SUPPLEMENTARY MATERIAL FOR "CHANGE-POINT DETECTION IN MULTINOMIAL DATA WITH A LARGE NUMBER OF CATEGORIES"

By Guanghui Wang*, Changliang Zou* and Guosheng Yin†

Nankai University and The University of Hong Kong†*

This Supplementary Material contains all theoretical proofs of Theorems 1–5, Proposition 1 and Corollary 1, and additional simulation results.

The key step of the proofs is the decomposition of the observations into $X_{tj} = \sum_{i=N_{0,t-1}+1}^{N_{0t}} Y_{ij}$ for $j = 1, \ldots, p,\ t = 1, \ldots, T$ with the convention of $N_{00} = 0$ and $N_{0T} = N$, where $\{(Y_{i1}, \ldots, Y_{ip})^\top\}_{i=1}^N$ are independent according to a sequence of multinomial distributions, $\mathrm{Multi}\big(1, (\tilde{p}_{i1}, \ldots, \tilde{p}_{ip})^\top\big)$, where

$$\tilde{p}_{ij} = \begin{cases} q_{0j}, & \text{for } i = 1, \ldots, N_{0\tau^*}, \\ q_{1j}, & \text{for } i = N_{0\tau^*} + 1, \ldots, N. \end{cases}$$

Denote $\tilde{Y}_{ij} = Y_{ij} - \tilde{p}_{ij}$, and let $\mathbf{q} \equiv \mathbf{q}_{0,\mathscr{B}}$ be the collection of q_{0j} with $j \in \mathscr{B}$ and $\mathbf{\Sigma} \equiv \mathbf{\Sigma}_{0,\mathscr{B}} = \mathrm{diag}(\mathbf{q}) - \mathbf{q}\mathbf{q}^\top$, where $\mathrm{diag}(\boldsymbol{a})$ is a diagonal matrix whose diagonal elements are the elements of vector $\boldsymbol{a}$. Denote $\widetilde{\mathbf{Y}}_i \equiv \widetilde{\mathbf{Y}}_{i,\mathscr{B}}$ as the collection of all $\tilde{Y}_{ij}$ with $j \in \mathscr{B}$.

Technical lemmas. We start with several necessary lemmas.

Lemma S.1. *Under Assumption (A1), we have*

$$\begin{aligned}
\mathrm{tr}(\mathbf{\Sigma}^2) &= \mathbf{q}^\top\mathbf{q}\{1 + o(1)\}, \\
\mathbf{q}^\top\mathbf{\Sigma}\mathbf{q}/\mathrm{tr}(\mathbf{\Sigma}^2) &= o(1), \\
\mathrm{tr}(\mathbf{\Sigma}^4)/(\mathbf{q}^\top\mathbf{q})^2 &= \Big(\sum_{j\in\mathscr{B}} q_{0j}^4\Big)\{1 + o(1)\}/(\mathbf{q}^\top\mathbf{q})^2 + o(1), \\
\mathrm{E}\{(\widetilde{\mathbf{Y}}_1^\top\mathbf{\Sigma}\widetilde{\mathbf{Y}}_1)^2\}/(\mathbf{q}^\top\mathbf{q})^2 &= \Big(\sum_{j\in\mathscr{B}} q_{0j}^3\Big)\{1 + o(1)\}/(\mathbf{q}^\top\mathbf{q})^2 + o(1), \\
\mathrm{E}\Big\{(\widetilde{\mathbf{Y}}_1^\top\widetilde{\mathbf{Y}}_2)^4/(\mathbf{q}^\top\mathbf{q})^2\Big\} &= (\mathbf{q}^\top\mathbf{q})^{-1} + O(1) + O\Big\{\Big(\sum_{j\in\mathscr{B}} q_{0j}^3\Big)/(\mathbf{q}^\top\mathbf{q})^2\Big\}.
\end{aligned}$$

Lemma S.2 (Bernstein's inequality). *Let $X_1, \ldots, X_n$ be independent centered random variables a.s. bounded by $A < \infty$ in absolute value. Let $\sigma^2 =$*

$n^{-1}\sum_{i=1}^{n} \mathrm{E}(X_i^2)$. *Then for all* $x > 0$,

$$\Pr\Big(\sum_{i=1}^{n} X_i \geq x\Big) \leq \exp\Big(-\frac{x^2}{2n\sigma^2 + 2Ax/3}\Big).$$

LEMMA S.3 (A Bernstein-type inequality for degenerate U-statistics of order two). *Consider the degenerate U-statistic of order two, i.e.,* $\mathbb{U}_n = \sum_{1\leq i<k\leq n} h_{ik}(\mathbf{X}_i, \mathbf{X}_k)$, *where* $\mathbf{X}_i$*'s are independent random vectors and* h_{ik} *satisfies that, for all* i, k, $\mathrm{E}|h_{ik}(\mathbf{X}_i, \mathbf{X}_k)| < \infty$ *and* $\mathrm{E}\{h_{ik}(\mathbf{x}_1, \mathbf{X}_k)\} = \mathrm{E}\{h_{ik}(\mathbf{X}_i, \mathbf{x}_2)\} = 0$. *Then for all* $x > 0$,

$$\Pr(\mathbb{U}_n \geq x) \leq \exp\Big[1 - \min\Big\{C_1\Big(\frac{x}{C}\Big)^2, C_2\frac{x}{D}, C_3\Big(\frac{x}{B}\Big)^{2/3}, C_4\Big(\frac{x}{A}\Big)^{1/2}\Big\}\Big],$$

where

$$\begin{aligned}
A &= \max_{i,k}\sup_{\mathbf{x}_1,\mathbf{x}_2} |h_{ik}(\mathbf{x}_1, \mathbf{x}_2)|, \\
B^2 &= \max\Big\{\max_{k}\sup_{\mathbf{x}_2}\sum_{i=1}^{k-1}\mathrm{E}\{h_{ik}^2(\mathbf{X}_i, \mathbf{x}_2)\}, \max_{i}\sup_{\mathbf{x}_1}\sum_{k=i+1}^{n}\mathrm{E}\{h_{ik}^2(\mathbf{x}_1, \mathbf{X}_k)\}\Big\}, \\
C^2 &= \sum_{i<k}\mathrm{E}\{h_{ik}^2(\mathbf{X}_i, \mathbf{X}_k)\}, \\
D &= \sup_{f,g}\Big\{\sum_{i<k}\mathrm{E}\{h_{ik}(\mathbf{X}_i, \mathbf{X}_k)f_i(\mathbf{X}_i)g_k(\mathbf{X}_k)\} : \sum_{i=1}^{n-1}\mathrm{E}\{f_i^2(\mathbf{X}_i)\} \leq 1, \\
&\qquad \sum_{k=2}^{n}\mathrm{E}\{g_k^2(\mathbf{X}_k)\} \leq 1\Big\}
\end{aligned}$$

and C_1, C_2, C_3, C_4 *are positive constants.*

The first lemma can be verified by some algebraic calculations. The second and third lemmas are on concentration inequalities. The first one is the well-known Bernstein inequality for the sum of centered bounded independent random variables, and the second is its extension to degenerate U-statistics of order two based on an independent vector-valued sample. The proofs can be found in Theorem 3.1.7 and Theorem 3.4.8 of **?**, respectively.

