A Data-Adaptive Integrated Approach to Covariance Change Point Detection in High-Dimensional Settings

Canhuang Xu[1], Lei Shu[2†], Yu Chen[3*], Qing Yang[1]

[1]*School of Management, University of Science and Technology of China*

[2]*School of Big Data and Statistics, Anhui University*

[3]*School of Public Affairs, University of Science and Technology of China*

Supplementary Material

This supplement contains some lemmas that we will need to use in our proofs, as well as the proofs of the asymptotic properties introduced in Section 3.

S1 Technical Lemma

Lemma 1 (Adapted from Vershynin (2018)). *Let $W_1, \ldots, W_p$ be mean-zero, sub-Gaussian random variables. There exist positive constants $c_1, c_2 > 0$, independent of the indices a and b, such that for any $1 \leq a, b \leq p$ and all $t > 0$,*

$$\mathbb{P}\Big\{|W_a W_b| > t\Big\} \leq c_1 e^{-c_2 t}.$$

[†]Canhuang Xu and Lei Shu are co-first authors

*Corresponding author: cyu@ustc.edu.cn

Furthermore, if these random variables are independent, there exist positive constants $c_1, c_2 > 0$ *such that*

$$\mathbb{P}\{|\frac{1}{p}\sum_{i=1}^{p} W_i| > t\} \le c_1 e^{-c_2 p t^2}.$$

The first inequality follows from the fact that each product $W_a W_b$ *is sub-exponential (Lemma 2.7.7 in Vershynin (2018)), while the second is a direct application of the concentration bound for sub-Gaussian variables (Theorem 2.6.3 in Vershynin (2018)).*

Alternatively, if $W_1, \ldots, W_p$ *are independent, mean-zero, sub-exponential random variables, then, adapted from Corollary 2.8.3 in Vershynin (2018), there exist positive constants* $c_1, c_2 > 0$ *such that for all* $t > 0$,

$$\mathbb{P}\{|\frac{1}{p}\sum_{i=1}^{p} W_i| > t\} \le c_1 e^{-c_2 p \cdot \min\{t^2, t\}}.$$

S2 Proof of Proposition 1

Denote $\mu_a = \mathbb{E}X_{ia}$, $\bar{X}_a = \sum_{i=1}^n X_{ia}$, $\dot{X}_{ia} = X_{ia} - \bar{X}_a$, $\bar{X}_{k,h} = \frac{1}{k}\sum_{i=1}^k X_{ia}X_{ib}$, $\bar{Y}_{k,h} = \frac{1}{n-k}\sum_{i=k+1}^n X_{ia}X_{ib}$, $\dot{\bar{X}}_{k,h} = \frac{1}{k}\sum_{i=1}^k \dot{X}_{ia}\dot{X}_{ib}$, and $\dot{\bar{Y}}_{k,h} = \frac{1}{n-k}\sum_{i=k+1}^n \dot{X}_{ia}\dot{X}_{ib}$. Note that $\dot{X}_{ia} = X_{ia} - \mu_a + \mu_a - \bar{X}_a$. To streamline notation in subsequent proofs, we use X_{ia} to denote $X_{ia} - \mu_a$ and $\bar{X}_a$ to denote $\bar{X}_a - \mu_a$. Thus,

$$\dot{X}_{ia}\dot{X}_{ib} = X_{ia}X_{ib} - X_{ia}\bar{X}_b - \bar{X}_a X_{ib} + \bar{X}_a\bar{X}_b. \tag{S2.1}$$

We decompose $v_{k,h}$ as $v_{k,h} = v_{k1,h} + v_{k2,h} - v_{k3,h}$, where

$$v_{k1,h} = \big[\dot{\bar{X}}_{k,h} - \dot{\bar{Y}}_{k,h}\big]^2, \quad v_{k2,h} = \big[\frac{1}{k-1}\dot{\bar{X}}_{k,h}^2 + \frac{1}{n-k-1}\dot{\bar{Y}}_{k,h}^2\big],$$
$$v_{k3,h} = \big[\frac{1}{(k)_2}\sum_{i=1}^{k}(\dot{X}_{ia}\dot{X}_{ib})^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}(\dot{X}_{ia}\dot{X}_{ib})^2\big].$$

Then, d_h can be decomposed as follows:

$$\begin{aligned}
d_h &= \frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}v_{k,h} \\
&= \frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}v_{k1,h} + \frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}v_{k2,h} - \frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}v_{k3,h} \\
&= d_{1,h} + d_{2,h} - d_{3,h}.
\end{aligned}$$

For Proposition 1(a): Considering the positions (a,b) satisfying $\Sigma_{1,ab} = \Sigma_{2,ab} = \Sigma_{ab}$. First, it is evident that $v_{k_1,h} = \big[\dot{\bar{X}}_{k,h} - \dot{\bar{Y}}_{k,h}\big]^2 = \big[(\dot{\bar{X}}_{k,h} - \Sigma_{ab}) - (\dot{\bar{Y}}_{k,h} - \Sigma_{ab})\big]^2$. Then we calculate

$$\begin{aligned}
&\Big[\frac{1}{k-1}(\dot{\bar{X}}_{k,h} - \Sigma_{ab})^2 + \frac{1}{n-k-1}(\dot{\bar{Y}}_{k,h} - \Sigma_{ab})^2\Big] \\
&\quad - \Big[\frac{1}{(k)_2}\sum_{i=1}^{k}(\dot{X}_{ia}\dot{X}_{ib} - \Sigma_{ab})^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}(\dot{X}_{ia}\dot{X}_{ib} - \Sigma_{ab})^2\Big] \\
=&\Big[\frac{1}{k-1}\dot{\bar{X}}_{k,h}^2 + \frac{1}{n-k-1}\dot{\bar{Y}}_{k,h}^2 - \frac{2}{k-1}\Sigma_{ab}\dot{\bar{X}}_{k,h} - \frac{2}{n-k-1}\Sigma_{ab}\dot{\bar{Y}}_{k,h} + (\frac{1}{k-1} + \frac{1}{n-k-1})\Sigma_{ab}^2\Big] \\
&\quad - \Big[\frac{1}{(k)_2}\sum_{i=1}^{k}(\dot{X}_{ia}\dot{X}_{ib})^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}(\dot{X}_{ia}\dot{X}_{ib})^2 - \frac{2}{(k)_2}\Sigma_{ab}\sum_{i=1}^{k}\dot{X}_{ia}\dot{X}_{ib} \\
&\qquad - \frac{2}{(n-k)_2}\Sigma_{ab}\sum_{i=k+1}^{n}\dot{X}_{ia}\dot{X}_{ib} + (\frac{1}{k-1} + \frac{1}{n-k-1})\Sigma_{ab}^2\Big].
\end{aligned}$$

