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4}) \xrightarrow{\mathcal{P}} \iota^2.$$

If $\iota > 0$, then

$$n^{-4} D^{(s)}(k^*; 1, n) \xrightarrow{\mathcal{P}} n^{-4} k^* (k^* - 1) (n - k^*) (n - k^* - 1) \frac{\iota^2}{\iota^2 + 2\sigma^2} > 0,$$

and

$$p^{-1}n^{-4}D(k^*; 1, n) \xrightarrow{\mathcal{P}} n^{-4}(k^*)(k^* - 1)(n - k^*)(n - k^* - 1)t^2 > 0.$$

Hence,

$$pn^{-3}\|\Sigma\|_F^{-1}D^{(s)}(k^*; 1, n) = (pn\|\Sigma\|_F^{-1})n^{-4}D^{(s)}(k^*; 1, n) \xrightarrow{\mathcal{P}} \infty,$$

and

$$n^{-3}\|\Sigma\|_F^{-1}D(k^*; 1, n) = (pn\|\Sigma\|_F^{-1})p^{-1}n^{-4}D(k^*; 1, n) \xrightarrow{\mathcal{P}} \infty. \quad \square$$

Proof of Theorem 3. By symmetry, we only consider the case $l < k \leq k^* < m$. Since under [Assumption 4](#), (11) still holds by Cauchy–Schwartz inequality, then using similar arguments in the proof of [Theorem 1](#), we have

$$\begin{aligned} 2p\sigma^2 D^{(s)}(k; l, m) &= \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} (X_{j_1} - X_{j_2})^\top (X_{j_3} - X_{j_4})(1 + o(1)) \\ &\quad + (k - l + 1)(k - l)(m - k^*)(m - k^* - 1)\|\delta\|^2(1 + o(1)) \\ &\quad - \left(2(k - l)(m - k^*)(m - k - 2) \sum_{j=l}^k X_j^\top \delta + 4(k - l)(k - l - 1)(m - k^*) \sum_{j=k+1}^{k^*} X_j^\top \delta \right) (1 + o(1)) \\ &:= D_{(1)}^{(s)}(k; l, m) + D_{(2)}^{(s)}(k; l, m) - D_{(3)}^{(s)}(k; l, m). \end{aligned} \quad (16)$$

That is, $2p\sigma^2 D^{(s)}(k; l, m) = D(k; l, m)(1 + o(1))$ for any triplet (k, l, m) , hence it suffices to consider $T_n^{(s)}$ as the results of T_n are similar.

We first note that

$$\text{Var}(X_i^\top \delta) = \delta^\top \Sigma \delta = o(\|\Sigma\|_F^2),$$

hence by Chebyshev inequality, for any triplet (k, l, m) , we have

$$n^{-3}\|\Sigma\|_F^{-1}D_{(3)}^{(s)}(k; l, m) = o_p(1). \quad (17)$$

(i) By similar arguments in the proof of [Theorem 2](#), it suffices to show

$$2p\sigma^2 n^{-3}\|\Sigma\|_F^{-1}D^{(s)}(k^*; 1, n) \xrightarrow{\mathcal{P}} \infty.$$

In fact, by similar arguments used in the proof of [Theorem 1](#), we can show that

$$n^{-3}\|\Sigma\|_F^{-1}D_{(1)}^{(s)}(k; l, m) = O_p(1).$$

Then, recall [\(16\)](#), the result follows by noting

$$n^{-3}\|\Sigma\|_F^{-1}D_{(2)}^{(s)}(k^*; 1, n) = n^{-3}\|\Sigma\|_F^{-1}(k - l + 1)(k - l)(m - k^*)(m - k^* - 1)\|\delta\|^2(1 + o(1)) \rightarrow \infty.$$

(ii) As $n\|\Sigma\|_F^{-1}\|\delta\|^2 \rightarrow 0$, it follows from the same argument as [\(14\)](#).

(iii) As $n\|\Sigma\|_F^{-1}\|\delta\|^2 \rightarrow c_n \in (0, \infty)$, then we have

$$n^{-3}\|\Sigma\|_F^{-1}D_{(2)}^{(s)}(k^*; m, l) = n^{-3}\|\Sigma\|_F^{-1}(k - l + 1)(k - l)(m - k^*)(m - k^* - 1)\|\delta\|^2(1 + o(1)) \rightarrow c_n \frac{4 \binom{k-l+1}{2} \binom{m-k^*}{2}}{n^4}.$$

Therefore, continuous mapping theorem together with [Lemma 1](#) indicate that

$$T_n^{(s)} \xrightarrow{\mathcal{D}} \sup_{k=4, \dots, n-4} \frac{n[\sqrt{2}G_n(\frac{k}{n}; \frac{1}{n}, 1) + c_n \Delta_n(\frac{k}{n}; \frac{1}{n}, 1)]^2}{\sum_{t=2}^{k-2} [\sqrt{2}G_n(\frac{t}{n}; \frac{1}{n}, \frac{k}{n}) + c_n \Delta_n(\frac{t}{n}; \frac{1}{n}, \frac{k}{n})]^2 + \sum_{t=k+2}^{n-2} [\sqrt{2}G_n(\frac{t}{n}; \frac{k+1}{n}, 1) + c_n \Delta_n(\frac{t}{n}; \frac{k+1}{n}, 1)]^2}.$$

The last part of the proof is similar to the proof of [Theorem 5\(ii\)](#) below, and is simpler, hence omitted. $\square$

Proof of Theorem 4. (i) Note that

$$\frac{1}{p}\|Y_i - Y_j\|^2 = \frac{1}{p} \sum_{\ell=1}^p \left(\frac{X_{i,\ell}}{R_i} - \frac{X_{j,\ell}}{R_j} \right)^2, \quad (18)$$

hence given $\mathcal{R}_n$, as $p \rightarrow \infty$, we have almost surely

$$\frac{1}{p}\|Y_i - Y_j\|^2 \rightarrow \sigma^2(R_i^{-2} + R_j^{-2}). \quad (19)$$

Note that

$$\begin{aligned} p\sigma^2 D^{(s)}(k; l, m) &= \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} \\ &\quad + \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} \left\{ \frac{p\sigma^2 (R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}}{\|Y_{j_1} - Y_{j_2}\| \|Y_{j_3} - Y_{j_4}\|} - 1 \right\} \\ &=: [D_3(k; l, m) + D_4(k; l, m)]. \end{aligned} \quad (20)$$

Let

$$\begin{aligned} A_{j_1, j_3}(k; l, m) &= \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2}, \\ B_{j_2, j_4}(k; l, m) &= \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2}, \\ C_{j_1, j_2}(k; l, m) &= -2 \sum_{\substack{l \leq j_3 \leq k \\ j_3 \neq j_1}} \sum_{\substack{k+1 \leq j_4 \leq m \\ j_4 \neq j_2}} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2}. \end{aligned}$$

Then under H_0 ,

$$\begin{aligned} D_3(k; l, m) &= \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} X_{j_1}^\top X_{j_3} (R_{j_1} R_{j_3})^{-1} A_{j_1, j_3}(k; l, m) + \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} X_{j_2}^\top X_{j_4} (R_{j_2} R_{j_4})^{-1} B_{j_2, j_4}(k; l, m) \\ &\quad + \sum_{l \leq j_1 \leq k} \sum_{k+1 \leq j_2 \leq m} X_{j_1}^\top X_{j_2} (R_{j_1} R_{j_2})^{-1} C_{j_1, j_2}(k; l, m). \end{aligned}$$

Denote that $U_1 = (X_1^\top X_2, \dots, X_1^\top X_n, X_2^\top X_3, \dots, X_2^\top X_n, \dots, X_{n-1}^\top X_n)^\top$ which contains all inner products of X_i and X_j for all $i \neq j$, and $U_2 = (R_1, \dots, R_n)^\top$. By definition, $\sigma(U_1) \perp \sigma(U_2)$, where $\sigma(U)$ is the σ -field generated by U , and we further observe that $2p\sigma^2 D_3(k; l, m)$ is a continuous functional of U_1 and U_2 . Hence to derive the limiting distribution of $2p\sigma^2 D_3(k; l, m)$ when $p \rightarrow \infty$, it suffices to derive the limiting distribution of $(U_1, U_2)^\top$.