Proof of Theorem 1. We only give a detailed proof of (ii). Let $q_{*j} = \frac{N_{0\tau^*}}{N}q_{0j} + \frac{N_{1\tau^*}}{N}q_{1j}$ and $\sigma_{*j}^2 = \frac{N_{0\tau^*}}{N}\sigma_{0j}^2 + \frac{N_{1\tau^*}}{N}\sigma_{1j}^2$, where $\sigma_{lj}^2 = q_{lj}(1-q_{lj})$ for $l = 0, 1$. We observe that $|\mathscr{A}| < 2a_p/\varepsilon$ and $|\hat{\mathscr{A}}| < a_p/(C\varepsilon)$, where $|\mathscr{A}|$ denotes the cardinal number of a set $\mathscr{A}$.

Assume $\mathscr{A} \neq \emptyset$. We first show that $\Pr(\mathscr{A} \subset \hat{\mathscr{A}}) \to 1$, as $(p, N) \to \infty$. If $\mathscr{A} \not\subset \hat{\mathscr{A}}$, then there exists some $j \in \mathscr{A}_0 \cup \mathscr{A}_1$ but $j \notin \hat{\mathscr{A}}$, i.e., $q_{lj} a_p > \varepsilon$ for $l = 0$ or 1 and $\hat{q}_j a_p \leq C\varepsilon$. We can obtain that $(q_{*j} - \hat{q}_j) a_p > (C^* - C)\varepsilon$ for some $j \in \mathscr{A}$ as $(p, N) \to \infty$, where $C^* = \min(\kappa_0, 1 - \kappa_0)/2$. It follows that $\{\mathscr{A} \not\subset \hat{\mathscr{A}}\} \subset \{(q_{*j} - \hat{q}_j) a_p > (C^* - C)\varepsilon \text{ for some } j \in \mathscr{A}\}$. Hence $\big\{\max_{j \in \mathscr{A}} (q_{*j} - \hat{q}_j) a_p \leq (C^* - C)\varepsilon\big\} \subset \{\mathscr{A} \subset \hat{\mathscr{A}}\}$. Consequently, for any $C < C^*$,

$$
\begin{aligned}
\Pr(\mathscr{A} \subset \hat{\mathscr{A}}) &\geq \Pr\left\{\max_{j \in \mathscr{A}} (q_{*j} - \hat{q}_j) a_p \leq (C^* - C)\varepsilon\right\} \\
&\geq 1 - \sum_{j \in \mathscr{A}} \Pr\left\{(q_{*j} - \hat{q}_j) a_p > (C^* - C)\varepsilon\right\} \\
&= 1 - \sum_{j \in \mathscr{A}} \Pr\left\{\sum_{i=1}^{N} (-\tilde{Y}_{ij}) > N a_p^{-1} (C^* - C)\varepsilon\right\} \\
&\geq 1 - |\mathscr{A}| \exp\left\{-\frac{N a_p^{-2} (C^* - C)^2 \varepsilon^2}{2 + 2a_p^{-1}(C^* - C)\varepsilon/3}\right\} \to 1,
\end{aligned}
$$

provided that $N a_p^{-2} (\log a_p)^{-1} \to \infty$. The last inequality follows from Lemma S.2 and, to be specific, for all $x > 0$,

$$
\Pr\left\{\sum_{i=1}^{N} (-\tilde{Y}_{ij}) > x\right\} \leq \exp\Big(-\frac{x^2}{2N\sigma_{*j}^2 + 2x/3}\Big),
$$

and the fact that $\sigma_{lj}^2 \leq 1$.

Next, we demonstrate that $\Pr(\hat{\mathscr{A}} \subset \mathscr{A}) \to 1$, as $(p, N) \to \infty$. If $\hat{\mathscr{A}} \not\subset \mathscr{A}$, then there exists some $j \in \hat{\mathscr{A}}$ but $j \notin \mathscr{A}$, i.e., $\hat{q}_j a_p > C\varepsilon$ and, by Assumption (A1), for any $C > 0$ and large enough p, $q_{lj} a_p < C\varepsilon/2$ for any $l = 0, 1$. We have that $(\hat{q}_j - q_{*j}) a_p > C\varepsilon/2$ for some $j \in \hat{\mathscr{A}}$. Similarly, we can show that

$$
\begin{aligned}
\Pr(\hat{\mathscr{A}} \subset \mathscr{A}) &\geq \Pr\left\{\max_{j \in \hat{\mathscr{A}}} (\hat{q}_j - q_{*j}) a_p \leq C\varepsilon/2\right\} \\
&\geq 1 - \sum_{j \in \hat{\mathscr{A}}} \Pr\Big(\sum_{i=1}^{N} \tilde{Y}_{ij} > N a_p^{-1} C\varepsilon/2\Big) \to 1,
\end{aligned}
$$

provided that $N a_p^{-2} (\log a_p)^{-1} \to \infty$. Finally, it is easy to check that $\hat{\mathscr{A}} = \emptyset$ if $\mathscr{A} = \emptyset$. Hence (ii) follows. $\square$

Proof of Theorem 2. Proof of part (i) Let $c_\tau = N_{0\tau}N_{1\tau}/N$ and we introduce

$$a_{i,\tau} = \begin{cases} N_{0\tau}^{-1}, & i = 1, \dots, N_{0\tau}, \\ -N_{1\tau}^{-1}, & i = N_{0\tau}+1, \dots, N. \end{cases}$$

For simplicity, we use $\sum_\tau$ and $\sum_{\tau<\tau'}$ to denote $\sum_{\tau\in\mathcal{T}}$ and $\sum_{\substack{\tau<\tau' \\ \tau,\tau'\in\mathcal{T}}}$, respectively. Also let $b_{ik} = \sum_\tau c_\tau a_{i,\tau}a_{k,\tau}$ for $1 \le i,k \le N$. Under the null hypothesis, we can rewrite $L_{\tau j}$ as

$$L_{\tau j} - L_{\tau j}^{(0)} = -q_{0j}^2 + 2\sum_{1\le i<k\le N} c_\tau a_{i,\tau}a_{k,\tau}\tilde{Y}_{ij}\tilde{Y}_{kj} - 2\sum_{i=1}^N c_\tau a_{i,\tau}^2\tilde{Y}_{ij}q_{0j},$$

and $S_p \equiv S_{p,\mathscr{A}}$ as

$$S_p = -\Lambda_T\sum_{j\in\mathscr{B}} q_{0j}^2 + S_{p,1} + S_{p,2},$$

where $S_{p,1} = 2\sum_{1\le i<k\le N} b_{ik}\sum_{j\in\mathscr{B}}\tilde{Y}_{ij}\tilde{Y}_{kj}$ and $S_{p,2} = -2\sum_{i=1}^N b_{ii}\sum_{j\in\mathscr{B}}\tilde{Y}_{ij}q_{0j}$.