Recall that $\dot{\bar{X}}_{k,h} = \frac{1}{k}\sum_{i=1}^{k}\dot{X}_{ia}\dot{X}_{ib}$, and $\dot{\bar{Y}}_{k,h} = \frac{1}{n-k}\sum_{i=k+1}^{n}\dot{X}_{ia}\dot{X}_{ib}$, thus we have$\big[\frac{1}{k-1}(\dot{\bar{X}}_{k,h} - \Sigma_{ab})^2 + \frac{1}{n-k-1}(\dot{\bar{Y}}_{k,h} - \Sigma_{ab})^2\big] - \big[\frac{1}{(k)_2}\sum_{i=1}^{k}(\dot{X}_{ia}\dot{X}_{ib} - \Sigma_{ab})^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}(\dot{X}_{ia}\dot{X}_{ib} - \Sigma_{ab})^2\big] = v_{k2,h} - v_{k3,h}$. Summarizing we have

$$\begin{aligned}
v_{k,h} =&\big[(\dot{\bar{X}}_{k,h} - \Sigma_{1,ab}) - (\dot{\bar{Y}}_{k,h} - \Sigma_{2,ab})\big]^2 + \Big[\frac{1}{k-1}(\dot{\bar{X}}_{k,h} - \Sigma_{1,ab})^2 + \frac{1}{n-k-1}(\dot{\bar{Y}}_{k,h} - \Sigma_{2,ab})^2\Big] \\
&- \Big[\frac{1}{(k)_2}\sum_{i=1}^{k}(\dot{X}_{ia}\dot{X}_{ib} - \Sigma_{1,ab})^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}(\dot{X}_{ia}\dot{X}_{ib} - \Sigma_{2,ab})^2\Big].
\end{aligned}$$

Hence we can assume that $\Sigma_{1,ab} = \Sigma_{2,ab} = 0$ in this proof, that is, $X_{ia}X_{ib}$ is a zero-mean sub-exponential random variable for any $1 \le a, b \le p$.

Observing the inclusion $\{d_h > \tau\} \subset \{d_{1,h} > \frac{1}{2}\tau\} \cup \{d_{2,h} > \frac{1}{2}\tau\}$, it's sufficient to calculate $\mathbb{P}\big\{d_{1,h} > \frac{1}{2}\tau\big\}$ and $\mathbb{P}\big\{d_{2,h} > \frac{1}{2}\tau\big\}$. Hereafter in the proof, c and $c_i(i = 1, 2, \dots)$ indicate some positive constants that may change from line to line. The first term can be handled as follows:

$$\begin{aligned}\mathbb{P}\Big\{d_{1,h} > \frac{1}{2}\tau\Big\} &= \mathbb{P}\Big\{\frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big[\dot{\bar{X}}_{k,h} - \dot{\bar{Y}}_{k,h}\big]^2 > \frac{1}{2}\tau\Big\}\\ &\le \sum_{k=2}^{n-2}\mathbb{P}\Big\{\frac{k(n-k)}{n}\big[\dot{\bar{X}}_{k,h} - \dot{\bar{Y}}_{k,h}\big]^2 > \frac{1}{2}\tau\Big\}\\ &\le \sum_{k=2}^{n-2}\Big[\mathbb{P}\Big\{\big|\dot{\bar{X}}_{k,h}\big| > \frac{1}{2}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\} + \mathbb{P}\Big\{\big|\dot{\bar{Y}}_{k,h}\big| > \frac{1}{2}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\}\Big].\end{aligned}$$

Consider the components in (S2.1). We use Lemma 1 to calculate the following probabilities

$$\mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k}X_{ia}X_{ib}\big| > \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\} \le c_1 e^{-c_2 k \min\{\frac{n\tau}{128k(n-k)}, \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\}}, \tag{S2.2}$$

Similarly, it follows from Lemma 1

$$\begin{aligned}&\mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k}X_{ia}\bar{X}_b\big| > \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\}\\ \le&\mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k}X_{ia}\big| > \big(\frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\big)^{\frac{1}{2}}\Big\} + \mathbb{P}\Big\{\big|\frac{1}{n}\sum_{i=1}^{n}X_{ib}\big| > \big(\frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\big)^{\frac{1}{2}}\Big\}\\ \le&c_1 e^{-c_2 k \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}} + c_3 e^{-c_4 n \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}}.\end{aligned} \tag{S2.3}$$

Using similar arguments, we obtain that

$$\begin{aligned}&\mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k}\bar{X}_a X_{ib}\big| > \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\} \le c_1 e^{-c_2 k \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}} + c_3 e^{-c_4 n \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}},\\ &\mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k}\bar{X}_a\bar{X}_b\big| > \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\} \le c_1 e^{-c_2 n \frac{1}{8}\sqrt{\frac{n\tau}{2k(n-k)}}}.\end{aligned} \tag{S2.4}$$

Combining (S2.2)-(S2.4), we obtain the upper bound

$$\mathbb{P}\Big\{\big|\dot{\bar{X}}_{k,h}\big| > \frac{1}{2}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\} \le c_1 e^{-c_2 k \min\{\frac{n\tau}{16k(n-k)}, \sqrt{\frac{n\tau}{2k(n-k)}}\}} + c_3 e^{-c_4 k\sqrt{\frac{n\tau}{2k(n-k)}}} + c_5 e^{-c_6 n\sqrt{\frac{n\tau}{2k(n-k)}}},$$

and similarly for $\dot{\bar{Y}}_{k,h}$

$$\mathbb{P}\Big\{\big|\dot{\bar{Y}}_{k,h}\big| > \frac{1}{2}\sqrt{\frac{n\tau}{2k(n-k)}}\Big\} \leq c_1 e^{-c_2(n-k)\min\{\frac{n\tau}{16k(n-k)},\sqrt{\frac{n\tau}{2k(n-k)}}\}} + c_3 e^{-c_4(n-k)\sqrt{\frac{n\tau}{2k(n-k)}}} + c_5 e^{-c_6 n\sqrt{\frac{n\tau}{2k(n-k)}}}.$$

Therefore, we have $\mathbb{P}\{d_{1,h} > \frac{1}{2}\tau\} \leq c_1 n[e^{-c_2\sqrt{\tau}} + e^{-c_3\sqrt{n\tau}}]$.