For any $\alpha \in \mathbb{R}^{n(n-1)/2}$, similar to the proof of Theorem 1, by Theorem 4.0.1 in [15] we have

$$\|\Sigma\|_F^{-1} \alpha^\top U_1 \xrightarrow{D} \alpha^\top \mathcal{Z} := \alpha^\top (\mathcal{Z}_{1,2}, \mathcal{Z}_{1,3}, \dots, \mathcal{Z}_{1,n}, \mathcal{Z}_{2,3}, \dots, \mathcal{Z}_{2,n}, \dots, \mathcal{Z}_{n-1,n})^\top,$$

where $\mathcal{Z}_{1,2}, \dots, \mathcal{Z}_{n-1,n}$ are i.i.d. standard normal random variables, and we can assume $\mathcal{Z}$ is independent of U_2 . For the ease of our notation, we let $\mathcal{Z}_{i,j} = \mathcal{Z}_{j,i}$, for all $i > j$. Furthermore since $\sigma(U_1) \perp \sigma(U_2)$, for any $\alpha \in \mathbb{R}^{n(n-1)/2}$ and $\beta \in \mathbb{R}^n$, the characteristic function of $\alpha^\top U_1 + \beta^\top U_2$ is the product of the characteristic function of $\alpha^\top U_1$ and that of $\beta^\top U_2$. By applying the Cramér–Wold device, $(\|\Sigma\|_F^{-1} U_1, U_2) \xrightarrow{D} (\mathcal{Z}, U_2)$. Therefore, by continuous mapping theorem, as $p \rightarrow \infty$,

$$n^{-3} \|\Sigma\|_F^{-1} D_3(k; l, m) \xrightarrow{D} G_n^{(\mathcal{R}_n, s)}(k/n; l/n, m/n), \quad (21)$$

where

$$\begin{aligned} G_n^{(\mathcal{R}_n, s)}\left(\frac{k}{n}; \frac{l}{n}, \frac{m}{n}\right) &:= n^{-3} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \mathcal{Z}_{j_1, j_3} (R_{j_1} R_{j_3})^{-1} A_{j_1, j_3}(k, l, m) + n^{-3} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \mathcal{Z}_{j_2, j_4} (R_{j_2} R_{j_4})^{-1} B_{j_2, j_4}(k, l, m) \\ &\quad + n^{-3} \sum_{l \leq j_1 \leq k} \sum_{k+1 \leq j_2 \leq m} \mathcal{Z}_{j_1, j_2} (R_{j_1} R_{j_2})^{-1} C_{j_1, j_2}(k, l, m). \end{aligned} \quad (22)$$

It is clear that the conditional distribution of $G_n^{(\mathcal{R}_n, s)}(k/n; l/n, m/n)$ given $\mathcal{R}_n$ is Gaussian, and for any $l_1 < k_1 < m_1, l_2 < k_2 < m_2, k_1, k_2, l_1, l_2, m_1, m_2 \in \{1, 2, \dots, n\}$, the covariance structure is given by

$$\begin{aligned} &\text{Cov}(G_n^{(\mathcal{R}_n, s)}(k_1/n; l_1/n, m_1/n), G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_2/n) | \mathcal{R}_n) \\ &= 2n^{-6} \left\{ \sum_{\substack{(l_1 \vee l_2) \leq j_1, j_2 \leq (k_1 \wedge k_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} A_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_2, m_2) \right. \\ &\quad \left. + \sum_{\substack{(l_1 \vee k_2 + 1) \leq j_1, j_2 \leq (k_1 \wedge m_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} A_{j_1, j_2}(k_1; l_1, m_1) B_{j_1, j_2}(k_2; l_2, m_2) \right. \end{aligned} \quad (23)$$

$$\begin{aligned}
& + 2 \sum_{j_1=(l_1 \vee l_2)}^{k_2} \sum_{j_2=k_2+1}^{(m_2 \wedge k_1)} \mathbf{1}(k_1 > k_2) R_{j_1}^{-2} R_{j_2}^{-2} A_{j_1 j_2}(k_1; l_1, m_1) C_{j_1 j_2}(k_2; l_2, m_2) \\
& + \sum_{\substack{(k_1+1 \vee l_2) \leq j_1, j_2 \leq (m_1 \wedge k_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1 j_2}(k_1; l_1, m_1) A_{j_1 j_2}(k_2; l_2, m_2) \\
& + \sum_{\substack{(k_1+1 \vee k_2+1) \leq j_1, j_2 \leq (m_1 \wedge m_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1 j_2}(k_1; l_1, m_1) B_{j_1 j_2}(k_2; l_2, m_2) \\
& + 2 \sum_{j_1=(k_1+1 \vee l_2)}^{k_2} \sum_{j_2=k_2+1}^{(m_1 \wedge m_2)} \mathbf{1}(m_1 > k_2) R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1 j_2}(k_1; l_1, m_1) C_{j_1 j_2}(k_2; l_2, m_2) \\
& + 2 \sum_{j_1=(l_1 \vee l_2)}^{k_1} \sum_{j_2=k_1+1}^{(m_1 \wedge k_2)} \mathbf{1}(k_2 > k_1) R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1 j_2}(k_1; l_1, m_1) A_{j_1 j_2}(k_2; l_2, m_2) \\
& + 2 \sum_{j_1=(k_2+1 \vee l_1)}^{k_1} \sum_{j_2=k_1+1}^{(m_1 \wedge m_2)} \mathbf{1}(m_2 > k_1) R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1 j_2}(k_1; l_1, m_1) B_{j_1 j_2}(k_2; l_2, m_2) \\
& + \sum_{j_1=(l_1 \vee l_2)}^{(k_1 \wedge k_2)} \sum_{j_2=(k_1+1 \vee k_2+1)}^{(m_1 \wedge m_2)} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1 j_2}(k_1; l_1, m_1) C_{j_1 j_2}(k_2; l_2, m_2) \Big\}.
\end{aligned}$$

Clearly, when $R_i \equiv 1$, we have $2D_3(k; l, m) = D_1(k; l, m)$ where $D_1(k; l, m)$ is defined in (12), and the result reduces to (13).

Using (19), we can see that given $\mathcal{R}_n$, $\frac{D_4(k; l, m)}{n^3 \|\Sigma\|_F} = o_p(1)$. Hence, given $\mathcal{R}_n$, we have

$$T_n^{(s)} = \sup_{k=4, \dots, n-4} \frac{[D_3(k; 1, n)]^2}{\frac{1}{n} \sum_{t=l+1}^{k-2} D_3(t; l, k)^2 + \frac{1}{n} \sum_{t=k+2}^{m-2} D_3(t; k+1, m)^2} + o_p(1).$$

Then, by (21), we have that as $p \rightarrow \infty$,

$$T_n^{(s)} | \mathcal{R}_n \xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n, s)} := \sup_{k=4, \dots, n-4} \frac{n[G_n^{(\mathcal{R}_n, s)}(k/n; 1/n, 1)]^2}{\sum_{t=2}^{k-1} [G_n^{(\mathcal{R}_n, s)}(t/n; 1/n, k/n)]^2 + \sum_{t=k+2}^{n-2} [G_n^{(\mathcal{R}_n, s)}(t/n; (k+1)/n, 1)]^2}.$$

As for T_n , note that

$$\begin{aligned}
D(k; l, m) &= (m-k)(m-k-1) \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} X_{j_1}^T X_{j_3} (R_{j_1} R_{j_3})^{-1} + (k-l+1)(k-l) \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} X_{j_2}^T X_{j_4} (R_{j_2} R_{j_4})^{-1} \\
&\quad - 2(k-l)(m-k-1) \sum_{l \leq j_1 \leq k} \sum_{k+1 \leq j_2 \leq m} X_{j_1}^T X_{j_2} (R_{j_1} R_{j_2})^{-1}.
\end{aligned}$$

Using similar arguments as in (21), we have

$$n^{-3} \|\Sigma\|_F^{-1} D(k; l, m) \xrightarrow{\mathcal{D}} G_n^{(\mathcal{R}_n)}(k/n; l/n, m/n), \quad (24)$$

where

$$\begin{aligned}
G_n^{(\mathcal{R}_n)}\left(\frac{k}{n}; \frac{l}{n}, \frac{m}{n}\right) &= (m-k)(m-k-1)n^{-3} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \mathcal{Z}_{j_1 j_3} (R_{j_1} R_{j_3})^{-1} + (k-l+1)(k-l)n^{-3} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \mathcal{Z}_{j_2 j_4} (R_{j_2} R_{j_4})^{-1} \\
&\quad - 2(k-l)(m-k-1)n^{-3} \sum_{l \leq j_1 \leq k} \sum_{k+1 \leq j_2 \leq m} \mathcal{Z}_{j_1 j_2} (R_{j_1} R_{j_2})^{-1}.
\end{aligned} \quad (25)$$