Recall that $\mathbf{q} \equiv \mathbf{q}_{0,\mathscr{B}}$ is the collection of q_{0j} with $j \in \mathscr{B}$ and $\mathbf{\Sigma} \equiv \mathbf{\Sigma}_{0,\mathscr{B}} = \mathrm{diag}(\mathbf{q}) - \mathbf{q}\mathbf{q}^\top$. It is easy to verify that $\mathrm{E}(S_p) = -\Lambda_T\mathbf{q}^\top\mathbf{q}$ and $\mathrm{Var}(S_p) = \mathrm{Var}(S_{p,1}) + \mathrm{Var}(S_{p,2})$, where under Assumption (A2),

$$\mathrm{Var}(S_{p,1}) = \sum_{i\ne k} b_{ik}^2\cdot 2\mathrm{tr}(\mathbf{\Sigma}^2) = \Big\{2\sum_{\tau<\tau'}\frac{N_{0\tau}}{N_{1\tau}}\frac{N_{1\tau'}}{N_{0\tau'}} + \Lambda_T + O\Big(\frac{T^2}{N}\Big)\Big\}\cdot 2\mathrm{tr}(\mathbf{\Sigma}^2)$$

and

$$\mathrm{Var}(S_{p,2}) = 4\sum_i b_{ii}^2\mathbf{q}^\top\mathbf{\Sigma}\mathbf{q} = O\Big(\frac{T^2}{N}\Big)\mathbf{q}^\top\mathbf{\Sigma}\mathbf{q}.$$

Hence, we conclude that $\mathrm{Var}(S_{p,2})/\mathrm{Var}(S_{p,1}) \to 0$ and

$$\mathrm{Var}(S_p) = \Big(2\sum_{\tau<\tau'}\frac{N_{0\tau}}{N_{1\tau}}\Big/\frac{N_{0\tau'}}{N_{1\tau'}} + \Lambda_T\Big)\cdot 2\mathbf{q}^\top\mathbf{q}\{1+o(1)\},$$

under Assumption (A1) and by Lemma S.1. □

Proof of part (ii) Note that

$$\frac{S_p - \mathrm{E}(S_p)}{\sqrt{\mathrm{Var}(S_p)}} = S_{p,1}/\sigma_p + o_p(1),$$

where $\sigma_p^2 = 2c_{N,T}\sum_{j\in\mathscr{B}} q_{0j}^2$ and $c_{N,T} = 2\sum_{\tau<\tau'} N_{0\tau}N_{1\tau'}/(N_{1\tau}N_{0\tau'}) + \Lambda_T$. Recall that $\widetilde{\mathbf{Y}}_i \equiv \widetilde{\mathbf{Y}}_{i,\mathscr{B}}$ is the collection of all $\tilde{Y}_{ij}$ with $j \in \mathscr{B}$. Let $S_{N1} = 0$, and for $l = 2, \ldots, N$, let $V_{Nl} = \sum_{i=1}^{l-1} b_{il}\widetilde{\mathbf{Y}}_i^\top\widetilde{\mathbf{Y}}_l$ and $S_{Nl} = \sum_{k=2}^{l} V_{Nk}$. If we define $\mathcal{F}_{N,l-1} = \sigma\{\widetilde{\mathbf{Y}}_1, \ldots, \widetilde{\mathbf{Y}}_{l-1}\}$, the σ-algebra generated by $\{\widetilde{\mathbf{Y}}_1, \ldots, \widetilde{\mathbf{Y}}_{l-1}\}$, then $\mathrm{E}(V_{Nl}|\mathcal{F}_{N,l-1}) = 0$. Hence, for each N, $S_{N2}, \ldots, S_{NN}$ are martingales and $V_{N2}, \ldots, V_{NN}$ are the martingale differences. Based on Corollary 3.1 of **?**, it suffices to show that for any $\epsilon > 0$,

$$\sum_{l=2}^{N} \sigma_p^{-2}\mathrm{E}\{V_{Nl}^2\mathrm{I}(|V_{Nl}| > \epsilon\sigma_p)\} \xrightarrow{\mathcal{P}} 0, \tag{S.1}$$

and there exists $\gamma \in (0, \infty)$ such that

$$\sum_{l=2}^{N} \sigma_p^{-2}\mathrm{E}(V_{Nl}^2|\mathcal{F}_{N,l-1}) \xrightarrow{\mathcal{P}} \gamma. \tag{S.2}$$

Then we can assert that $S_{p,1}/\sigma_p \xrightarrow{\mathcal{D}} N(0,1)$ as $(p,N) \to \infty$.

We first verify (S.2). For simplicity, we denote $\mathrm{E}_{l-1}(\cdot) = \mathrm{E}(\cdot|\mathcal{F}_{N,l-1})$, and then

$$U_l = \mathrm{E}_{l-1}\left(V_{Nl}^2\right) = \sum_{1\le i_1,i_2\le l-1} b_{i_1 l}b_{i_2 l}\widetilde{\mathbf{Y}}_{i_1}^\top\mathbf{\Sigma}\widetilde{\mathbf{Y}}_{i_2}.$$

It is easy to show that

$$\mathrm{E}\Big(\sum_{l=2}^{N} U_l\Big) = \sum_{i_1<l} b_{i_1 l}^2\mathrm{tr}(\mathbf{\Sigma}^2) = \frac{1}{4}\sigma_p^2\{1 + o(1)\},$$

which indicates $\gamma = 1/4$.

We next demonstrate that $\mathrm{E}\Big\{\Big(\sigma_p^{-2}\sum_{l=2}^{N} U_l - 1/4\Big)^2\Big\} \to 0$ as $(p,N) \to \infty$. It suffices to show that

$$\mathrm{E}\Big\{\Big(\sum_{l=2}^{N} U_l\Big)^2\Big\}\Big/\sigma_p^4 = \sum_{l=2}^{N}\mathrm{E}(U_l^2)/\sigma_p^4 + 2\sum_{2\le l<l'\le N}\mathrm{E}(U_l U_l')/\sigma_p^4 \to 1/16.$$

Assume $l < l'$ and expand $\mathrm{E}(U_l U_l')$ as

$$\begin{aligned}
\mathrm{E}(U_l U_l') &= \mathrm{E}\Big\{\Big(\sum_{1\le i_1,i_2\le l-1} b_{i_1 l} b_{i_2 l} \widetilde{\mathbf{Y}}_{i_1}^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_{i_2}\Big)\Big(\sum_{1\le i_3,i_4\le l'-1} b_{i_3 l'} b_{i_4 l'} \widetilde{\mathbf{Y}}_{i_3}^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_{i_4}\Big)\Big\} \\
&= \sum_{\substack{1\le i_1,i_2\le l-1 \\ 1\le i_3,i_4\le l'-1}} b_{i_1 l} b_{i_2 l} b_{i_3 l'} b_{i_4 l'} \mathrm{E}\left(\widetilde{\mathbf{Y}}_{i_1}^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_{i_2} \widetilde{\mathbf{Y}}_{i_3}^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_{i_4}\right) \\
&= 2 \sum_{1\le i_1\ne i_2\le l-1} b_{i_1 l}^2 b_{i_2 l}^2 \mathrm{tr}(\mathbf{\Sigma}^4) + \sum_{i_1=1}^{l-1} \sum_{i_3=1, i_3\ne i_1}^{l'-1} b_{i_1 l}^2 b_{i_3 l'}^2 \mathrm{tr}^2(\mathbf{\Sigma}^2) \\
&\quad + \sum_{i_1=1}^{l-1} b_{i_1 l}^4 \mathrm{E}\Big\{\Big(\widetilde{\mathbf{Y}}_{i_1}^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_{i_1}\Big)^2\Big\}.
\end{aligned}$$

Consequently, we can write

$$\mathrm{E}\Big\{\Big(\sum_{l=2}^{N} U_l\Big)^2\Big\}\Big/\sigma_p^4 = A + B_1 + B_2 + B_3,$$

where

$$\begin{aligned}
A &= 2 \sum_{2\le l<l'\le N} \sum_{i_1=1}^{l-1} \sum_{i_3=1, i_3\ne i_1}^{l'-1} b_{i_1 l}^2 b_{i_3 l'}^2 \mathrm{tr}^2(\mathbf{\Sigma}^2)/\sigma_p^4, \\
B_1 &= 4 \sum_{l=2}^{N-1} (N-l) \sum_{1\le i_1\ne i_2\le l-1} b_{i_1 l}^2 b_{i_2 l}^2 \mathrm{tr}(\mathbf{\Sigma}^4)/\sigma_p^4, \\
B_2 &= 2 \sum_{l=2}^{N-1} (N-l) \sum_{i_1=1}^{l-1} b_{i_1 l}^4 \mathrm{E}\Big\{\Big(\widetilde{\mathbf{Y}}_1^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_1\Big)^2\Big\}\Big/\sigma_p^4,
\end{aligned}$$

and $B_3 = \sum_{l=2}^{N} \mathrm{E}(U_l^2)/\sigma_p^4$. Under Assumptions (A1)–(A2) and by Lemma S.1, tedious algebra yields that

$$A = \Big(\sum_{i<l} b_{il}^2\Big)^2 \mathrm{tr}^2(\mathbf{\Sigma}^2)/\sigma_p^4\{1+o(1)\} = \frac{1}{16}\{1+o(1)\}.$$