Similarly, the second term can be estimated as follows:

$$\begin{aligned}
\mathbb{P}\Big\{d_{2,h} > \frac{1}{2}\tau\Big\} &= \mathbb{P}\Big\{\frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big[\frac{1}{k-1}\dot{\bar{X}}^2_{k,h} + \frac{1}{n-k-1}\dot{\bar{Y}}^2_{k,h}\big] > \frac{1}{2}\tau\Big\} \\
&\leq \sum_{k=2}^{n-2}\mathbb{P}\Big\{\frac{k(n-k)}{n}\big[\frac{1}{k-1}\dot{\bar{X}}^2_{k,h} + \frac{1}{n-k-1}\dot{\bar{Y}}^2_{k,h}\big] > \frac{1}{2}\tau\Big\} \\
&\leq \sum_{k=2}^{n-2}\Big[\mathbb{P}\Big\{\big|\dot{\bar{X}}_{k,h}\big| > \sqrt{\frac{n(k-1)\tau}{4k(n-k)}}\Big\} + \mathbb{P}\Big\{\big|\dot{\bar{Y}}_{k,h}\big| > \sqrt{\frac{n(n-k-1)\tau}{4k(n-k)}}\Big\}\Big] \\
&\leq c_1 n\big[e^{-c_2\sqrt{\tau}} + e^{-c_3 n\sqrt{\tau}}\big].
\end{aligned}$$

Summarizing we have $p^2 \cdot \mathbb{P}\{d_h > \tau\} \leq c_1 p^2 n[e^{-c_2\sqrt{\tau}} + e^{-c_3\sqrt{n\tau}} + e^{-c_4 n\sqrt{\tau}}]$. In particular, if $p^2 n = o(e^{c\sqrt{\tau}})$ holds we have $p^2 \cdot \mathbb{P}\{d_h > \tau\} \to 0$.

For Proposition 1(b): Considering the positions (a,b) satisfying $\big|\Sigma_{1,ab} - \Sigma_{2,ab}\big| \geq c\sqrt{\frac{\tau}{n}}$ for some constant c. Observing the inclusion $\{d_h \leq \tau\} \subset \{d_{1,h} \leq 3\tau\} \cup \{d_{3,h} \geq 2\tau\}$, one can see that it is sufficient to calculate $\mathbb{P}\big\{d_{1,h} \leq 3\tau\big\}$ and $\mathbb{P}\big\{d_{3,h} \geq 2\tau\big\}$. The second term is estimated as follows:

$$\begin{aligned}
\mathbb{P}\Big\{d_{3,h} > 2\tau\Big\} &= \mathbb{P}\Big\{\frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big[\frac{1}{(k)_2}\sum_{i=1}^{k}\big(\dot{X}_{ia}\dot{X}_{ib}\big)^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}\big(\dot{X}_{ia}\dot{X}_{ib}\big)^2\big] \geq 2\tau\Big\} \\
&\leq \sum_{k=2}^{n-2}\mathbb{P}\Big\{\frac{k(n-k)}{n}\big[\frac{1}{(k)_2}\sum_{i=1}^{k}\big(\dot{X}_{ia}\dot{X}_{ib}\big)^2 + \frac{1}{(n-k)_2}\sum_{i=k+1}^{n}\big(\dot{X}_{ia}\dot{X}_{ib}\big)^2\big] \geq 2\tau\Big\} \\
&\leq \sum_{k=2}^{n-2}\Big[\mathbb{P}\Big\{\frac{1}{k}\sum_{i=1}^{k}\big(\dot{X}_{ia}\dot{X}_{ib}\big)^2 > \frac{n(k-1)\tau}{k(n-k)}\Big\} + \mathbb{P}\Big\{\frac{1}{n-k}\sum_{i=k+1}^{n}\big(\dot{X}_{ia}\dot{X}_{ib}\big)^2 > \frac{n(n-k-1)\tau}{k(n-k)}\Big\}\Big] \\
&\leq \sum_{k=2}^{n-2}\Big[k\mathbb{P}\Big\{\big|\dot{X}_{ia}\dot{X}_{ib}\big| > \sqrt{\frac{n(k-1)\tau}{k(n-k)}}\Big\} + (n-k)\mathbb{P}\Big\{\big|\dot{X}_{ia}\dot{X}_{ib}\big| > \sqrt{\frac{n(n-k-1)\tau}{k(n-k)}}\Big\}\Big].
\end{aligned}$$

Consider the components in (S2.1). We use Lemma 1 to calculate the following probability

$$\mathbb{P}\Big\{\big|X_{ia}X_{ib}\big| > \frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}\Big\} \leq c_1 e^{-c_2\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}}. \tag{S2.5}$$

Similarly, it follows from Assumption 1(a) and Lemma 1

$$\begin{aligned}\mathbb{P}\Big\{\big|X_{ia}\bar{X}_b\big| > \frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}\Big\} \leq & \mathbb{P}\Big\{\big|X_{ia}\big| > (\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}})^{\frac{1}{2}}\Big\} + \mathbb{P}\Big\{\big|\bar{X}_b\big| > (\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}})^{\frac{1}{2}}\Big\} \\ \leq & c_1 e^{-c_2\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}} + c_3 e^{-c_4 n\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}}.\end{aligned} \tag{S2.6}$$

And using similar arguments, we obtain for

$$\begin{aligned}&\mathbb{P}\Big\{\big|X_{ib}\bar{X}_a\big| > \frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}\Big\} \leq c_1 e^{-c_2\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}} + c_3 e^{-c_4 n\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}}, \\ &\mathbb{P}\Big\{\big|\bar{X}_a\bar{X}_b\big| > \frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}\Big\} \leq c_1 e^{-c_2 n\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}}.\end{aligned} \tag{S2.7}$$

Combining (S2.5)-(S2.7), we obtain the upper bound

$$\begin{aligned}&\mathbb{P}\Big\{\big|\dot{X}_{ia}\dot{X}_{ib}\big| > \sqrt{\frac{n(k-1)\tau}{k(n-k)}}\Big\} \leq c_1 e^{-c_2\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}} + c_3 e^{-c_4 n\frac{1}{4}\sqrt{\frac{n(k-1)\tau}{k(n-k)}}}, \\ &\mathbb{P}\Big\{\big|\dot{X}_{ia}\dot{X}_{ib}\big| > \sqrt{\frac{n(n-k-1)\tau}{k(n-k)}}\Big\} \leq c_1 e^{-c_2\frac{1}{4}\sqrt{\frac{n(n-k-1)\tau}{k(n-k)}}} + c_3 e^{-c_4 n\frac{1}{4}\sqrt{\frac{n(n-k-1)\tau}{k(n-k)}}}.\end{aligned}$$

Therefore, we have

$$\mathbb{P}\big\{d_{3,h} > 2\tau\big\} \leq c_1 n^2\big[e^{-c_2\sqrt{\tau}} + e^{-c_3 n\sqrt{\tau}}\big].$$

Then we estimate the term $\mathbb{P}\big\{d_{1,h} \leq 3\tau\big\}$. A straightforward calculation gives

$$\mathbb{E}\big(\bar{X}_{k,h} - \bar{Y}_{k,h}\big) = \begin{cases} \frac{n-k_0}{n-k}(\Sigma_{1,ab} - \Sigma_{2,ab}), & k \leq k_0, \\ \frac{k_0}{k}(\Sigma_{1,ab} - \Sigma_{2,ab}), & k > k_0. \end{cases}$$