Similar to $G_n^{(\mathcal{R}_n, s)}(k/n; l/n, m/n)$, the conditional distribution of $G_n^{(\mathcal{R}_n)}(k/n; l/n, m/n)$ given $\mathcal{R}_n$ is Gaussian, and for any $l_1 < k_1 < m_1$, $l_2 < k_2 < m_2$, $k_1, k_2, l_1, l_2, m_1, m_2 = 1, 2, \dots, n$, the covariance structure is given by

$$\begin{aligned}
& \text{Cov}(G_n^{(\mathcal{R}_n)}(k_1/n; l_1/n, m_1/n), G_n^{(\mathcal{R}_n)}(k_2/n; l_2/n, m_2/n) | \mathcal{R}_n) \\
&= 8n^{-6} \Big\{ \sum_{\substack{(l_1 \vee l_2) \leq j_1, j_2 \leq (k_1 \wedge k_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} \binom{m_1 - k_1}{2} \binom{m_2 - k_2}{2} + \sum_{\substack{(l_1 \vee k_2+1) \leq j_1, j_2 \leq (k_1 \wedge m_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} \binom{m_1 - k_1}{2} \binom{k_1 - l_1 + 1}{2} \Big\}
\end{aligned} \quad (26)$$

$$\begin{aligned}
& -2 \sum_{j_1=(l_1 \vee l_2)}^{k_2} \sum_{j_2=k_2+1}^{(m_2 \wedge k_1)} \mathbf{1}(k_1 > k_2) R_{j_1}^{-2} R_{j_2}^{-2} \binom{m_1 - k_1}{2} (k_2 - l_2)(m_2 - k_2 - 1) \\
& + \sum_{\substack{(k_1+1 \vee l_2) \leq j_1, j_2 \leq (m_1 \wedge k_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} \binom{k_1 - l_1 + 1}{2} \binom{m_2 - k_2}{2} + \sum_{\substack{(k_1+1 \vee k_2+1) \leq j_1, j_2 \leq (m_1 \wedge m_2) \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} \binom{k_1 - l_1 + 1}{2} \binom{k_2 - l_2 + 1}{2} \\
& -2 \sum_{j_1=(k_1+1 \vee l_2)}^{k_2} \sum_{j_2=k_2+1}^{(m_1 \wedge m_2)} \mathbf{1}(m_1 > k_2) R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1, j_2} \binom{k_1 - l_1 + 1}{2} (k_2 - l_2)(m_2 - k_2 - 1) \\
& -2 \sum_{j_1=(l_1 \vee l_2)}^{k_1} \sum_{j_2=k_1+1}^{(m_1 \wedge k_2)} \mathbf{1}(k_2 > k_1) R_{j_1}^{-2} R_{j_2}^{-2} (k_1 - l_1)(m_1 - k_1 - 1) \binom{m_2 - k_2}{2} \\
& -2 \sum_{j_1=(k_2+1 \vee l_1)}^{k_1} \sum_{j_2=k_1+1}^{(m_1 \wedge m_2)} \mathbf{1}(m_2 > k_1) R_{j_1}^{-2} R_{j_2}^{-2} (k_1 - l_1)(m_1 - k_1 - 1) \binom{k_2 - l_2 + 1}{2} \\
& + \sum_{j_1=(l_1 \vee l_2)}^{(k_1 \wedge k_2)} \sum_{j_2=(k_1+1 \vee k_2+1)}^{(m_1 \wedge m_2)} R_{j_1}^{-2} R_{j_2}^{-2} (k_1 - l_1)(m_1 - k_1 - 1)(k_2 - l_2)(m_2 - k_2 - 1) \Big\}.
\end{aligned}$$

Hence, as $p \rightarrow \infty$,

$$T_n | \mathcal{R}_n \xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n)} := \sup_{k=4, \dots, n-4} \frac{n[G_n^{(\mathcal{R}_n)}(k/n; 1/n, 1)]^2}{\sum_{t=2}^{k-1} [G_n^{(\mathcal{R}_n)}(t/n; 1/n, k/n)]^2 + \sum_{t=k+2}^{n-2} [G_n^{(\mathcal{R}_n)}(t/n; (k+1)/n, 1)]^2}.$$

(ii) We shall only show the process convergence $G_n^{(\mathcal{R}_n, s)}(\cdot) \rightsquigarrow \mathbb{E} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} G(\cdot)$, because that $G_n^{(\mathcal{R}_n)}(\cdot) \rightsquigarrow \mathbb{E}(R_1^{-2}) \sqrt{2} G(\cdot)$ is similar and simpler. Once the process convergence is obtained, the limiting distributions of $\mathcal{T}_n^{(\mathcal{R}_n, s)}$ and $\mathcal{T}_n^{(\mathcal{R}_n)}$ can be easily obtained by the continuous mapping theorem.

The proof for the process convergence contains two parts: the finite dimensional convergence and the tightness.

To show the finite dimensional convergence, we need to show that for any positive integer N , any fixed $u_1, u_2, \dots, u_N \in [0, 1]^3$ and any $\alpha_1, \dots, \alpha_N \in \mathbb{R}$,

$$\alpha_1 G_n^{(\mathcal{R}_n, s)}(u_1) + \dots + \alpha_N G_n^{(\mathcal{R}_n, s)}(u_N) \xrightarrow{\mathcal{D}} \mathbb{E}^2 \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} [\alpha_1 G(u_1) + \dots + \alpha_N G(u_N)],$$

where for $u = (u^{(1)}, u^{(2)}, u^{(3)})^T$, $G_n(u) = G_n(u_1; u_2, u_3)$. Since both $G_n^{(\mathcal{R}_n, s)}(\cdot) | \mathcal{R}_n$ and $G(\cdot)$ are Gaussian processes, by [Lemma 2](#) we have

$$\begin{aligned}
& \mathbb{P}_r(\alpha_1 G_n^{(\mathcal{R}_n, s)}(u_1) + \dots + \alpha_k G_n^{(\mathcal{R}_n, s)}(u_k) < x | \mathcal{R}_n) \\
& \xrightarrow{\mathcal{P}} \mathbb{P}_r \left(\mathbb{E} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} [\alpha_1 G(u_1) + \dots + \alpha_N G(u_N)] < x \right).
\end{aligned}$$

Then by bounded convergence theorem we have

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \mathbb{P}_r(\alpha_1 G_n^{(\mathcal{R}_n, s)}(u_1) + \dots + \alpha_k G_n^{(\mathcal{R}_n, s)}(u_k) < x) = \lim_{n \rightarrow \infty} \mathbb{E}[\mathbb{P}_r(\alpha_1 G_n^{(\mathcal{R}_n, s)}(u_1) + \dots + \alpha_k G_n^{(\mathcal{R}_n, s)}(u_k) < x | \mathcal{R}_n)] \\
& = \mathbb{E}[\lim_{n \rightarrow \infty} \mathbb{P}_r(\alpha_1 G_n^{(\mathcal{R}_n, s)}(u_1) + \dots + \alpha_k G_n^{(\mathcal{R}_n, s)}(u_k) < x | \mathcal{R}_n)] \\
& = \mathbb{E} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} [\alpha_1 G(u_1) + \dots + \alpha_k G(u_k)] < x \\
& = \mathbb{P}_r \left(\mathbb{E} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} [\alpha_1 G(u_1) + \dots + \alpha_k G(u_k)] < x \right).
\end{aligned}$$

This completes the proof of the finite dimensional convergence.

To show the tightness, it suffices to show that there exists $C > 0$ such that

$$\mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(u) - G_n^{(\mathcal{R}_n, s)}(v))^8] \leq C(\|u - v\|^4 + 1/n^4),$$

for any $u, v \in [0, 1]^3$ (see the proof of equation S8.12 in [\[28\]](#)).

Since given $\mathcal{R}_n$, $G_n^{(\mathcal{R}_n, s)}(\cdot)$ is a Gaussian process, we have

$$\mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(u) - G_n^{(\mathcal{R}_n, s)}(v))^8] = \mathbb{E}[\mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(u) - G_n^{(\mathcal{R}_n, s)}(v))^8 | \mathcal{R}_n]] = C\mathbb{E}[\text{Var}((G_n^{(\mathcal{R}_n, s)}(u) - G_n^{(\mathcal{R}_n, s)}(v)) | \mathcal{R}_n)^4].$$

By (23), for $u = (k_1/n, l_1/n, m_1/n)$ (and similar for $v = (k_2/n, l_2/n, m_2/n)$) this reduces to

$$\begin{aligned} \text{Var}(G_n^{(\mathcal{R}_n, s)}(u) | \mathcal{R}_n) = & 2n^{-6} \left\{ \sum_{\substack{l_1 \leq j_1, j_3 \leq k_1 \\ j_1 \neq j_3}} (R_{j_1} R_{j_3})^{-2} A_{j_1, j_3}(k_1, l_1, m_1)^2 + \sum_{\substack{k_1+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} (R_{j_2} R_{j_4})^{-2} B_{j_2, j_4}(k_1, l_1, m_1)^2 \right. \\ & \left. + \sum_{l_1 \leq j_1 \leq k_1} \sum_{k_1+1 \leq j_2 \leq m_1} (R_{j_1} R_{j_2})^{-2} C_{j_1, j_2}(k_1, l_1, m_1)^2 \right\}. \end{aligned}$$

Note that

$$\begin{aligned} & \mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(k_1/n; l_1/n, m_1/n) - G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_2/n))^8] \\ & \lesssim \mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(k_1/n; l_1/n, m_1/n) - G_n^{(\mathcal{R}_n, s)}(k_2/n; l_1/n, m_1/n))^8] \\ & + \mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(k_2/n; l_1/n, m_1/n) - G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_1/n))^8] \\ & + \mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_1/n) - G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_2/n))^8] := I_1 + I_2 + I_3. \end{aligned}$$