Additionally, under Assumptions (A3)–(A4), we conclude that

$$B_1 = O\Big(\frac{\mathrm{tr}(\mathbf{\Sigma}^4)}{(\mathbf{q}^\top\mathbf{q})^2}\Big) = o(1) \text{ and } B_2 = O\Big(N^{-1}\frac{\mathrm{E}\{(\widetilde{\mathbf{Y}}_1^\top \mathbf{\Sigma} \widetilde{\mathbf{Y}}_1)^2\}}{(\mathbf{q}^\top\mathbf{q})^2}\Big) = o(1).$$

Similarly, we can show that $B_3 = o(1)$, and thus (S.2) follows.

The Lindeberg condition (S.1) can be established by showing that a stronger Liapounov condition, $\sum_{l=2}^{N} \mathrm{E}(V_{Nl}^4)/\sigma_p^4 = o(1)$, holds. We expand $\sum_{l=2}^{N} \mathrm{E}(V_{Nl}^4)$ as

$$\begin{aligned}
& \sum_{l=2}^{N} \mathrm{E}(V_{Nl}^4) \\
= \ & \sum_{l=2}^{N} \mathrm{E}\Big\{\Big(\sum_{i=1}^{l-1} b_{il}\widetilde{\mathbf{Y}}_i^\top \widetilde{\mathbf{Y}}_l\Big)^4\Big\} \\
= \ & \sum_{l=2}^{N}\sum_{i=1}^{l-1} b_{il}^4 \mathrm{E}(\widetilde{\mathbf{Y}}_i^\top \widetilde{\mathbf{Y}}_l)^4 + \sum_{l=2}^{N} 3 \sum_{1\le i_1 \ne i_2 \le l-1} b_{i_1 l}^2 b_{i_2 l}^2 \mathrm{E}(\widetilde{\mathbf{Y}}_l^\top \widetilde{\mathbf{Y}}_{i_1} \widetilde{\mathbf{Y}}_{i_1}^\top \widetilde{\mathbf{Y}}_l \widetilde{\mathbf{Y}}_l^\top \widetilde{\mathbf{Y}}_{i_2} \widetilde{\mathbf{Y}}_{i_2}^\top \widetilde{\mathbf{Y}}_l) \\
= \ & C + 3D.
\end{aligned}$$

Following similar argument as that in the verification of (S.2), we can show that

$$C/\sigma_p^4 = O\Big(N^{-2}\frac{\mathrm{E}(\widetilde{\mathbf{Y}}_1^\top \widetilde{\mathbf{Y}}_2)^4}{(\mathbf{q}^\top \mathbf{q})^2}\Big) \text{ and } D/\sigma_p^4 = O\Big(N^{-1}\frac{\mathrm{E}\{(\widetilde{\mathbf{Y}}_1^\top \boldsymbol{\Sigma} \widetilde{\mathbf{Y}}_1)^2\}}{(\mathbf{q}^\top \mathbf{q})^2}\Big).$$

Again by Lemma S.1 and Assumptions (A3)–(A4), we obtain $(C+3D)/\sigma_N^4 = o(1)$, from which Theorem 2 follows. □

<u>Proof of part (iii)</u> It suffices to show that, under H_0, $\Pr(E_{p,\mathscr{A}} = 0) \to 1$ as $(p, N) \to \infty$, i.e.,

$$\Pr\left(\max_{\tau \in \mathcal{T}} \max_{j \in \mathscr{A}_0} R_{\tau j} > r_p\right) \to 0.$$

Let $\tilde{R}_{\tau j} = L_{\tau j}/q_{0j}$. First note that, for any $s_p > 0$,

$$\Pr(\max_{\tau\in\mathcal{T}} \max_{j\in\mathscr{A}_0} R_{\tau j} > r_p) \le \Pr(\max_{\tau\in\mathcal{T}} \max_{j\in\mathscr{A}_0} \tilde{R}_{\tau j} > r_p/s_p) + \Pr\left(\max_{j\in\mathscr{A}_0} \frac{q_{0j}}{\hat{q}_j} > s_p\right).$$

On the one hand, for $s_p = 2$,

$$\begin{aligned}
\Pr\left(\max_{j\in\mathscr{A}_0} \frac{q_{0j}}{\hat{q}_j} > s_p\right) &= \Pr\left\{\max_{j\in\mathscr{A}_0}\left(1 - \frac{\hat{q}_j}{q_{0j}}\right) > 1 - s_p^{-1}\right\} \\
&\le \sum_{j\in\mathscr{A}_0} \Pr\left\{\sum_{i=1}^{N}(-\tilde{Y}_{ij}) > (1 - s_p^{-1}) N q_{0j}\right\} \\
&\le \sum_{j\in\mathscr{A}_0} \exp\left\{-\frac{N q_{0j}}{8(1-q_{0j}) + 4/3}\right\} \\
&\to 0,
\end{aligned}$$

by Lemma S.2 and the fact that $Na_p^{-1}(\log a_p)^{-1} \to \infty$. On the other hand,

$$\Pr\Big(\max_{\tau\in\mathcal{T}}\max_{j\in\mathscr{A}_0}\tilde{R}_{\tau j} > r_p/2\Big) \le \sum_{\tau\in\mathcal{T}}\sum_{j\in\mathscr{A}_0}\Pr\Big(\Big|\sum_{i=1}^{N}\sqrt{c_\tau}a_{i,\tau}\tilde{Y}_{ij}\Big| > \sqrt{r_p/2}q_{0j}^{1/2}\Big) \to 0,$$

again by Lemma S.2 and Assumptions (A2) and (A5). □

Proof of Proposition 1. To show that $U_{N,\mathscr{A}}/\mathbf{q}^\top\mathbf{q} \xrightarrow{\mathcal{P}} 1$, we let $\mathbf{Y}_{i,\mathscr{B}}$ be the collection of all Y_{ij} with $j\in\mathscr{B}$ and thus we write $U_{N,\mathscr{A}} = \{N(N-1)\}^{-1}\sum_{i\neq k}\mathbf{Y}_{i,\mathscr{B}}^\top\mathbf{Y}_{k,\mathscr{B}}$. It suffices to note that $\mathrm{E}(U_{N,\mathscr{A}}) = \mathbf{q}^\top\mathbf{q}$ and

$$\mathrm{E}(U_{N,\mathscr{A}}^2)/(\mathbf{q}^\top\mathbf{q})^2 = \frac{(N)_4}{(N)_2^2} + O\Big\{\frac{\sum_{j\in\mathscr{B}}q_{0j}^3}{N(\mathbf{q}^\top\mathbf{q})^2}\Big\} + O\Big(\frac{1}{N^2\mathbf{q}^\top\mathbf{q}}\Big),$$

where $(x)_m = x(x-1)\cdots(x-m+1)$ denotes the Pochhammer symbol. By Assumption (A3), $\mathrm{E}\{(U_{N,p}/\mathbf{q}^\top\mathbf{q}-1)^2\}\to 0$ as $(p,N)\to\infty$. □