Let $A_{k,h} = \dot{\bar{X}}_{k,h} - \dot{\bar{Y}}_{k,h} - \mathbb{E}\big(\bar{X}_{k,h} - \bar{Y}_{k,h}\big)$, $B_{k,h} = \mathbb{E}\big(\bar{X}_{k,h} - \bar{Y}_{k,h}\big)$. One can observe that

$$|B_{k,h}| \geq \min\big\{\frac{n-k_0}{n}, \frac{k_0}{n}\big\} \cdot \big|\Sigma_{1,h} - \Sigma_{2,ab}\big| = c_0\big|\Sigma_{1,ab} - \Sigma_{2,ab}\big|$$

where $c_0 = \min\{\frac{k_0}{n}, \frac{n-k_0}{n}\}$ and

$$\frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big(B_{k,h}\big)^2 \geq c_0^2\cdot n(\Sigma_{1,ab}-\Sigma_{2,ab})^2\cdot\frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n^2} \geq c_0^2\cdot\frac{n(\Sigma_{1,ab}-\Sigma_{2,ab})^2}{6}.$$

Consequently, by Assumption 1(b) (with a sufficiently large constant c) we obtain the estimate

$$3\tau \leq \frac{1}{2(n-3)}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big(B_{k,h}\big)^2.$$

Then we can obtain that

$$\begin{aligned}
\mathbb{P}\{d_{1,h}\leq 3\tau\} &= \mathbb{P}\Big\{\frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big[A_{k,h}+B_{k,h}\big]^2 \leq 3\tau\Big\}\\
&\leq \mathbb{P}\Big\{\frac{2}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}A_{k,h}B_{k,h} \leq 3\tau - \frac{1}{n-3}\sum_{k=2}^{n-2}\frac{k(n-k)}{n}\big(B_{k,h}\big)^2\Big\}\\
&\leq \mathbb{P}\Big\{\frac{2}{n-3}\Big|\sum_{k=2}^{n-2}\frac{k(n-k)}{n}A_{k,h}B_{k,h}\Big| \geq \frac{1}{2(n-3)}\Big|\sum_{k=2}^{n-2}\frac{k(n-k)}{n}(B_{k,h})^2\Big|\Big\}\\
&\leq n\cdot\mathbb{P}\Big\{\Big|\frac{k(n-k)}{n}A_{k,h}\Big| \geq \frac{c}{|B_{k,h}|}\cdot n(\Sigma_{1,ab}-\Sigma_{2,ab})^2\Big\} = n\cdot\mathbb{P}\big\{|A_{k,h}|\geq c\phi\big\},
\end{aligned}$$

where $\phi = \frac{1}{|B_{k,h}|}\frac{n^2}{k(n-k)}(\Sigma_{1,ab}-\Sigma_{2,ab})^2$. Since $|B_{k,h}| \leq |\Sigma_{1,ab}-\Sigma_{2,ab}|$, under Assumption 1(b) we have

$$\phi \geq \frac{n^2}{k(n-k)}\big|\Sigma_{1,ab}-\Sigma_{2,ab}\big| \geq C\frac{n}{k(n-k)}\sqrt{n\tau}.$$

Observing the decomposition (S2.1), the term $A_{k,h}$ can be written as

$$\begin{aligned}
A_{k,h} = \big(\bar{X}_{k,h}-\mathbb{E}\bar{X}_{k,h}\big) - \big(\bar{Y}_{k,h}-\mathbb{E}\bar{Y}_{k,h}\big) - \frac{1}{k}\sum_{i=1}^{k}X_{ia}\bar{X}_b - \frac{1}{k}\sum_{i=1}^{k}X_{ib}\bar{X}_a\\
+\frac{1}{n-k}\sum_{i=1}^{n-k}X_{ia}\bar{X}_b + \frac{1}{n-k}\sum_{i=1}^{n-k}X_{ib}\bar{X}_a.
\end{aligned}$$

Considering the components of this decomposition, we use Lemma 1 to calculate the following probabilities

$$\mathbb{P}\Big\{\big|\bar{X}_{k,h}-\mathbb{E}\bar{X}_{k,h}\big| > \frac{c}{6}\phi\Big\} \leq c_1 e^{-c_2 k\min\{\phi^2,\phi\}}, \quad \mathbb{P}\Big\{\big|\bar{Y}_{k,h}-\mathbb{E}\bar{Y}_{k,h}\big| > \frac{c}{6}\phi\Big\} \leq c_1 e^{-c_2(n-k)\min\{\phi^2,\phi\}}. \tag{S2.8}$$

Similarly, we can obtain that

$$\mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k} X_{ia}\bar{X}_b\big| > \frac{c}{6}\phi\Big\} \le c_1 e^{-c_2 k\phi} + c_3 e^{-c_4 n\phi}, \quad \mathbb{P}\Big\{\big|\frac{1}{k}\sum_{i=1}^{k} X_{ib}\bar{X}_a\big| > \frac{c}{6}\phi\Big\} \le c_1 e^{-c_2 k\phi} + c_3 e^{-c_4 n\phi},$$
$$\mathbb{P}\Big\{\big|\frac{1}{n-k}\sum_{i=1}^{n-k} X_{ia}\bar{X}_b\big| > \frac{c}{6}\phi\Big\} \le c_1 e^{-c_2 (n-k)\phi} + c_3 e^{-c_4 n\phi},$$
$$\mathbb{P}\Big\{\big|\frac{1}{n-k}\sum_{i=1}^{n-k} X_{ib}\bar{X}_a\big| > \frac{c}{6}\phi\Big\} \le c_1 e^{-c_2 (n-k)\phi} + c_3 e^{-c_4 n\phi}. \tag{S2.9}$$

Combining (S2.8) and (S2.9), we can obtain that

$$\begin{aligned}\mathbb{P}\{d_{1,h} \le 3\tau\} &\le n\cdot\mathbb{P}\big\{|A_{k,h}| \ge c\phi\big\} \\ &\le c_1 n\big[e^{-c_2 k\min\{\phi^2,\phi\}} + e^{-c_3 (n-k)\min\{\phi^2,\phi\}} + e^{-c_4 k\phi} + e^{-c_5 (n-k)\phi} + e^{-c_6 n\phi}\big] \\ &\le c_1 n\big[e^{-c_2 \min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\big].\end{aligned}$$

Summarizing we have $p^2\cdot\mathbb{P}\big\{d_h \le \tau\big\} \le c_1 p^2 n\big[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}} + ne^{-c_4\sqrt{\tau}} + ne^{-c_5 n\sqrt{\tau}}\big]$. In particular, if $p^2 n^2 = o(e^{c\sqrt{\tau}})$ holds we have $p^2\cdot\mathbb{P}\big\{d_h \le \tau\big\} \to 0$.