We shall analyze I_1 first, and WLOG we let $k_1 < k_2$. Then we have (with $l_1 = l_2, m_1 = m_2$)

$$\begin{aligned} & \text{Cov}(G_n^{(\mathcal{R}_n, s)}(k_1/n; l_1/n, m_1/n), G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_2/n) | \mathcal{R}_n) \\ = & 2n^{-6} \left\{ \sum_{\substack{l_1 \leq j_1, j_3 \leq k_1 \\ j_1 \neq j_3}} R_{j_1}^{-2} R_{j_3}^{-2} A_{j_1, j_3}(k_1; l_1, m_1) A_{j_1, j_3}(k_2; l_1, m_1) + \sum_{\substack{k_1+1 \leq j_1, j_2 \leq k_2 \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_2, m_2) \right. \\ & + \sum_{\substack{k_2+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} R_{j_2}^{-2} R_{j_4}^{-2} B_{j_2, j_4}(k_1; l_1, m_1) B_{j_2, j_4}(k_2; l_2, m_2) + 2 \sum_{j_1=k_1+1}^{k_2} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_2; l_2, m_2) B_{j_1, j_2}(k_1; l_1, m_1) \\ & \left. + 2 \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_1+1}^{k_2} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_2, m_2) + \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_1; l_1, m_1) C_{j_1, j_2}(k_2; l_2, m_2) \right\}. \end{aligned}$$

Hence,

$$\begin{aligned} & \text{Var}(G_n^{(\mathcal{R}_n, s)}(k_1/n; l_1/n, m_1/n) - G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_2/n) | \mathcal{R}_n) \\ = & 2n^{-6} \left\{ \sum_{\substack{l_1 \leq j_1, j_3 \leq k_1 \\ j_1 \neq j_3}} (R_{j_1} R_{j_3})^{-2} A_{j_1, j_3}(k_1, l_1, m_1)^2 + \sum_{\substack{k_1+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} (R_{j_2} R_{j_4})^{-2} B_{j_2, j_4}(k_1, l_1, m_1)^2 \right. \\ & \left. + \sum_{l_1 \leq j_1 \leq k_1} \sum_{k_1+1 \leq j_2 \leq m_1} (R_{j_1} R_{j_2})^{-2} C_{j_1, j_2}(k_1, l_1, m_1)^2 \right\} \\ & + 2n^{-6} \left\{ \sum_{\substack{l_1 \leq j_1, j_3 \leq k_2 \\ j_1 \neq j_3}} (R_{j_1} R_{j_3})^{-2} A_{j_1, j_3}(k_2, l_1, m_1)^2 + \sum_{\substack{k_2+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} (R_{j_2} R_{j_4})^{-2} B_{j_2, j_4}(k_2, l_1, m_1)^2 \right. \\ & \left. + \sum_{l_1 \leq j_1 \leq k_2} \sum_{k_2+1 \leq j_2 \leq m_1} (R_{j_1} R_{j_2})^{-2} C_{j_1, j_2}(k_2, l_1, m_1)^2 \right\} \\ & - 4n^{-6} \left\{ \sum_{\substack{l_1 \leq j_1, j_3 \leq k_1 \\ j_1 \neq j_3}} R_{j_1}^{-2} R_{j_3}^{-2} A_{j_1, j_3}(k_1; l_1, m_1) A_{j_1, j_3}(k_2; l_1, m_1) + \sum_{\substack{k_1+1 \leq j_1, j_2 \leq k_2 \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_2, m_2) \right. \\ & + \sum_{\substack{k_2+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} R_{j_2}^{-2} R_{j_4}^{-2} B_{j_2, j_4}(k_1; l_1, m_1) B_{j_2, j_4}(k_2; l_2, m_2) + 2 \sum_{j_1=k_1+1}^{k_2} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_2; l_2, m_2) B_{j_1, j_2}(k_1; l_1, m_1) \\ & \left. + 2 \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_1+1}^{k_2} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_2, m_2) + \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_1; l_1, m_1) C_{j_1, j_2}(k_2; l_2, m_2) \right\}. \end{aligned}$$

By rearranging terms we have

$$\begin{aligned}
& \text{Var}(G_n^{(\mathcal{R}_n, s)}(k_1/n; l_1/n, m_1/n) - G_n^{(\mathcal{R}_n, s)}(k_2/n; l_2/n, m_2/n) | \mathcal{R}_n) \\
&= 2n^{-6} \left\{ \sum_{\substack{l_1 \leq j_1, j_3 \leq k_1 \\ j_1 \neq j_3}} (R_{j_1} R_{j_3})^{-2} (A_{j_1, j_3}(k_1, l_1, m_1) - A_{j_1, j_3}(k_2, l_1, m_1))^2 \right. \\
&+ \sum_{j_1=k_1+1}^{k_2} \sum_{j_3=l_1}^{j_1-1} (R_{j_1} R_{j_3})^{-2} (A_{j_1, j_3}(k_2, l_1, m_1)^2 + A_{j_3, j_1}(k_2, l_1, m_1)^2) \\
&+ \sum_{\substack{k_2+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} R_{j_2}^{-2} R_{j_4}^{-2} (B_{j_2, j_4}(k_1; l_1, m_1) - B_{j_2, j_4}(k_2; l_1, m_1))^2 \\
&+ \sum_{j_2=k_1+1}^{k_2-1} \sum_{j_4=j_2+1}^{m_1} R_{j_2}^{-2} R_{j_4}^{-2} (B_{j_2, j_4}(k_1; l_1, m_1)^2 + B_{j_4, j_2}(k_1; l_1, m_1)^2) \\
&+ \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} (C_{j_1, j_2}(k_1, l_1, m_1) - C_{j_1, j_2}(k_2, l_1, m_1))^2 \\
&+ \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_1+1}^{k_2} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_1, l_1, m_1)^2 + \sum_{j_1=k_1+1}^{k_2} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_2, l_1, m_1)^2 \\
&- 2 \sum_{\substack{k_1+1 \leq j_1, j_2 \leq k_2 \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} B_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_1, m_1) - 4 \sum_{j_1=k_1+1}^{k_2} \sum_{j_2=k_2+1}^{m_1} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_2; l_1, m_1) B_{j_1, j_2}(k_1; l_1, m_1) \\
&- 4 \sum_{j_1=l_1}^{k_1} \sum_{j_2=k_1+1}^{k_2} R_{j_1}^{-2} R_{j_2}^{-2} C_{j_1, j_2}(k_1; l_1, m_1) A_{j_1, j_2}(k_2; l_1, m_1) =: \sum_{l=1}^{10} J_l.
\end{aligned}$$

Thus by CR-inequality we have $I_1 = \mathbb{E}[(\sum_{l=1}^{10} J_l)^4] \lesssim \sum_{l=1}^{10} \mathbb{E}[J_l^4]$. We shall analyze $\mathbb{E}[J_1^4]$ first. Note that

$$\begin{aligned}
& A_{j_1, j_3}(k_1, l_1, m_1) - A_{j_1, j_3}(k_2, l_1, m_1) \\
&= \sum_{\substack{k_1+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2} - \sum_{\substack{k_2+1 \leq j_2, j_4 \leq m_1 \\ j_2 \neq j_4}} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2} \\
&= \sum_{j_2=k_1+1}^{k_2} \sum_{j_4=j_2+1}^{m_1} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2} + \sum_{j_4=k_1+1}^{k_2} \sum_{j_2=j_4+1}^{m_1} (R_{j_1}^{-2} + R_{j_2}^{-2})^{-1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{-1/2} \\
&\leq 2 \sum_{j_2=k_1+1}^{k_2} \sum_{j_4=j_2+1}^{m_1} (2|R_{j_1}|^{-1} |R_{j_2}|^{-1})^{-1/2} (2|R_{j_3}|^{-1} |R_{j_4}|^{-1})^{-1/2} \\
&\leq |R_{j_1} R_{j_3}|^{1/2} \sum_{j_2=k_1+1}^{k_2} \sum_{j_4=j_2+1}^{m_1} |R_{j_2} R_{j_4}|^{1/2}.
\end{aligned}$$

Since $A_{j_1, j_3}(k_1, l_1, m_1) - A_{j_1, j_3}(k_2, l_1, m_1) > 0$ and $J_1 > 0$ almost surely, we have

$$\mathbb{E}[J_1^4] \leq 2^4 n^{-24} \mathbb{E} \left[\prod_{i=1}^4 \left(\sum_{\substack{l_1 \leq j_{1,i}, j_{3,i} \leq k_1 \\ j_{1,i} \neq j_{3,i}}} \sum_{j_{2,i}=k_1+1}^{k_2} \sum_{j_{4,i}=j_{2,i}+1}^{m_1} \sum_{j'_{2,i}=k_1+1}^{k_2} \sum_{j'_{4,i}=j'_{2,i}+1}^{m_1} |R_{j_{1,i}} R_{j_{3,i}}|^{-1} |R_{j_{2,i}} R_{j_{4,i}}|^{1/2} |R_{j'_{2,i}} R_{j'_{4,i}}|^{1/2} \right) \right].$$