Proof of Theorem 3. Proof of part (i) Let $\tilde{\mathbf{p}}_i \equiv \tilde{\mathbf{p}}_{i,\mathscr{B}}$ be the collection of $\tilde{p}_{ij}$ with $j\in\mathscr{B}$. Similar to the arguments in the proof of Theorem 2, we first write $S_p \equiv S_{p,\mathscr{A}}$ as

$$S_p = -\epsilon_p + S_{p,u} - 2S_{p,v} + 2S_{p,x} + s_p,$$

where $\epsilon_p = \sum_{i=1}^N b_{ii}\tilde{\mathbf{p}}_i^\top\tilde{\mathbf{p}}_i$, $S_{p,u} = \sum_{1\le i\neq k\le N} b_{ik}\widetilde{\mathbf{Y}}_i^\top\widetilde{\mathbf{Y}}_k$, $S_{p,v} = \sum_{i=1}^N b_{ii}\widetilde{\mathbf{Y}}_i^\top\tilde{\mathbf{p}}_i$,

$$S_{p,x} = \sum_{i=1}^N\Big(\sum_{\tau=\lceil a(T-1)\rceil}^{\tau^*} c_\tau a_{i,\tau}\frac{N_{1\tau^*}}{N_{1\tau}} + \sum_{\tau=\tau^*+1}^{\lceil b(T-1)\rceil} c_\tau a_{i,\tau}\frac{N_{0\tau^*}}{N_{0\tau}}\Big)\widetilde{\mathbf{Y}}_i^\top\boldsymbol{\delta}_{\mathscr{B}},$$

$$s_p = \Big(\sum_{\tau=\lceil a(T-1)\rceil}^{\tau^*}\frac{N_{0\tau}}{N_{1\tau}}\frac{N_{1\tau^*}^2}{N^2} + \sum_{\tau=\tau^*+1}^{\lceil b(T-1)\rceil}\frac{N_{1\tau}}{N_{0\tau}}\frac{N_{0\tau^*}^2}{N^2}\Big)\cdot N\sum_{j\in\mathscr{B}}\delta_j^2,$$

and $\boldsymbol{\delta}_{\mathscr{B}}$ denotes the collection of all δ_j's with $j\in\mathscr{B}$.

Let $\mathbf{q}_a$ be the collection of all q_{1j}'s with $j\in\mathscr{B}$ and $\boldsymbol{\Sigma}_a = \mathrm{diag}(\mathbf{q}_a) - \mathbf{q}_a\mathbf{q}_a^\top$. By Assumption (A2), $\tau^* = \lceil\gamma(T-1)\rceil$ and $N_{0\tau^*}/N\to\kappa_0$, we have

$$\begin{aligned}
\epsilon_p &= O[T\{\kappa_0\mathbf{q}^\top\mathbf{q} + (1-\kappa_0)\mathbf{q}_a^\top\mathbf{q}_a\}],\\
\mathrm{Var}(S_{p,u}) &= O\{T^2\mathrm{tr}(\boldsymbol{\Sigma}_{\kappa_0}^2)\},\\
\mathrm{Var}(S_{p,v}) &= O[T^2N^{-1}\{\kappa_0\mathbf{q}^\top\boldsymbol{\Sigma}\mathbf{q} + (1-\kappa_0)\mathbf{q}_a^\top\boldsymbol{\Sigma}_a\mathbf{q}_a\}],\\
\mathrm{Var}(S_{p,x}) &= O(T^2N\boldsymbol{\delta}_{\mathscr{B}}^\top\boldsymbol{\Sigma}_{\kappa_0}\boldsymbol{\delta}_{\mathscr{B}}),\\
s_p &= O(TN\boldsymbol{\delta}_{\mathscr{B}}^\top\boldsymbol{\delta}_{\mathscr{B}}),
\end{aligned}$$

where $\mathbf{\Sigma}_{\kappa_0} = \kappa_0\mathbf{\Sigma} + (1-\kappa_0)\mathbf{\Sigma}_a$. By Assumption (A1), it is straightforward to show that

$$\mathrm{tr}(\mathbf{\Sigma}_{\kappa_0}^2) = \mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}\{1+o(1)\} \text{ and } \kappa_0\mathbf{q}^\top\mathbf{\Sigma}\mathbf{q} + (1-\kappa_0)\mathbf{q}_a^\top\mathbf{\Sigma}_a\mathbf{q}_a = o(\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}),$$

where $\mathbf{q}_{\kappa_0} = \kappa_0\mathbf{q} + (1-\kappa_0)\mathbf{q}_a$. Hence, we conclude that

$$\begin{aligned}\mathrm{Var}(S_{p,u} - 2S_{p,v}) &= \mathrm{Var}(S_{p,u}) + 4\mathrm{Var}(S_{p,v}) \\ &= O[T^2\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}\{1+o(1)\}],\end{aligned}$$

by noting that $\mathrm{Cov}(S_{p,u}, S_{p,v}) = 0$, which results in that $S_{p,u} - 2S_{p,v} = O_p\left(T\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}}\right)$. Obviously, $\epsilon_p = o\left(T\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}}\right)$, and furthermore since $N\boldsymbol{\delta}_{\mathscr{B}}^\top\boldsymbol{\delta}_{\mathscr{B}}/(\max\limits_{l}\sum_{j\in\mathscr{B}} q_{lj}^2)^{1/2} \to \infty$, we have $N\boldsymbol{\delta}_{\mathscr{B}}^\top\boldsymbol{\delta}_{\mathscr{B}}/\max\limits_{l;j\in\mathscr{B}} q_{lj} \to \infty$ and $S_{p,x} = o_p(s_p)$. Now it is clear that $S_p = s_p\{1+o_p(1)\} + O_p\left(T\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}}\right)$.

Next we focus on the asymptotic analysis of $U_{N,\mathscr{A}}$ under the alternative. Tedious yet straightforward calculations yield that

$$\begin{aligned}&\mathrm{E}(U_{N,\mathscr{A}}) = \mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0} + o(1), \\ &\mathrm{E}(U_{N,\mathscr{A}}^2) = (\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0})^2 + 4N^{-1}\mathbf{q}_{\kappa_0}^\top\mathbf{D}_{\kappa_0}\mathbf{q}_{\kappa_0} + 2N^{-2}\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0} + o(1),\end{aligned}$$

since $N_{0\tau^*}/N \to \kappa_0$, where $\mathbf{D}_{\kappa_0} = \kappa_0\mathrm{diag}(\mathbf{q}) + (1-\kappa_0)\mathrm{diag}(\mathbf{q}_a)$. By Assumption (A3), we can show that

$$\mathrm{E}\{(U_{N,\mathscr{A}}/\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0} - 1)^2\} \to 0,$$

and thus $U_{N,\mathscr{A}}/\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0} \xrightarrow{\mathcal{P}} 1$. Also, by Assumption (A2), we have $c_{N,T} = O(T^2)$. Combining these results, we have

$$\begin{aligned}\frac{S_{p,\mathscr{A}}}{\sqrt{2c_{N,T}U_{N,\mathscr{A}}}} &= O\left[\frac{N\boldsymbol{\delta}_{\mathscr{B}}^\top\boldsymbol{\delta}_{\mathscr{B}}}{\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}}}\{1+o_p(1)\} + O_p(1)\right]\Bigg/\sqrt{\frac{U_{N,\mathscr{A}}}{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}}} \\ &\xrightarrow{\mathcal{P}} \infty,\end{aligned}$$

as $(p,N) \to \infty$, by the assumption that $N\boldsymbol{\delta}_{\mathscr{B}}^\top\boldsymbol{\delta}_{\mathscr{B}}/(\max\limits_{l}\sum_{j\in\mathscr{B}} q_{lj}^2)^{1/2} \to \infty$.