S3 Proof of Theorem 1

Recalling the definition of the statistic $T_{\boldsymbol{W}}(k)$ given in (1) and the change point estimator $\hat{k}_0$ defined in (3), it follows that

$$\mathbb{P}\big\{\big|\frac{\hat{k}_0}{k_0} - 1\big| \ge \epsilon\big\} \le \mathbb{P}\big\{\hat{k}_0 \ge (1+\epsilon)k_0\big\} + \mathbb{P}\big\{\hat{k}_0 \le (1-\epsilon)k_0\big\}.$$

To establish that this probability converges to zero asymptotically, we bound the two terms on the right-hand side. For brevity, we focus on the first term, as the second can be analyzed analogously. Specifically, it suffices to show that

$$\mathbb{P}\Big\{\bigcup_{k\geq(1+\epsilon)k_0}\big(T_{\boldsymbol{W}}(k)\geq T_{\boldsymbol{W}}(k_0)\big)\Big\}\leq\sum_{k\geq(1+\epsilon)k_0}\mathbb{P}\big\{T_{\boldsymbol{W}}(k)\geq T_{\boldsymbol{W}}(k_0)\big\}.$$

Denote $\tilde{\boldsymbol{\sigma}}_{(i)}=\dot{\boldsymbol{\sigma}}_{(i)}-\mathbb{E}\boldsymbol{\sigma}_{(i)}=(\tilde{\sigma}_{ij})_{j=1,\ldots,\frac{p(p+1)}{2}}$ and $\boldsymbol{\sigma}=\mathrm{vech}(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2)=\big(\sigma_h\big)_{1\leq h\leq\frac{p(p+1)}{2}}$, where $\boldsymbol{\sigma}_{(i)}=\mathrm{vech}\big((\boldsymbol{X}_i-\boldsymbol{\mu})(\boldsymbol{X}_i-\boldsymbol{\mu})^\top\big)$. We will make use of the decomposition

$$\begin{aligned}T_{\boldsymbol{W}}(k)&=\frac{1}{(k)_2(n-k)_2}\sum_{(i\neq t)=1}^{k}\sum^{k}\sum_{(j\neq l)=k+1}^{n}\sum^{n}\Big[\big(\tilde{\boldsymbol{\sigma}}_{(i)}-\tilde{\boldsymbol{\sigma}}_{(j)}\big)+\big(\mathbb{E}\boldsymbol{\sigma}_{(i)}-\mathbb{E}\boldsymbol{\sigma}_{(j)}\big)\Big]^T\boldsymbol{W}\Big[\big(\tilde{\boldsymbol{\sigma}}_{(t)}-\tilde{\boldsymbol{\sigma}}_{(l)}\big)+\big(\mathbb{E}\boldsymbol{\sigma}_{(t)}-\mathbb{E}\boldsymbol{\sigma}_{(l)}\big)\Big]\\&=A(k)+B(k)+C(k)+D(k),\end{aligned}$$

where $\boldsymbol{W}=\boldsymbol{\Gamma}+\boldsymbol{\delta}\boldsymbol{\delta}^\top$, $\boldsymbol{\Gamma}=\big(\gamma_{(a,b)}\big)\in\mathbb{R}^{\frac{p(p+1)}{2}\times\frac{p(p+1)}{2}}$, $\boldsymbol{\delta}=\big(\delta_{(a,b)}\big)\in\mathbb{R}^{\frac{p(p+1)}{2}}$ and

$$\begin{aligned}A(k)&=\frac{1}{(k)_2(n-k)_2}\sum_{(i\neq t)=1}^{k}\sum^{k}\sum_{(j\neq l)=k+1}^{n}\sum^{n}\big(\tilde{\boldsymbol{\sigma}}_{(i)}-\tilde{\boldsymbol{\sigma}}_{(j)}\big)^\top\boldsymbol{W}\big(\tilde{\boldsymbol{\sigma}}_{(t)}-\tilde{\boldsymbol{\sigma}}_{(l)}\big),\\B(k)&=\frac{1}{(k)_2(n-k)_2}\sum_{(i\neq t)=1}^{k}\sum^{k}\sum_{(j\neq l)=k+1}^{n}\sum^{n}\big(\mathbb{E}\boldsymbol{\sigma}_{(t)}-\mathbb{E}\boldsymbol{\sigma}_{(l)}\big)^\top\boldsymbol{W}\big(\tilde{\boldsymbol{\sigma}}_{(i)}-\tilde{\boldsymbol{\sigma}}_{(j)}\big),\\C(k)&=\frac{1}{(k)_2(n-k)_2}\sum_{(i\neq t)=1}^{k}\sum^{k}\sum_{(j\neq l)=k+1}^{n}\sum^{n}\big(\mathbb{E}\boldsymbol{\sigma}_{(i)}-\mathbb{E}\boldsymbol{\sigma}_{(j)}\big)^\top\boldsymbol{W}\big(\tilde{\boldsymbol{\sigma}}_{(t)}-\tilde{\boldsymbol{\sigma}}_{(l)}\big),\\D(k)&=\frac{1}{(k)_2(n-k)_2}\sum_{(i\neq t)=1}^{k}\sum^{k}\sum_{(j\neq l)=k+1}^{n}\sum^{n}\big(\mathbb{E}\boldsymbol{\sigma}_{(i)}-\mathbb{E}\boldsymbol{\sigma}_{(j)}\big)^\top\boldsymbol{W}\big(\mathbb{E}\boldsymbol{\sigma}_{(t)}-\mathbb{E}\boldsymbol{\sigma}_{(l)}\big).\end{aligned}$$

Starting to study the constant terms $D(k)$ and $D(k_0)$. Notice that

$$\mathbb{E}\boldsymbol{\sigma}_{(i)}=\begin{cases}\mathrm{vech}(\boldsymbol{\Sigma}_1),\quad i\leq k_0;\\[2ex]\mathrm{vech}(\boldsymbol{\Sigma}_2),\quad i>k_0.\end{cases}$$

We calculate that

$$\begin{aligned}
D = D(k_0) - D(k) =& \frac{1}{(k_0)_2(n-k_0)_2} \sum_{(i\neq t)=1}^{k_0}\sum^{k_0} \sum_{(j\neq l)=k_0+1}^{n}\sum^{n} \big(\mathbb{E}\boldsymbol{\sigma}_{(i)} - \mathbb{E}\boldsymbol{\sigma}_{(j)}\big)^\top \boldsymbol{W} \big(\mathbb{E}\boldsymbol{\sigma}_{(t)} - \mathbb{E}\boldsymbol{\sigma}_{(l)}\big) \\
& - \frac{1}{(k)_2(n-k)_2} \sum_{(i\neq t)=1}^{k}\sum^{k} \sum_{(j\neq l)=k+1}^{n}\sum^{n} \big(\mathbb{E}\boldsymbol{\sigma}_{(i)} - \mathbb{E}\boldsymbol{\sigma}_{(j)}\big)^\top \boldsymbol{W} \big(\mathbb{E}\boldsymbol{\sigma}_{(t)} - \mathbb{E}\boldsymbol{\sigma}_{(l)}\big) \\
=& \boldsymbol{\sigma}^\top \boldsymbol{W}\boldsymbol{\sigma} - \frac{1}{(k)_2(n-k)_2} \sum_{(i\neq t)=1}^{k_0}\sum^{k_0} \sum_{(j\neq l)=k+1}^{n}\sum^{n} \boldsymbol{\sigma}^\top \boldsymbol{W}\boldsymbol{\sigma} \\
=& \big(1 - \frac{(k_0)_2}{(k)_2}\big)\boldsymbol{\sigma}^\top \boldsymbol{W}\boldsymbol{\sigma}.
\end{aligned}$$