By the Hölder's inequality, and the fact that $j_{1,s} \neq j_{3,s}, j_{2,s} \neq j_{4,s}, j'_{2,s} \neq j'_{4,s}$ and $j_{1,s}, j_{3,s}$ are not identical to any of $\{j_{2,s}, j_{4,s}, j_{2,s}, j_{4,s}\}$ for any $s = 1, 2, 3, 4$, we have

$$\begin{aligned}
& \mathbb{E}[|R_{j_{1,1}} R_{j_{3,1}}|^{-1} |R_{j_{2,1}} R_{j_{4,1}}|^{1/2} |R_{j'_{2,1}} R_{j'_{4,1}}|^{1/2} \cdots |R_{j_{1,4}} R_{j_{3,4}}|^{-1} |R_{j_{2,4}} R_{j_{4,4}}|^{1/2} |R_{j'_{2,4}} R_{j'_{4,4}}|^{1/2}] \\
&\leq \prod_{s=1}^4 \mathbb{E}[(|R_{j_{1,s}}| |R_{j_{3,s}}|)^{-1} |R_{j_{2,s}} R_{j_{4,s}} R_{j'_{2,s}} R_{j'_{4,s}}|^{1/2}]^{1/4}
\end{aligned}$$

$$\begin{aligned}
&= \prod_{s=1}^4 \mathbb{E}[(R_{j_{1,s}} R_{j_{3,s}})^{-4} R_{j_{2,s}}^2 R_{j_{4,s}}^2 R_{j_{2,s}}^2 R_{j_{4,s}}^2]^{1/4} = \prod_{s=1}^4 \left\{ \mathbb{E}[R_{j_{1,s}}^{-4}] \mathbb{E}[R_{j_{3,s}}^{-4}] \mathbb{E}[R_{j_{2,s}}^2 R_{j_{4,s}}^2 R_{j_{2,s}}^2 R_{j_{4,s}}^2] \right\}^{1/4} \\
&\leq \prod_{s=1}^4 \left\{ \mathbb{E}[R_{j_{1,s}}^{-4}] \mathbb{E}[R_{j_{3,s}}^{-4}] \mathbb{E}[R_{j_{2,s}}^4 R_{j_{4,s}}^4]^{1/2} \mathbb{E}[R_{j_{2,s}}^4 R_{j_{4,s}}^4]^{1/2} \right\}^{1/4} = \mathbb{E}[R_1^{-4}]^2 \mathbb{E}[R_2^4]^2.
\end{aligned}$$

Therefore,

$$\mathbb{E}[J_s^4] \lesssim 2^4 n^{-24} (k_2 - k_1)^8 n^{16} \mathbb{E}[R_1^{-4}]^2 \mathbb{E}[R_2^4]^2 = 2^4 \mathbb{E}[R_1^{-4}]^2 \mathbb{E}[R_2^4]^2 (k_2/n - k_1/n)^8.$$

We repeatedly apply the Hölder's inequality and the above bound for the expectation, and we have $\mathbb{E}[J_s^4] \lesssim (k_2/n - k_1/n)^8$ for $s = 1, 3, 5, 8$ since there are 8 summations in each $\mathbb{E}[J_s^4]$ which take the sum from $k_1 + 1$ to k_2 , and $\mathbb{E}[J_s^4] \lesssim (k_2/n - k_1/n)^4$ for $s = 2, 4, 6, 7, 9, 10$ since there are only 4 summations in each $\mathbb{E}[J_s^4]$ which take the sum from $k_1 + 1$ to k_2 . Combining these results we have $I_1 \lesssim (k_2/n - k_1/n)^4$.

We can also show $I_2 \lesssim (l_2/n - l_1/n)^4$, and $I_3 \lesssim (m_2/n - m_1/n)^4$. Since the steps are very similar to the arguments for I_1 , we omit the details here. Thus, for any $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3) \in [0, 1]^3$, we have

$$\mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(u) - G_n^{(\mathcal{R}_n, s)}(v))^8] \leq C'((\lfloor nu_1 \rfloor / n - \lfloor nv_1 \rfloor / n)^4 + (\lfloor nu_2 \rfloor / n - \lfloor nv_2 \rfloor / n)^4 + (\lfloor nu_3 \rfloor / n - \lfloor nv_3 \rfloor / n)^4),$$

for some positive constant $C' > 0$. It is easy to see that

$$\begin{aligned}
(\lfloor nu_1 \rfloor / n - \lfloor nv_1 \rfloor / n)^4 &= ((u_1 - v_1) - (\{u_1\} - \{v_1\})/n)^4 \lesssim (u_1 - v_1)^4 + (\{u_1\} - \{v_1\})^4 / n^4 \\
&\lesssim (u_1 - v_1)^4 + 1/n^4.
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{E}[(G_n^{(\mathcal{R}_n, s)}(u) - G_n^{(\mathcal{R}_n, s)}(v))^8] &\leq C((u_1 - v_1)^4 + (u_2 - v_2)^4 + (u_3 - v_3)^4 + 1/n^4) \\
&= C(\|u - v\|_4^4 + 1/n^4) \leq C(\|u - v\|^4 + 1/n^4),
\end{aligned}$$

since $\|u - v\|_4^4 = \sum_{i=1}^3 (u_i - v_i)^4 \leq \sum_{i=1}^3 (u_i - v_i)^2 (u_i + v_i)^2 = (\sum_{i=1}^3 (u_i - v_i)^2)^2 = \|u - v\|^4$. This completes the proof of tightness. $\square$

Proof of Theorem 5. (i) Under Assumption 4, conditional on $\mathcal{R}_n$, we still have almost surely

$$\frac{1}{p} \|Y_i - Y_j\|^2 = \frac{1}{p} \sum_{k=1}^p \left(\frac{X_{i,k}}{R_i} - \frac{X_{j,k}}{R_j} \right)^2 + \frac{2}{p} (\mu_i - \mu_j) \left(\frac{X_{i,k}}{R_i} - \frac{X_{j,k}}{R_j} \right) + \frac{1}{p} \|\mu_i - \mu_j\|^2 \xrightarrow{p} \sigma^2(R_i^{-2} + R_j^{-2})$$

as conditional on $\mathcal{R}_n$, $\{R_i^{-1} X_{i,k}\}_{k=1}^p$ is still a ρ -mixing sequence.

Recall (20), conditional on $\mathcal{R}_n$, we mainly work on $D_3(k; l, m)$ since $D_4(k; l, m)$ is of a smaller order, where

$$D_3(k; l, m) = \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(Y_{j_1} - Y_{j_2})^\top (Y_{j_3} - Y_{j_4})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}}.$$

By symmetry, we only consider the case $l < k \leq k^* < m$, and the summation in $D_3(k; l, m)$ can be decomposed into

$$\sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} = \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \left\{ \sum_{\substack{k+1 \leq j_2, j_4 \leq k^* \\ j_2 \neq j_4}} + \sum_{\substack{k^*+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} + \sum_{j_2=k+1}^{k^*} \sum_{j_4=k^*+1}^m + \sum_{j_4=k+1}^{k^*} \sum_{j_2=k^*+1}^m \right\},$$

according to the relative location of j_2, j_4 and k^* .

Then, it is not hard to see that

$$\begin{aligned}
&D_3(k; l, m) \\
&= \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(X_{j_1}/R_{j_1} - X_{j_2}/R_{j_2})^\top (X_{j_3}/R_{j_3} - X_{j_4}/R_{j_4})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} + \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k^*+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{\|\delta\|^2}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} \\
&\quad - \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{j_2=k^*+1}^m \sum_{j_4=k+1, j_4 \neq j_2}^m \frac{\delta^\top (X_{j_1}/R_{j_1} - X_{j_2}/R_{j_2})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} - \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{j_4=k^*+1}^{k^*} \sum_{j_2=k+1, j_2 \neq j_4}^m \frac{\delta^\top (X_{j_3}/R_{j_3} - X_{j_4}/R_{j_4})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} \\
&:= \sum_{i=1}^4 D_{3,i}(k; l, m).
\end{aligned}$$

Under [Assumption 4](#), and conditional on $\mathcal{R}_n$, similar to (17), we can show for $i = 3, 4$.

$$n^{-3} \|\Sigma\|_F^{-1} D_{3,i}(k; l, m) = o_p(1),$$

while by (21), $n^{-3} \|\Sigma\|_F^{-1} D_{3,1}(k; l, m) \xrightarrow{\mathcal{D}} G^{(\mathcal{R}_n, s)}(k/n; l/n, m/n)$.