<u>Proof of part (ii)</u> We observe that

$$\Pr\left(\max_{\tau\in\mathcal{T}}\max_{j\in\mathscr{A}} R_{\tau j} > r_p\right) \geq \Pr(R_{\tau^* j'} > r_p).$$

It is straightforward to show that

$$\begin{aligned}\sqrt{\frac{N_{0\tau^*}N_{1\tau^*}}{N}}\Big(\frac{Z_{0\tau^*j'}}{N_{0\tau^*}}-\frac{Z_{1\tau^*j'}}{N_{1\tau^*}}\Big) &= \sqrt{c_{\tau^*}}\sum_{i=1}^N a_{i,\tau^*}\tilde{Y}_{ij}+\sqrt{c_{\tau^*}}\delta_{j'} \\ &= O_p\Big\{\Big(\sum_{i=1}^N c_\tau^* a_{i,\tau^*}^2\tilde{p}_{ij}\Big)^{1/2}\Big\}+\sqrt{c_{\tau^*}}\delta_{j'}\end{aligned}$$

and $\hat{q}_{j'}/q_{*j'} \xrightarrow{\mathcal{P}} 1$ since $Na_p^{-1}\to\infty$ by Assumption (A1). Hence, by Assumption (A2) and the condition that $N_{0\tau^*}/N\to\kappa_0$,

$$\begin{aligned}&\sqrt{\frac{N_{0\tau^*}N_{1\tau^*}}{N}}\Big(\frac{Z_{0\tau^*j'}}{N_{0\tau^*}}-\frac{Z_{1\tau^*j'}}{N_{1\tau^*}}\Big)\Big/\hat{q}_{j'}^{1/2} \\ = \ & \frac{O_p\Big\{\Big(\sum_{i=1}^N c_\tau^* a_{i,\tau^*}^2\tilde{p}_{ij}\Big)^{1/2}\Big\}+\sqrt{c_{\tau^*}}\delta_{j'}}{q_{*j'}^{1/2}\{1+o_p(1)\}} \\ = \ & O_p(1)+\frac{\sqrt{c_{\tau^*}}\delta_{j'}}{q_{*j'}^{1/2}}\{1+o_p(1)\} \\ = \ & O_p(1)+\frac{\{\kappa_0(1-\kappa_0)\}^{1/2}N^{1/2}\delta_{j'}}{\{\kappa_0 q_{0j}+(1-\kappa_0)q_{1j}\}^{1/2}}\{1+o_p(1)\}.\end{aligned}$$

Since $N\delta_{j'}^2 r_p^{-1}\to\infty$, we conclude that $\Pr(R_{\tau^*j'}>r_p)\to 1$. □

Proof of Corollary 1. This proof contains two parts (i)–(ii).

(i) If there exists some $j'\in\mathscr{A}$ such that $\delta_j'\neq 0$ and further we assume that $N\delta_{j'}^2 r_p^{-1}\to\infty$, from the proof of Theorem 3, we know $\Pr(E_{p,\hat{\mathscr{A}}}=e_p)\to 1$.

Assume $\tau<\tau^*$. Following similar steps in the proof of Theorem 3, we can show that, for any $\tau\in\mathcal{T}$ and $j\in\mathscr{A}$,

$$\begin{aligned}\sqrt{\frac{N_{0\tau}N_{1\tau}}{N}}\Big(\frac{Z_{0\tau j}}{N_{0\tau}}-\frac{Z_{1\tau j}}{N_{1\tau}}\Big) &= \sqrt{c_\tau}\sum_{i=1}^N a_{i,\tau}\tilde{Y}_{ij}+\sqrt{c_\tau}\frac{N_{1\tau^*}}{N_{1\tau}}\delta_j \\ &= O_p\Big\{\Big(\sum_{i=1}^N c_\tau a_{i,\tau}^2\tilde{p}_{ij}\Big)^{1/2}\Big\}+\sqrt{c_\tau}\frac{N_{1\tau^*}}{N_{1\tau}}\delta_j \\ &= O_p\{(\max_{lj} q_{lj})^{1/2}\}+\sqrt{c_\tau}\frac{N_{1\tau^*}}{N_{1\tau}}\delta_j\end{aligned}$$

and $\hat{q}_j/q_{*j}\xrightarrow{\mathcal{P}} 1$ if $Na_p^{-1}\to\infty$ by Assumption (A1), where $q_{*j}=\frac{N_{0\tau^*}}{N}q_{0j}+$

$\frac{N_{1\tau^*}}{N} q_{1j}$. Hence, by Assumption (A2),

$$\begin{aligned}\sqrt{\frac{N_{0\tau}N_{1\tau}}{N}}\Big(\frac{Z_{0\tau j}}{N_{0\tau}} - \frac{Z_{1\tau j}}{N_{1\tau}}\Big)\Big/\hat{q}_j^{1/2} &= \frac{O_p\Big\{\Big(\sum_{i=1}^N c_\tau a_{i,\tau}^2 \tilde{p}_{ij}\Big)^{1/2}\Big\} + \sqrt{c_\tau}\frac{N_{1\tau^*}}{N_{1\tau}}\delta_j}{q_{*j}^{1/2}\{1+o_p(1)\}} \\ &= O_p(1) + \frac{\sqrt{c_\tau}\frac{N_{1\tau^*}}{N_{1\tau}}\delta_j}{q_{*j}^{1/2}}\{1+o_p(1)\}.\end{aligned}$$

We aim to demonstrate that $\Pr(\max_{j\in\mathscr{A}} R_{\tau j} < \max_{j\in\mathscr{A}} R_{\tau^* j}) \to 1$. Denote $j_\tau = \arg\max_{j\in\mathscr{A}} R_{\tau j}$. We claim that $N\delta_{j_\tau}^2 r_p^{-1} \to \infty$, since if $\delta_{j_\tau}^2 = o(\delta_{j'}^2)$,

$$\begin{aligned}\Pr(R_{\tau j_\tau} > R_{\tau j'}) &= \Pr\Big[\frac{o(\delta_{j'}^2)}{\hat{q}_{j_\tau}} > \frac{\delta_{j'}^2\{1+o_p(1)\}}{\hat{q}_{j'}}\Big] \\ &= \Pr\Big[o_p(1) > \frac{q_{*j_\tau}\{1+o_p(1)\}}{q_{*j'}\{1+o_p(1)\}}\Big] \\ &\le \Pr\{o_p(1) > C^*\varepsilon a_p^{-1}\} \to 0,\end{aligned}$$

by Assumptions (A1)–(A2) and (A5), where $C^* = \min(N_{0\tau^*}, N_{1\tau^*})$ and $a_p^{-1} = O(1)$. Then noting that

$$\begin{aligned}R_{\tau j_\tau} - R_{\tau^* j_\tau} &= \frac{\Big(c_\tau\frac{N_{1\tau^*}^2}{N_{1\tau}^2} - c_\tau^*\Big)\delta_{j_\tau}^2\{1+o_p(1)\}}{\hat{q}_{j_\tau}} \\ &= \frac{\Big(\frac{N_{0\tau}}{N_{0\tau^*}}\frac{N_{1\tau^*}}{N_{1\tau}} - 1\Big)\kappa_0(1-\kappa_0)N\delta_{j_\tau}^2\{1+o_p(1)\}}{\hat{q}_{j_\tau}},\end{aligned}$$

we have $\Pr(R_{\tau j_\tau} < R_{\tau^* j_\tau}) \to 1$ if

$$\lim_{N\to\infty}(N_{0\tau}/N - \kappa_0)N\delta_{j'}^2 \to \infty.$$

Hence $\Pr(R_{\tau j_\tau} < \max_{j\in\mathscr{A}} R_{\tau^* j}) \ge \Pr(R_{\tau j_\tau} < R_{\tau^* j_\tau}) \to 1$. A similar conclusion can be drawn if $\tau > \tau^*$.