Observing the inclusion

$$\big\{T_{\boldsymbol{W}}(k) \geq T_{\boldsymbol{W}}(k_0)\big\} \subset \big\{A(k) + B(k) + C(k) - \big(A(k_0) + B(k_0) + C(k_0)\big) \geq D\big\},$$

we now investigate the other terms $A(k) - A(k_0), B(k) - B(k_0), C(k) - C(k_0)$. By calculating this, we can obtain the decomposition

$$B(k) - B(k_0) = B_1 + B_2 + B_3$$

where

$$\begin{aligned}
& B_1 = \big(\frac{k_0 - 1}{k(k-1)} - \frac{1}{k_0}\big) \sum_{i=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W}\boldsymbol{\sigma}, \quad B_2 = -\big(\frac{k_0}{k(n-k)} - \frac{1}{n-k_0}\big) \sum_{i=k_0+1}^{n} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W}\boldsymbol{\sigma}, \\
& B_3 = \big(\frac{k_0}{k(k-1)} + \frac{k_0}{k(n-k)}\big) \sum_{i=k_0+1}^{k} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W}\boldsymbol{\sigma}.
\end{aligned}$$

The term B_1 can be handled as follows

$$\begin{aligned}
\mathbb{P}\Big\{|B_1| \geq cD\Big\} &= \mathbb{P}\Big\{\big|\sum_{i=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W}\boldsymbol{\sigma}\big| \geq ck_0 \boldsymbol{\sigma}^\top \boldsymbol{W}\boldsymbol{\sigma}\Big\} \\
&\leq \mathbb{P}\Big\{\big\|\boldsymbol{W}^{\frac{1}{2}} \sum_{i=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}\big\|_2 \big\|\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\sigma}\big\|_2 \geq ck_0 \big\|\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\sigma}\big\|_2^2\Big\} \\
&\leq \mathbb{P}\Big\{\sum_{j\in J_1} \gamma_j \big(\sum_{i=1}^{k_0} \tilde{\sigma}_{ij}\big)^2 \geq ck_0^2 \sum_{j\in J_1} \gamma_j \sigma_j^2\Big\} + \mathbb{P}\Big\{\big|\sum_{j\in J_2}\sum_{i=1}^{k_0} \delta_j \tilde{\sigma}_{ij}\big| \geq ck_0 \big|\sum_{j\in J_2} \delta_j \sigma_j\big|\Big\},
\end{aligned}$$

where J_1 and J_2 represent the sets of subscripts for which γ_h and δ_h are non-zero, respectively. Note that for each j there exist a position (a,b) such that $\tilde{\sigma}_{ij}$ can be written as

$$\tilde{\sigma}_{ij} = \big(X_{ia}X_{ib} - \mathbb{E}X_{ia}X_{ib}\big) - X_{ia}\bar{X}_b - X_{ib}\bar{X}_a + \bar{X}_a\bar{X}_b.$$

Consider this decomposition. We can use Lemma 1 to handle the first term as follows:

$$\begin{aligned}
&\mathbb{P}\Big\{\sum_{j\in J_1}\gamma_j\big(\sum_{i=1}^{k_0}\tilde{\sigma}_{ij}\big)^2 \geq ck_0^2\sum_{j\in J_1}\gamma_j\sigma_j^2\Big\}\\
\leq&\sum_{j\in J_1}\mathbb{P}\Big\{\big|\frac{1}{k_0}\sum_{i=1}^{k_0}\tilde{\sigma}_{ij}\big| \geq c|\sigma_j|\Big\}\\
\leq&\mathbb{P}\Big\{\big|\frac{1}{k_0}\sum_{i=1}^{k_0}\big(X_{ia}X_{ib} - \mathbb{E}X_{ia}X_{ib}\big)\big| \geq c|\sigma_j|\Big\} + \mathbb{P}\Big\{\big|\frac{1}{k_0}\sum_{i=1}^{k_0}X_{ia}\bar{X}_b\big| \geq c|\sigma_j|\Big\}\\
&+\mathbb{P}\Big\{\big|\frac{1}{k_0}\sum_{i=1}^{k_0}X_{ib}\bar{X}_a\big| \geq c|\sigma_j|\Big\} + \mathbb{P}\Big\{\big|\bar{X}_a\bar{X}_b\big| \geq c|\sigma_j|\Big\}\\
\leq&\sum_{j\in J_1}c_1\big[e^{-c_2k_0\min\{\sigma_j^2,|\sigma_j|\}} + e^{-c_3k_0|\sigma_j|} + e^{-c_4n|\sigma_j|}\big].
\end{aligned}$$

For the second term, note that δ_j and σ_j share the same sign, and thus we can similarly proceed as follows:

$$\begin{aligned}
\mathbb{P}\Big\{\big|\sum_{j\in J_2}\sum_{i=1}^{k_0}\delta_j\tilde{\sigma}_{ij}\big| \geq ck_0\big|\sum_{j\in J_2}\delta_j\sigma_j\big|\Big\} &\leq \mathbb{P}\Big\{\sum_{j\in J_2}\big|\delta_j\sum_{i=1}^{k_0}\tilde{\sigma}_{ij}\big| \geq ck_0\sum_{j\in J_2}\big|\delta_j\big|\big|\sigma_j\big|\Big\}\\
&\leq \sum_{j\in J_2}\mathbb{P}\Big\{\frac{1}{k_0}\big|\sum_{i=1}^{k_0}\tilde{\sigma}_{ij}\big| \geq c\big|\sigma_j\big|\Big\}\\
&\leq \sum_{j\in J_2}c_1\big[e^{-c_2k_0\min\{\sigma_j^2,|\sigma_j|\}} + e^{-c_3k_0|\sigma_j|} + e^{-c_4n|\sigma_j|}\big].
\end{aligned}$$

Under Assumption 1(b), we have

$$\mathbb{P}\big\{|B_1| \geq cD\big\} \leq c_1p^2\big[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\big]. \tag{S3.10}$$