Hence, if $n\mathbb{E}(R^{-2})^{-1} \|\Sigma\|_F^{-1} \|\delta\|^2 \rightarrow c_n$ as $p \rightarrow \infty$, then conditional on $\mathcal{R}_n$, we obtain that

$$\begin{aligned} n^{-3} \|\Sigma\|_F^{-1} p\sigma^2 D^{(s)}(k; l, m) &= n^{-3} \|\Sigma\|_F^{-1} D_3(k; l, m) + o_p(1) \\ &= n^{-3} \|\Sigma\|_F^{-1} [D_{3,1}(k; l, m) + D_{3,2}(k; l, m)] + o_p(1) \\ &= n^{-3} \|\Sigma\|_F^{-1} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{(X_{j_1}/R_{j_1} - X_{j_2}/R_{j_2})^\top (X_{j_3}/R_{j_3} - X_{j_4}/R_{j_4})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} \\ &\quad + n^{-4} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k^*+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{n \|\Sigma\|_F^{-1} \|\delta\|^2}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}} + o_p(1) \\ &\xrightarrow{\mathcal{D}} G^{(\mathcal{R}_n, s)}(k/n; l/n, m/n) + c_n \Delta_n^{(\mathcal{R}_n, s)}(k/n; l/n, m/n), \end{aligned}$$

where

$$\Delta_n^{(\mathcal{R}_n, s)}(k/n; l/n, m/n) = \begin{cases} n^{-4} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k^*+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{\mathbb{E}(R^{-2})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}}, & l < k \leq k^* < m \\ n^{-4} \sum_{\substack{l \leq j_1, j_3 \leq k^* \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} \frac{\mathbb{E}(R^{-2})}{(R_{j_1}^{-2} + R_{j_2}^{-2})^{1/2} (R_{j_3}^{-2} + R_{j_4}^{-2})^{1/2}}, & l < k^* < k < m \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

Hence, we have

$$\begin{aligned} T_n^{(s)} | \mathcal{R}_n &\xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n, s)}(c_n, \Delta_n^{(\mathcal{R}_n, s)}) \\ &:= \sup_{k=4, \dots, n-4} \frac{n[G_n^{(\mathcal{R}_n, s)}(\frac{k}{n}, \frac{1}{n}, 1) + c_n \Delta_n^{(\mathcal{R}_n, s)}(\frac{k}{n}, \frac{1}{n}, 1)]^2}{\sum_{t=2}^{k-1} [G_n^{(\mathcal{R}_n, s)}(\frac{t}{n}, \frac{1}{n}, \frac{k}{n}) + c_n \Delta_n^{(\mathcal{R}_n, s)}(\frac{t}{n}, \frac{1}{n}, \frac{k}{n})]^2 + \sum_{t=k+2}^{n-2} [G_n^{(\mathcal{R}_n, s)}(\frac{t}{n}, \frac{(k+1)}{n}, 1) + c_n \Delta_n^{(\mathcal{R}_n, s)}(\frac{t}{n}, \frac{(k+1)}{n}, 1)]^2}. \end{aligned}$$

For T_n , by similar arguments as above, we have

$$\begin{aligned} n^{-3} \|\Sigma\|_F^{-1} D(k; l, m) &= n^{-3} \|\Sigma\|_F^{-1} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} (X_{j_1}/R_{j_1} - X_{j_2}/R_{j_2})^\top (X_{j_3}/R_{j_3} - X_{j_4}/R_{j_4}) \\ &\quad + n^{-4} \sum_{\substack{l \leq j_1, j_3 \leq k \\ j_1 \neq j_3}} \sum_{\substack{k^*+1 \leq j_2, j_4 \leq m \\ j_2 \neq j_4}} n \|\Sigma\|_F^{-1} \|\delta\|^2 + o_p(1) \xrightarrow{\mathcal{D}} G^{(\mathcal{R}_n)}(k/n; l/n, m/n) + c_n \mathbb{E}(R^{-2}) \Delta_n(k/n; l/n, m/n). \end{aligned}$$

Hence, we have

$$\begin{aligned} T_n^{(s)} | \mathcal{R}_n &\xrightarrow{\mathcal{D}} \mathcal{T}_n^{(\mathcal{R}_n, s)}(c_n, \Delta_n^{(\mathcal{R}_n, s)}) \\ &:= \sup_{k=4, \dots, n-4} \frac{n[G_n^{(\mathcal{R}_n)}(\frac{k}{n}, \frac{1}{n}, 1) + c_n \mathbb{E}(R^{-2}) \Delta_n(\frac{k}{n}, \frac{1}{n}, 1)]^2}{\sum_{t=2}^{k-1} [G_n^{(\mathcal{R}_n)}(\frac{t}{n}, \frac{1}{n}, \frac{k}{n}) + c_n \mathbb{E}(R^{-2}) \Delta_n(\frac{t}{n}, \frac{1}{n}, \frac{k}{n})]^2 + \sum_{t=k+2}^{n-2} [G_n^{(\mathcal{R}_n)}(\frac{t}{n}, \frac{(k+1)}{n}, 1) + c_n \mathbb{E}(R^{-2}) \Delta_n(\frac{t}{n}, \frac{(k+1)}{n}, 1)]^2}. \end{aligned}$$

(ii) Note that for any $u = (u_1, u_2, u_3)^\top \in [0, 1]^3$ such that $u_2 \leq u_1 \leq u_3$, as $n \rightarrow \infty$,

$$\Delta_n^{(\mathcal{R}_n, s)}(\lfloor nu_1 \rfloor/n; \lfloor nu_2 \rfloor/n, \lfloor nu_3 \rfloor/n) \xrightarrow{\mathcal{P}} \mathbb{E}(R_1^{-2}) E^2 \left[\frac{R_1 R_2}{\sqrt{R_1^2 + R_2^2}} \right] \Delta(u_1; u_2, u_3).$$

by the law of large numbers for U -statistics (since $\Delta_n^{(\mathcal{R}_n, s)}$ can be viewed as a two sample U -statistic). Then using the similar arguments in the proof of [Theorem 4](#) (ii), we have

$$\Delta_n^{(\mathcal{R}_n, s)}(\cdot) \rightsquigarrow \mathbb{E}(R_1^{-2}) E^2 \left[\frac{R_1 R_2}{\sqrt{R_1^2 + R_2^2}} \right] \Delta(\cdot).$$

Note that $\Delta(\cdot)$ is deterministic, and recall $G_n^{(\mathcal{R}_n, s)}(\cdot) \rightsquigarrow \mathbb{E} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} G(\cdot)$ in the proof of [Theorem 4](#)(ii), by similar arguments in the proof of [Theorem 3.6](#) in [28], we have

$$G_n^{(\mathcal{R}_n, s)}(\cdot) + c_n \Delta_n^{(\mathcal{R}_n, s)}(\cdot) \rightsquigarrow \mathbb{E} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \sqrt{2} G(\cdot) + c \mathbb{E}(R_1^{-2}) E^2 \left[\frac{R_1 R_2}{\sqrt{R_1^2 + R_2^2}} \right] \Delta(\cdot).$$

Similarly,

$$G_n^{(\mathcal{R}_n)}(\cdot) + c_n \mathbb{E}(R^{-2}) \Delta_n^{(\mathcal{R}_n)}(\cdot) \rightsquigarrow \mathbb{E}(R^{-2}) \sqrt{2} G(\cdot) + c \mathbb{E}(R_1^{-2}) \Delta(\cdot).$$

The result follows by the continuous mapping theorem. Here, the multiplicative $K > 1$ follows by the proof of Theorem 3.2 in [5]. $\square$

Appendix B. Auxiliary lemmas

Lemma 1. Under Assumptions 1 and 2, let $n \geq 8$ be a fixed number, and for any $0 \leq k < m \leq n$, let $Z(k, m) = \sum_{i=k+1}^m \sum_{j=k}^{i-1} X_i^\top X_j$. Then, as $p \rightarrow \infty$,

$$\frac{\sqrt{2}}{n \|\Sigma\|_F} Z(k, m) \xrightarrow{\mathcal{D}} Q_n\left(\frac{k}{n}, \frac{m}{n}\right),$$

where $Q_n(a, b)$ is a centered Gaussian process defined on $[0, 1]^2$ with covariance structure given by:

$$\begin{aligned} & \text{Cov}(Q_n(a_1, b_1), Q_n(a_2, b_2)) \\ &= n^{-2} (\lfloor nb_1 \rfloor \wedge \lfloor nb_2 \rfloor - \lfloor na_1 \rfloor \vee \lfloor na_2 \rfloor) (\lfloor nb_1 \rfloor \wedge \lfloor nb_2 \rfloor - \lfloor na_1 \rfloor \vee \lfloor na_2 \rfloor + 1) \mathbf{1}(b_1 \wedge b_2 > a_1 \vee a_2). \end{aligned}$$

Proof. By Cramér–Wold device, it suffices to show that for fixed n and N , any sequences of $\{\alpha_i\}_{i=1}^N$, $\alpha_i \in \mathbb{R}$,

$$\sum_{i=1}^N \alpha_i \frac{\sqrt{2}}{n \|\Sigma\|_F} Z(k_i, m_i) \xrightarrow{\mathcal{D}} \sum_{i=1}^N \alpha_i Q_n\left(\frac{k_i}{n}, \frac{m_i}{n}\right),$$

where $1 \leq k_i \leq m_i \leq n$ are integers.