(ii) If all $\delta_j = 0$ for $j \in \mathscr{A}$, from the proof of Theorem 2 (iii), we know $\Pr(E_{p,\hat{\mathscr{A}}} = e_p) \to 0$. Denote $M_\tau = \sum_{j\in\mathscr{B}}(L_{\tau j} - L_{\tau j}^{(0)})$. Following similar steps in the proof of Theorem 3, we can show that

$$M_\tau = c_\tau \frac{N_{1\tau^*}^2}{N_{1\tau}^2}\sum_{j\in\mathscr{B}}\delta_j^2\{1+o_p(1)\} + O_p\Big(\sqrt{\mathbf{q}_{\kappa_0}^\top \mathbf{q}_{\kappa_0}}\Big)$$

and thus

$$M_\tau - M_\tau^* = \Big(\frac{N_{0\tau}}{N_{0\tau^*}}\frac{N_{1\tau^*}}{N_{1\tau}} - 1\Big)\kappa_{00}(1-\kappa_{00})N\sum_{j\in\mathscr{B}}\delta_j^2\{1+o_p(1)\} + O_p\Big(\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}}\Big).$$

Hence, with probability tending to one, $M_\tau < M_\tau^*$ if $\tau \le \tau^* - \zeta_{\mathscr{B},T}$, where $\zeta_{\mathscr{B},T} > 0$ satisfies

$$(\lim_{N\to\infty} N_{0,\tau^*-\zeta_{\mathscr{B},T}}/N - \kappa_{00})N\sum_{j\in\mathscr{B}}\delta_j^2/\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}} \to \infty,$$

as $(p,N)\to\infty$. Similarly, with probability tending to one, $M_\tau < M_\tau^*$ if $\tau \ge \tau^* + \zeta_{\mathscr{B},T}$, where $\zeta_{\mathscr{B},T} > 0$ satisfies

$$(\lim_{N\to\infty} N_{0,\tau^*+\zeta_{\mathscr{B},T}}/N - \kappa_{00})N\sum_{j\in\mathscr{B}}\delta_j^2/\sqrt{\mathbf{q}_{\kappa_0}^\top\mathbf{q}_{\kappa_0}} \to \infty,$$

as $(p,N)\to\infty$. That is, the probability that M_τ attains its maximum at τ such that $|\tau-\tau^*| \ge \zeta_{\mathscr{B},T}$ goes to zero as $(p,N)\to\infty$. □

Proof of Theorem 5. Given $a = 0,1,\dots,A^*$, we rewrite $\tau_b^0 = \tau^*_{\mathscr{B},a,b}, b = 0,1,\dots,B_a^*$ for ease of notation. Then for B_a candidate change-points $\{\tau_b\}_{b=1}^{B_a}$ satisfying that $\tau^*_{\mathscr{A},a} \equiv \tau_0 < \tau_1 < \cdots < \tau_{B_a} < \tau_{B_a+1} \equiv \tau^*_{\mathscr{A},a+1}$, let

$$\mathcal{S}_{\mathscr{B}}(\tau_1,\dots,\tau_{B_a}) = \sum_{b=0}^{B_a}\sum_{t=\tau_b+1}^{\tau_{b+1}}\sum_{j\in\mathscr{B}}\{X_{tj} - n_t\bar{X}_j(\tau_b,\tau_{b+1})\}^2/n_t,$$

and

$$\mathcal{S}_{B_a}^{\mathrm{Pen}} = \mathcal{S}_{\mathscr{B}}(\hat\tau_{B_a,1},\dots,\hat\tau_{B_a,B_a}) + B_a\{\hat{Q}_{\mathscr{B}}(\tau_0,\tau_{B_a+1}) + \eta_{p,N}\},$$

where $(\hat\tau_{B_a,1},\dots,\hat\tau_{B_a,B_a}) = \operatorname*{argmin}_{\tau_1<\cdots<\tau_{B_a}} \mathcal{S}_{\mathscr{B}}(\tau_1,\dots,\tau_{B_a})$ and $\hat{Q}_{\mathscr{B}}(\tau_0,\tau_{B_a+1}) = \sum_{t=\tau_0}^{\tau_{B_a+1}}\sum_{j\in\mathscr{B}}X_{tj}/\sum_{t=\tau_0}^{\tau_{B_a+1}} n_t$. By Corollary 3 and Theorem 4, and under the condition (10) in Theorem 5, it suffices to demonstrate that,

$$\Pr\big(\hat{B}_a = B_a^*;\ |\hat\tau_{\hat{B}_a,b} - \tau_b^0| \le \delta_T, b = 0,1,\dots,B_a^*\big) \to 1, \tag{S.3}$$

where $\hat{B}_a = \operatorname{argmin}_{B_a}\mathcal{S}_{B_a}^{\mathrm{Pen}}$. We further extend the definition of $\mathcal{S}_{\mathscr{B}}(\cdot)$ for a set of candidate change-points $\{\nu_i\}_{i=1}^I \subset (\tau_0, \tau_{B_a+1}]$,

$$\mathcal{S}_{\mathscr{B}}(\nu_1,\dots,\nu_I) = \sum_{i=0}^{I}\sum_{t=\nu_{(i)}+1}^{\nu_{(i+1)}}\sum_{j\in\mathscr{B}}[X_{tj} - n_t\bar{X}_j\{\nu_{(i)},\nu_{(i+1)}\}]^2/n_t,$$

where $\nu_{(0)} = \tau_0$, $\nu_{(I+1)} = \tau_{B_a+1}$ and $\nu_{(1)} < \cdots < \nu_{(I)}$ are the ordered version of ν_i's.

The proof of (S.3) contains three parts (see Propositions S.1, S.2 and S.3 below). Firstly, we demonstrate that the minimization of $\mathcal{S}_{B_a}^{\mathrm{Pen}}$ includes no less than B_a^* change-point estimators asymptotically.

Proposition S.1. *Under the conditions in Theorem 5,* $\Pr(\hat{B}_a \geq B_a^*) \to 1$ *as* $(p, N, T) \to \infty$.