For the term B_2, it is similar to B_1 to obtain the result as follows

$$\begin{aligned}\mathbb{P}\Big\{|B_2| \geq cD\Big\} &= \mathbb{P}\Big\{\big|\sum_{i=k_0+1}^{n} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W}\boldsymbol{\sigma}\big| \geq ca(k,k_0)\boldsymbol{\sigma}^\top \boldsymbol{W}\boldsymbol{\sigma}\Big\} \\ &\leq \mathbb{P}\Big\{\big\|\boldsymbol{W}^{\frac{1}{2}} \sum_{i=k_0+1}^{n} \tilde{\boldsymbol{\sigma}}_{(i)}\big\|_2 \big\|\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\sigma}\big\|_2 \geq ca(k,k_0)\big\|\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\sigma}\big\|_2^2\Big\} \\ &\leq \sum_{j\in J_1\cup J_2} c_1\Big[e^{-c_2(n-k_0)\min\left\{\frac{a(k,k_0)^2}{(n-k_0)^2}\sigma_j^2, \frac{a(k,k_0)}{n-k_0}|\sigma_j|\right\}} + e^{-c_3 a(k,k_0)|\sigma_j|} + e^{-c_4 n \frac{a(k,k_0)}{n-k_0}|\sigma_j|}\Big],\end{aligned}$$

where $a(k,k_0) = \frac{(k(k-1)-k_0(k_0-1))(n-k)(n-k_0)}{|k(n-k)-k_0(n-k_0)|(k-1)}$. Lemma 1 give the last inequality, and we have

$$\mathbb{P}\Big\{|B_2| \geq cD\Big\} \leq c_1 p^2 \big[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\big]. \tag{S3.11}$$

For the term B_3 similarly, we have

$$\mathbb{P}\Big\{|B_3| \geq cD\Big\} = \mathbb{P}\Big\{\big|\sum_{i=k_0+1}^{k} \tilde{\boldsymbol{\sigma}}_i^\top \boldsymbol{W}\tilde{\boldsymbol{\sigma}}\big| \geq ca(k,k_0)\tilde{\boldsymbol{\sigma}}^\top \boldsymbol{W}\tilde{\boldsymbol{\sigma}}\Big\},$$

where $a(k,k_0) = \frac{(k(k-1)-k_0(k_0-1))(n-k)}{k_0(n-1)}$ and similarly we can obtain the result

$$\mathbb{P}\Big\{|B_3| \geq cD\Big\} \leq c_1 p^2 \Big[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\Big]. \tag{S3.12}$$

Combining (S3.10)-(S3.12), we can conclude that

$$\mathbb{P}\Big\{|B(k) - B(k_0)| \geq cD\Big\} \leq c_1 p^2 \Big[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\Big]. \tag{S3.13}$$

The terms $C(k)$ and $C(k_0)$ are the same, that is

$$\mathbb{P}\Big\{|C(k) - C(k_0)| \geq cD\Big\} \leq c_1 p^2 \Big[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\Big]. \tag{S3.14}$$

Finally, we investigate the terms $A(k)$ and $A(k_0)$ introducing the decomposition

$$A(k) - A(k_0) = H_1 + H_2 + H_3 + H_4 + H_5 + H_6,$$

where

$$H_1 = a_1 \sum_{(i \neq t)=1}^{k_0} \sum^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)}, \quad H_2 = a_2 \sum_{(i \neq t)=k_0+1}^{k} \sum^{k} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \tilde{\boldsymbol{\sigma}}_{(t)}, \quad H_3 = a_3 \sum_{(i \neq t)=k+1}^{n} \sum^{n} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)},$$
$$H_4 = a_4 \sum_{i=1}^{k_0} \sum_{t=k_0+1}^{k} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)}, \quad H_5 = a_5 \sum_{i=k_0+1}^{k} \sum_{t=k+1}^{n} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)}, \quad H_6 = a_6 \sum_{i=1}^{k_0} \sum_{t=k+1}^{n} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)},$$

and

$$a_1 = \frac{1}{(k)_2} - \frac{1}{(k_0)_2}, \quad a_2 = \frac{1}{(k)_2} - \frac{1}{(n-k_0)_2}, \quad a_3 = \frac{1}{(n-k)_2} - \frac{1}{(n-k_0)_2},$$
$$a_4 = \frac{2}{(k)_2} + \frac{2}{k_0(n-k_0)}, \quad a_5 = -\frac{2}{k(n-k)} - \frac{2}{(n-k_0)_2}, \quad a_6 = \frac{2}{k_0(n-k_0)} - \frac{2}{k(n-k)}.$$

In order to prove that $\mathbb{P}\big\{|A(k) - A(k_0)| \geq cD\big\} = o\big(\frac{1}{n}\big)$, it is sufficient to show that $\mathbb{P}\big\{|H_i| \geq cD\big\} = o\big(\frac{1}{n}\big), i = 1, \ldots, 6$ uniformly with respect to $(1+\epsilon)k_0 \leq k \leq c_2 n$.

The term H_1 can be handled as follows:

$$\mathbb{P}\big\{\big|H_1\big| \geq cD\big\} \leq \mathbb{P}\Big\{\big|\sum_{i=1}^{k_0}\sum_{t=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)}\big| \geq \frac{c}{|a_1|} D\Big\} + \mathbb{P}\Big\{\big|\sum_{i=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(i)}\big| \geq \frac{c}{|a_1|} D\Big\}.$$

For the first term, we can find that

$$\mathbb{P}\Big\{\big|\sum_{i=1}^{k_0}\sum_{t=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_{(t)}\big| \geq \frac{c}{|a_1|} D\Big\} = \mathbb{P}\Big\{\big\|\boldsymbol{W}^{\frac{1}{2}} \sum_{i=1}^{k_0} \tilde{\boldsymbol{\sigma}}_{(i)}\big\|_2^2 \geq c b_1(k, k_0) \big\|\boldsymbol{W}^{\frac{1}{2}} \boldsymbol{\sigma}\big\|_2^2\Big\},$$

where $b_1(k, k_0) = \frac{1 - \frac{(k_0)_2}{(k)_2}}{|a_1|}$ and similarly we can obtain the result

$$\mathbb{P}\Big\{\big|\sum_{i=1}^{k_0}\sum_{t=1}^{k_0} \tilde{\boldsymbol{\sigma}}_i^\top \boldsymbol{W} \tilde{\boldsymbol{\sigma}}_t\big| \geq \frac{c}{|a_1|} D\Big\} \leq c_1 p^2 [e^{-c_2 \min\{\tau, \sqrt{n\tau}\}} + e^{-c_3 \sqrt{n\tau}}].$$