For simplicity, we consider the case of $N = 2$, and by symmetry there are basically three types of enumerations of (k_1, m_1, k_2, m_2) : (1) $k_1 \leq m_1 \leq k_2 \leq m_2$; (2) $k_1 \leq k_2 \leq m_1 \leq m_2$; (3) $k_1 \leq k_2 \leq m_2 \leq m_1$.

Define $\xi_{i,t}^{(1)} = X_{i,t} \sum_{j=k_1}^{i-1} X_{j,t}$, and $\xi_{i,t}^{(2)} = X_{i,t} \sum_{j=k_2}^{i-1} X_{j,t}$. Then, we can show

$$\begin{aligned} & \frac{\sqrt{2}}{n \|\Sigma\|_F} [\alpha_1 Z(k_1, m_1) + \alpha_2 Z(k_2, m_2)] = \frac{\sqrt{2}}{n \|\Sigma\|_F} \left(\alpha_1 \sum_{i=k_1+1}^{m_1} \sum_{j=k_1}^{i-1} X_i^\top X_j + \alpha_2 \sum_{i=k_2+1}^{m_2} \sum_{j=k_2}^{i-1} X_i^\top X_j \right) \\ &= \begin{cases} \frac{\sqrt{2}}{n \|\Sigma\|_F} \sum_{t=1}^p \left(\sum_{i=k_1+1}^{m_1} \alpha_1 \xi_{i,t}^{(1)} + \sum_{i=k_2+1}^{m_2} \alpha_2 \xi_{i,t}^{(2)} \right), & \text{Case (1)} \\ \frac{\sqrt{2}}{n \|\Sigma\|_F} \sum_{t=1}^p \left(\sum_{i=k_1+1}^{k_2} \alpha_1 \xi_{i,t}^{(1)} + \sum_{i=k_2+1}^{m_1} [\alpha_1 \xi_{i,t}^{(1)} + \alpha_2 \xi_{i,t}^{(2)}] + \sum_{i=m_1+1}^{m_2} \alpha_2 \xi_{i,t}^{(2)} \right), & \text{Case (2)} \\ \frac{\sqrt{2}}{n \|\Sigma\|_F} \sum_{t=1}^p \left(\sum_{i=k_1+1}^{k_2} \alpha_1 \xi_{i,t}^{(1)} + \sum_{i=k_2+1}^{m_2} [\alpha_1 \xi_{i,t}^{(1)} + \alpha_2 \xi_{i,t}^{(2)}] + \sum_{i=m_2+1}^{m_1} \alpha_1 \xi_{i,t}^{(1)} \right), & \text{Case (3)} \end{cases} \end{aligned}$$

For simplicity, we consider the Case (2), and using the independence of X_i , one can show that

$$S_1 = \frac{\sqrt{2}}{n \|\Sigma\|_F} \sum_{t=1}^p \sum_{i=k_1+1}^{k_2} \alpha_1 \xi_{i,t}^{(1)}, \quad S_2 = \frac{\sqrt{2}}{n \|\Sigma\|_F} \sum_{t=1}^p [\alpha_1 \xi_{i,t}^{(1)} + \alpha_2 \xi_{i,t}^{(2)}], \quad S_3 = \frac{\sqrt{2}}{n \|\Sigma\|_F} \sum_{t=1}^p \sum_{i=k_2+1}^{m_2} \alpha_2 \xi_{i,t}^{(2)}$$

are independent. Then by Theorem 4.0.1 in Lin and Lu (2010), they are asymptotically normal with variances given by

$$\begin{aligned} \text{Var}(S_1) &= n^{-2} \alpha_1^2 (k_2 - k_1)(k_2 - k_1 + 1), \\ \text{Var}(S_2) &= n^{-2} [\alpha_1^2 (m_1 - k_2)(k_2 - k_1 + 1 + m_1 - k_1) + 2\alpha_1 \alpha_2 (m_1 - k_2)(m_1 - k_2 + 1) + \alpha_2^2 (m_1 - k_2)(m_1 - k_2 + 1)], \\ \text{Var}(S_3) &= n^{-2} \alpha_2^2 (m_2 - m_1)(m_2 - k_2 + m_1 - k_2 + 1). \end{aligned}$$

Similarly, we can obtain the asymptotic normality for Case (1) and Case (3).

Hence,

$$\frac{\sqrt{2}}{n \|\Sigma\|_F} [\alpha_1 Z(k_1, m_1) + \alpha_2 Z(k_2, m_2)] \xrightarrow{\mathcal{D}} N(0, \frac{\tau^2}{n^2}).$$

where

$$\tau^2 = \begin{cases} \alpha_1^2(m_1 - k_1)(m_1 - k_1 + 1) + \alpha_2^2(m_2 - k_2)(m_2 - k_2 + 1), & \text{Case (1)} \\ \alpha_1^2(m_1 - k_1)(m_1 - k_1 + 1) + \alpha_2^2(m_2 - k_2)(m_2 - k_2 + 1) + 2\alpha_1\alpha_2(m_1 - k_2)(m_1 - k_2 + 1), & \text{Case (2)} \\ \alpha_1^2(m_1 - k_1)(m_1 - k_1 + 1) + \alpha_2^2(m_2 - k_2)(m_2 - k_2 + 1) + 2\alpha_1\alpha_2(m_2 - k_2)(m_2 - k_2 + 1), & \text{Case (3)}. \end{cases}$$

Hence, the case of $N = 2$ is proved by examining the covariance structure of Q_n defined in Theorem 1. The cases $N > 2$ are similar. $\square$

Lemma 2. As $n \rightarrow \infty$, we have for any $0 \leq a_1 < r_1 < b_1 \leq 1$ and $0 \leq a_2 < r_2 < b_2 \leq 1$, as $n \rightarrow \infty$,

$$\text{Cov}(G_n^{(\mathcal{R}_{n,s})}(r_1; a_1, b_1), G_n^{(\mathcal{R}_{n,s})}(r_2; a_2, b_2)) \xrightarrow{\mathcal{P}} 2\mathbb{E}^2 \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \text{Cov}(G(r_1; a_1, b_1), G(r_2; a_2, b_2)).$$

Proof. There are 9 terms in the covariance structure given in (23), for first one, we have

$$\begin{aligned} & 2n^{-6} \sum_{\substack{[n(a_1 \vee a_2)] \leq j_1, j_2 \leq [n(r_1 \wedge r_2)] \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} A_{j_1 j_2}(\lfloor nr_1 \rfloor; \lfloor na_1 \rfloor, \lfloor nb_1 \rfloor) A_{j_1 j_2}(\lfloor nr_2 \rfloor; \lfloor na_2 \rfloor, \lfloor nb_2 \rfloor) \\ &= 2n^{-6} \sum_{\substack{[n(a_1 \vee a_2)] \leq j_1, j_2 \leq [n(r_1 \wedge r_2)] \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} \sum_{\substack{[nr_1] + 1 \leq j_3, j_4 \leq [nb_1] \\ j_3 \neq j_4}} \frac{R_{j_1} R_{j_3}}{\sqrt{(R_{j_1}^2 + R_{j_3}^2)}} \frac{R_{j_2} R_{j_4}}{\sqrt{(R_{j_2}^2 + R_{j_4}^2)}} \\ &\quad \times \sum_{\substack{[nr_2] + 1 \leq j_5, j_6 \leq [nb_2] \\ j_5 \neq j_6}} \frac{R_{j_1} R_{j_5}}{\sqrt{(R_{j_1}^2 + R_{j_5}^2)}} \frac{R_{j_2} R_{j_6}}{\sqrt{(R_{j_2}^2 + R_{j_6}^2)}} \\ &= 2n^{-2} \sum_{\substack{[n(a_1 \vee a_2)] \leq j_1, j_2 \leq [n(r_1 \wedge r_2)] \\ j_1 \neq j_2}} R_{j_1}^{-2} R_{j_2}^{-2} (b_1 - r_1)^2 \left\{ \mathbb{E} \left[\frac{R_{j_1} R_{j_3}}{\sqrt{(R_{j_1}^2 + R_{j_3}^2)}} \frac{R_{j_2} R_{j_4}}{\sqrt{(R_{j_2}^2 + R_{j_4}^2)}} | R_{j_1}, R_{j_2} \right] + o_p(1) \right\} \\ &\quad \times (b_2 - r_2)^2 \left\{ \mathbb{E} \left[\frac{R_{j_1} R_{j_5}}{\sqrt{(R_{j_1}^2 + R_{j_5}^2)}} \frac{R_{j_2} R_{j_6}}{\sqrt{(R_{j_2}^2 + R_{j_6}^2)}} | R_{j_1}, R_{j_2} \right] + o_p(1) \right\} \\ &\xrightarrow{\mathcal{P}} 2[(r_1 \wedge r_2) - (a_1 \vee a_2)]^2 (b_1 - r_1)^2 (b_2 - r_2)^2 \mathbb{E}^2 \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right]. \end{aligned}$$

where the last equality holds by applying the law of large numbers for U -statistics to R_{j_3}, R_{j_4} and R_{j_5}, R_{j_6} , and the last holds by the law of large numbers of U -statistics to R_{j_1}, R_{j_2} .