For the second term,

$$
\begin{aligned}
&\mathbb{P}\Big\{\big|\sum_{i=1}^{k_0}\tilde{\boldsymbol{\sigma}}_{(i)}^{\top}\boldsymbol{W}\tilde{\boldsymbol{\sigma}}_{(i)}\big| \geq \frac{c}{a_1}D\Big\}\\
\leq&\mathbb{P}\Big\{\sum_{i=1}^{k_0}\sum_{j\in J_1}\gamma_j\tilde{\sigma}_{ij}^2 \geq cb_1(k,k_0)\sum_{j\in J_1}\gamma_j\sigma_j^2\Big\} + \mathbb{P}\Big\{\sum_{i=1}^{k_0}\big(\sum_{j\in J_2}\delta_j\tilde{\sigma}_{ij}\big)^2 \geq cb_1(k,k_0)\big(\sum_{j\in J_2}^{m}\delta_j\sigma_j\big)^2\Big\}\\
\leq&\sum_{i=1}^{k_0}\sum_{j\in J_1\cup J_2}\mathbb{P}\Big\{|\tilde{\sigma}_{ij}| \geq c\sqrt{\frac{b_1(k,k_0)}{k_0}}|\sigma_j|\Big\}\\
\leq&\sum_{i=1}^{k_0}\sum_{j\in J_1\cup J_2}\Big[\mathbb{P}\Big\{\big|X_{ia}X_{ib}-\mathbb{E}X_{ia}X_{ib}\big| \geq c\sqrt{\frac{b_1(k,k_0)}{k_0}}|\sigma_j|\Big\} + \mathbb{P}\Big\{\big|X_{ia}\bar{X}_b\big| \geq c\sqrt{\frac{b_1(k,k_0)}{k_0}}|\sigma_j|\Big\}\\
&+\mathbb{P}\Big\{\big|X_{ib}\bar{X}_a\big| \geq c\sqrt{\frac{b_1(k,k_0)}{k_0}}|\sigma_j|\Big\} + \mathbb{P}\Big\{\big|\bar{X}_a\bar{X}_b\big| \geq c\sqrt{\frac{b_1(k,k_0)}{k_0}}|\sigma_j\Big\}\Big]\\
\leq&\sum_{i=1}^{k_0}\sum_{j\in J_1\cup J_2}c_1\big[e^{-c_2\sqrt{\frac{b_1(k,k_0)}{k_0}}|\tilde{\sigma}_j|} + e^{-c_3 n\sqrt{\frac{b_1(k,k_0)}{k_0}}|\tilde{\sigma}_j|}\big] \leq c_1p^2n[e^{-c_2\sqrt{\tau}} + e^{-c_3 n\sqrt{\tau}}].
\end{aligned}
$$

The penultimate inequality is given by Assumption 1(a) and Lemma 1, hence we have

$$
\mathbb{P}\{|H_1| \geq cD\} \leq c_1p^2[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}} + ne^{-c_4\sqrt{\tau}} + ne^{-c_5 n\sqrt{\tau}}]. \tag{S3.15}
$$

For the terms H_2 and H_3, similarly we can obtain the results

$$
\begin{aligned}
&\mathbb{P}\{|H_2| \geq cD\} \leq c_1p^2[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}} + ne^{-c_4\sqrt{\tau}} + ne^{-c_5 n\sqrt{\tau}}],\\
&\mathbb{P}\{|H_3| \geq cD\} \leq c_1p^2[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}} + ne^{-c_4\sqrt{\tau}} + ne^{-c_5 n\sqrt{\tau}}].
\end{aligned} \tag{S3.16}
$$

The term H_4 can be handled as follows:

$$
\begin{aligned}
\mathbb{P}\big\{\big|H_4\big| \geq cD\big\} &= \mathbb{P}\Big\{\big|a_4\sum_{i=1}^{k_0}\sum_{t=k_0+1}^{k}\tilde{\boldsymbol{\sigma}}_{(i)}^{\top}\boldsymbol{W}\tilde{\boldsymbol{\sigma}}_{(t)}\big| \geq cD\Big\}\\
&\leq \mathbb{P}\Big\{\big\|\boldsymbol{W}^{\frac{1}{2}}\sum_{i=1}^{k_0}\tilde{\boldsymbol{\sigma}}_{(i)}\big\|_2\big\|\boldsymbol{W}^{\frac{1}{2}}\sum_{t=k_0+1}^{k}\tilde{\boldsymbol{\sigma}}_{(t)}\big\|_2 \geq \frac{c}{|a_4|}D\Big\}\\
&\leq \mathbb{P}\Big\{\big\|\boldsymbol{W}^{\frac{1}{2}}\sum_{i=1}^{k_0}\tilde{\boldsymbol{\sigma}}_{(i)}\big\|_2^2 \geq \frac{c}{a_4}D\Big\} + \mathbb{P}\Big\{\big\|\boldsymbol{W}^{\frac{1}{2}}\sum_{t=k_0+1}^{k}\tilde{\boldsymbol{\sigma}}_{(t)}\big\|_2^2 \geq \frac{c}{a_4}D\Big\}\\
&\leq c_1p^2[e^{-c_2\min\{\tau,\sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}],
\end{aligned} \tag{S3.17}
$$

For the terms H_5 and H_6, similarly we can obtain the results

$$\begin{aligned} &\mathbb{P}\Big\{\big|H_5\big| \geq cD\Big\} \leq c_1 p^2 \big[e^{-c_2 \min\{\tau, \sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\big] \\ &\mathbb{P}\Big\{\big|H_6\big| \geq cD\Big\} \leq c_1 p^2 \big[e^{-c_2 \min\{\tau, \sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}}\big]. \end{aligned} \tag{S3.18}$$

Combining (S3.15)-(S3.18), we can conclude that

$$\mathbb{P}\Big\{|A(k) - A(k_0)| \geq cD\Big\} \leq c_1 p^2 \big[e^{-c_2 \min\{\tau, \sqrt{n\tau}\}} + e^{-c_3\sqrt{n\tau}} + ne^{-c_4\sqrt{\tau}} + ne^{-c_5 n\sqrt{\tau}}\big]. \tag{S3.19}$$

Combining (S3.13), (S3.14) and (S3.19), we have

$$\begin{aligned} \mathbb{P}\Big\{\bigcup_{k \geq (1+\epsilon)k_0} \big(T_{\boldsymbol{W}}(k) \geq T_{\boldsymbol{W}}(k_0)\big)\Big\} \leq &\sum_{k \geq (1+\epsilon)k_0} \mathbb{P}\big\{T_{\boldsymbol{W}}(k) \geq T_{\boldsymbol{W}}(k_0)\big\} \\ \leq & c_1 p^2 n\big[e^{-c_2\tau} + e^{-c_3\sqrt{n\tau}} + ne^{-c_4\sqrt{\tau}} + ne^{-c_5 n\sqrt{\tau}}\big]. \end{aligned}$$

In particular, if $p^2n^2 = o(e^{c\sqrt{\tau}})$ holds, $\mathbb{P}\big\{\big|\frac{\hat{k}_0}{k_0} - 1\big| \geq \epsilon\big\} \to 0$. This completes the proof of Theorem 1.

Bibliography

Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*, Volume 47. Cambridge University Press.