Therefore, similar arguments for other terms indicate that

$$\begin{aligned} & 2^{-1} \mathbb{E}^{-2} \left[\frac{R_1 R_2}{\sqrt{(R_1^2 + R_3^2)(R_2^2 + R_3^2)}} \right] \lim_{n \rightarrow \infty} \text{Cov}(G_n^{(\mathcal{R}_{n,s})}(\lfloor nr_1 \rfloor; \lfloor na_1 \rfloor, \lfloor nb_1 \rfloor), G_n^{(\mathcal{R}_{n,s})}(\lfloor nr_2 \rfloor; \lfloor na_2 \rfloor, \lfloor nb_2 \rfloor)) \\ &= [(r_1 \wedge r_2) - (a_1 \vee a_2)]^2 (b_1 - r_1)^2 (b_2 - r_2)^2 \mathbf{1}((r_1 \wedge r_2) > (a_1 \vee a_2)) \\ &\quad + [(r_1 \wedge b_2) - (a_1 \vee r_2)]^2 (b_1 - r_1)^2 (r_2 - a_2)^2 \mathbf{1}((r_1 \wedge b_2) > (a_1 \vee r_2)) \\ &\quad - 4[r_2 - (a_1 \vee a_2)][(b_2 \wedge r_1) - r_2](b_1 - r_1)^2 (b_2 - r_2)(r_2 - a_2) \mathbf{1}(r_1 > r_2, r_2 > (a_1 \vee a_2), (b_2 \wedge r_1) > r_2) \\ &\quad + [(b_1 \wedge r_2) - (r_1 \vee a_2)]^2 (r_1 - a_1)^2 (b_2 - r_2)^2 \mathbf{1}((b_1 \wedge r_2) > (r_1 \vee a_2)) \\ &\quad + [(b_1 \wedge b_2) - (r_1 \vee r_2)]^2 (r_1 - a_1)^2 (r_2 - a_2)^2 \mathbf{1}((b_1 \wedge b_2) - (r_1 \vee r_2)) \\ &\quad - 4[r_2 - (r_1 \vee a_2)][(b_1 \wedge b_2) - r_2](r_2 - a_2)(b_2 - r_2)(r_1 - a_1)^2 \mathbf{1}(b_1 > r_2, r_2 > (r_1 \vee a_2), (b_1 \wedge b_2) > r_2) \\ &\quad - 4[r_1 - (a_1 \vee a_2)][(b_1 \wedge r_2) - r_1](r_1 - a_1)(b_1 - r_1)(b_2 - r_2)^2 \mathbf{1}(r_2 > r_1, r_1 > (a_1 \vee a_2), (b_1 \vee r_2) > r_1) \\ &\quad - 4[r_1 - (r_2 \vee a_1)][(b_1 \wedge b_2) - r_1](r_1 - a_1)(b_1 - r_1)(r_2 - a_2)^2 \mathbf{1}(b_2 > r_1, r_1 > (r_2 \vee a_1), (b_1 \wedge b_2) > r_1) \\ &\quad + 4[(r_1 \wedge r_2) - (a_1 \vee a_2)][(b_1 \wedge b_2) - (r_1 \wedge r_2)](r_1 - a_1)(b_1 - r_1)(r_2 - a_2)(b_2 - r_2) \\ &\quad \times \mathbf{1}((r_1 \wedge r_2) > (a_1 \vee a_2), (b_1 \wedge b_2) > (r_1 \wedge r_2)). \end{aligned}$$

This is indeed the covariance structure of $G(\cdot)$ after tedious algebra. $\square$

Appendix C. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jmva.2022.105114>.

References

- [1] A. Aue, L. Horváth, Structural breaks in time series, *J. Time Series Anal.* 34 (1) (2013) 1–16.
- [2] J.-M. Bardet, V. Brault, S. Dachian, F. Enikeeva, B. Sausseureau, Change-point detection, segmentation, and related topics, *ESAIM Proc. Surv.* 68 (2020) 97–122.
- [3] R.C. Bradley, Basic properties of strong mixing conditions, in: *Dependence in Probability and Statistics*, Springer, Birkhäuser, Boston, MA, 1986, pp. 165–192.
- [4] T.T. Cai, W. Liu, Y. Xia, Two-sample test of high dimensional means under dependence, *J. R. Stat. Soc. Ser. B Stat. Methodol.* (2014) 349–372.
- [5] A. Chakraborty, P. Chaudhuri, Tests for high-dimensional data based on means, spatial signs and spatial ranks, *Ann. Statist.* 45 (2) (2017) 771–799.
- [6] S.X. Chen, Y.-L. Qin, A two-sample test for high-dimensional data with applications to gene-set testing, *Ann. Statist.* 38 (2) (2010) 808–835.
- [7] H. Cho, Change-point detection in panel data via double CUSUM statistic, *Electron. J. Stat.* 10 (2) (2016) 2000–2038.
- [8] Z. Fan, L. Mackey, Empirical Bayesian analysis of simultaneous changepoints in multiple data sequences, *Ann. Appl. Stat.* 11 (4) (2017) 2200–2221.
- [9] P. Fryzlewicz, Wild binary segmentation for multiple change-point detection, *Ann. Statist.* 42 (6) (2014) 2243–2281.
- [10] B.M. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* 3 (5) (1975) 1163–1174.
- [11] L. Horváth, M. Hušková, Change-point detection in panel data, *J. Time Series Anal.* 33 (4) (2012) 631–648.
- [12] L. Horváth, G. Rice, Extensions of some classical methods in change point analysis, *Test* 23 (2) (2014) 219–255.
- [13] M. Jirak, Uniform change point tests in high dimension, *Ann. Statist.* 43 (6) (2015) 2451–2483.
- [14] S. Kovács, H. Li, P. Bühlmann, A. Munk, Seeded binary segmentation: A general methodology for fast and optimal change point detection, *Biometrika* (2022) (In Press).
- [15] Z. Lin, C. Lu, *Limit Theory for Mixing Dependent Random Variables*, Springer Science & Business Media, 2010.
- [16] H. Liu, C. Gao, R.J. Samworth, Minimax rates in sparse, high-dimensional change point detection, *Ann. Statist.* 49 (2) (2021) 1081–1112.
- [17] B. Liu, X. Zhang, Y. Liu, High dimensional change point inference: Recent developments and extensions, *J. Multivariate Anal.* 188 (2022) 104833.
- [18] B. Liu, C. Zhou, X. Zhang, Y. Liu, A unified data-adaptive framework for high dimensional change point detection, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 82 (4) (2020) 933–963.
- [19] J. Möttönen, H. Oja, J. Tienari, On the efficiency of multivariate spatial sign and rank tests, *Ann. Statist.* 25 (2) (1997) 542–552.
- [20] H. Oja, *Multivariate Nonparametric Methods with R: An Approach Based on Spatial Signs and Ranks*, Springer Science & Business Media, New York, N.Y., 2010.
- [21] P. Phillips, H. Moon, Linear regression limit theory for nonstationary panel data, *Econometrica* 67 (5) (1999) 1057–1111.
- [22] X. Shao, A self-normalized approach to confidence interval construction in time series, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 72 (3) (2010) 343–366.
- [23] X. Shao, Self-normalization for time series: a review of recent developments, *J. Amer. Statist. Assoc.* 110 (512) (2015) 1797–1817.
- [24] X. Shao, X. Zhang, Testing for change points in time series, *J. Amer. Statist. Assoc.* 105 (491) (2010) 1228–1240.
- [25] C. Truong, L. Oudre, N. Vayatis, Selective review of offline change point detection methods, *Signal Process.* 167 (2020).
- [26] L. Wang, B. Peng, R. Li, A high-dimensional nonparametric multivariate test for mean vector, *J. Amer. Statist. Assoc.* 110 (512) (2015) 1658–1669.
- [27] T. Wang, R.J. Samworth, High dimensional change point estimation via sparse projection, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 80 (1) (2018) 57–83.
- [28] R. Wang, C. Zhu, S. Volgushev, X. Shao, Inference for change points in high-dimensional data via selfnormalization, *Ann. Statist.* 50 (2) (2022) 781–806.
- [29] M. Yu, X. Chen, Finite sample change point inference and identification for high-dimensional mean vectors, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 83 (2) (2021) 247–270.
- [30] M. Yu, X. Chen, A robust bootstrap change point test for high-dimensional location parameter, *Electron. J. Stat.* 16 (1) (2022) 1096–1152.
- [31] Y. Zhang, R. Wang, X. Shao, Adaptive inference for change points in high-dimensional data, *J. Amer. Statist. Assoc.* (2022) (In Press).