Proof. Assume that $B_a^* > 0$. Define

$$\mathcal{S}_{B_a^*}^{0\mathrm{Pen}} = \mathcal{S}_{\mathscr{B}}(\tau_1^0, \ldots, \tau_{B_a^*}^0) + B_a^*\{\hat{Q}_{\mathscr{B}}(\tau_0, \tau_{B_a+1}) + \eta_{p,N}\}.$$

Let ρ_T be the maximum integer less than $\lambda_{\mathscr{B},T}/2$ and consider $0 < B_a < B_a^*$. We introduce, for $r = 1, \ldots, B_a^*$,

$$\Theta_r(\rho_T) = \{\tau_1 < \cdots < \tau_{B_a} : |\tau_b - \tau_r^0| > \rho_T \text{ for } 1 \leq b \leq B_a\}.$$

Since $B_a < B_a^*$, $(\hat{\tau}_{B_a,1}, \ldots, \hat{\tau}_{B_a,B_a})$ must belong to one of the $\Theta_r(\rho_T)$'s. For every $(\tau_1, \ldots, \tau_{B_a}) \in \Theta_r(\rho_T)$, with probability approaching one,

$$\begin{aligned} &\mathcal{S}_{\mathscr{B}}(\tau_1, \ldots, \tau_{B_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1, \ldots, \tau_{B_a}, \tau_1^0, \ldots, \tau_{r-1}^0, \tau_r^0 - \rho_T, \tau_r^0, \tau_r^0 + \rho_T, \tau_{r+1}^0, \ldots, \tau_{B_a^*}^0) \\ &\geq (B_a^* + 2)Q_{\mathscr{B}} + \frac{\rho_T^2 \underline{n}^2}{N}\Delta_{\mathscr{B}} + o_p(\eta_{p,N}), \end{aligned} \tag{S.4}$$

and

$$\begin{aligned} &\mathcal{S}_{\mathscr{B}}(\tau_1, \ldots, \tau_{B_a}, \tau_1^0, \ldots, \tau_{r-1}^0, \tau_r^0 - \rho_T, \tau_r^0, \tau_r^0 + \rho_T, \tau_{r+1}^0, \ldots, \tau_{B_a^*}^0) \\ &\quad - \mathcal{S}_{\mathscr{B}}(\tau_1^0, \ldots, \tau_{B_a^*}^0) \geq -(B_a + 2)Q_{\mathscr{B}} + o_p(\eta_{p,N}), \end{aligned} \tag{S.5}$$

under Assumptions (B1)–(B2). Hence, with probability approaching one,

$$\mathcal{S}_{\mathscr{B}}(\hat{\tau}_{B_a,1}, \ldots, \hat{\tau}_{B_a,B_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1^0, \ldots, \tau_{B_a^*}^0) \geq -(B_a - B_a^*)Q_{\mathscr{B}} + \frac{\rho_T^2 \underline{n}^2}{N}\Delta_{\mathscr{B}} + o_p(\eta_{p,N}).$$

As a by-product in the proof of (S.4), we have

$$\hat{Q}_{\mathscr{B}}(\tau_0, \tau_{B_a+1}) - Q_{\mathscr{B}} = o_p(\eta_{p,N}).$$

Then, we conclude that

$$\begin{aligned} & \Pr(\mathcal{S}_{B_a}^{\mathrm{Pen}} > \mathcal{S}_{B_a^*}^{\mathrm{Pen}}) \geq \Pr(\mathcal{S}_{B_a}^{\mathrm{Pen}} > \mathcal{S}_{B_a^*}^{0\mathrm{Pen}}) \\ = \ & \Pr\{\mathcal{S}_{\mathscr{B}}(\hat{\tau}_{B_a,1}, \ldots, \hat{\tau}_{B_a,B_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1^0, \ldots, \tau_{B_a^*}^0) + (B_a - B_a^*)\hat{Q}_{\mathscr{B}}(\tau_0, \tau_{B_a+1}) \\ & +(B_a - B_a^*)\eta_{p,N} > 0\} \\ \geq \ & \Pr\left[\frac{\rho_T^2 \underline{n}^2}{N}\Delta_{\mathscr{B}} + o_p(\eta_{p,N}) + (B_a - B_a^*)\{\hat{Q}_{\mathscr{B}}(\tau_0, \tau_{B_a+1}) - Q_{\mathscr{B}}\} + (B_a - B_a^*)\eta_{p,N} > 0\right] \\ \to \ & 1, \end{aligned}$$

if $\lambda^2_{\mathscr{B},T}\underline{n}^2N^{-1}\Delta_{\mathscr{B}}/\eta_{p,N} \to \infty$. Thus, the proposition follows provided that (S.4)–(S.5) hold.

Now we consider the claim (S.4). Let $\mathcal{R}_j(s,e) = \sum_{t=s+1}^{e}\{X_{tj} - n_t\bar{X}_j(s,e)\}^2/n_t$ and $N_{t_1,t_2} = \sum_{t=t_1+1}^{t_2} n_t$. For every $(\tau_1,\ldots,\tau_{B_a}) \in \Theta_r(\rho_T)$, we have

$$
\begin{aligned}
&\mathcal{S}_{\mathscr{B}}(\tau_1,\ldots,\tau_{B_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1,\ldots,\tau_{B_a},\tau_1^0,\ldots,\tau_{r-1}^0,\tau_r^0-\rho_T,\tau_r^0,\tau_r^0+\rho_T,\tau_{r+1}^0,\ldots,\tau_{B_a^*}^0)\\
=&\sum_{b=0}^{B_a}\left\{\sum_{j\in\mathscr{B}}\mathcal{R}_j(\tau_b,\tau_{b+1}) - \sum_{i=0}^{M_b}\sum_{j\in\mathscr{B}}\mathcal{R}_j(\nu_{b,i},\nu_{b,i+1})\right\}\\
=&\sum_{b=0}^{B_a}\sum_{j\in\mathscr{B}}\left[\left\{\sum_{i=0}^{M_b}\mathcal{R}_j(\nu_{b,i},\nu_{b,i+1}) + \sum_{0\le i<k\le M_b}\frac{N_{\nu_{b,i},\nu_{b,i+1}}+N_{\nu_{b,k},\nu_{b,k+1}}}{N_{\tau_b,\tau_{b+1}}}L_{ik,j}^{(b)}\right\}\right.\\
&\qquad\qquad\left. - \sum_{i=0}^{M_b}\mathcal{R}_j(\nu_{b,i},\nu_{b,i+1})\right]\\
=&\sum_{b=0}^{B_a}\sum_{0\le i<k\le M_b}\frac{N_{\nu_{b,i},\nu_{b,i+1}}+N_{\nu_{b,k},\nu_{b,k+1}}}{N_{\tau_b,\tau_{b+1}}}\sum_{j\in\mathscr{B}}L_{ik,j}^{(b)},
\end{aligned}
$$

where $\{\nu_{b,i}\}_{i=1}^{M_b}$ satisfy $\tau_b \equiv \nu_{b,0} < \nu_{b,1} < \cdots < \nu_{b,M_b} < \nu_{b,M_b+1} \equiv \tau_{b+1}$ and take values in $\{\tau_1^0,\ldots,\tau_{r-1}^0,\tau_r^0-\rho_T,\tau_r^0,\tau_r^0+\rho_T,\tau_{r+1}^0,\ldots,\tau_{B_a^*}^0\}$, and

$$
L_{ik,j}^{(a)} = \frac{N_{\nu_i^a,\nu_{i+1}^a}N_{\nu_k^a,\nu_{k+1}^a}}{N_{\nu_i^a,\nu_{i+1}^a}+N_{\nu_k^a,\nu_{k+1}^a}}\{\bar{X}_j(\nu_i^a,\nu_{i+1}^a) - \bar{X}_j(\nu_k^a,\nu_{k+1}^a)\}^2.
$$

We observe that $Q_{i,\mathscr{B}} = Q_{k,\mathscr{B}}$ for any i,k and

$$
\sum_{b=0}^{B_a}\sum_{0\le i<k\le M_b}\frac{N_{\nu_{b,i},\nu_{b,i+1}}+N_{\nu_{b,k},\nu_{b,k+1}}}{N_{\tau_b,\tau_{b+1}}} = \sum_{b=0}^{B_a}M_b = B_a^*+2.
$$

By Lemma S.4 below, we have that uniformly,

$$
\begin{aligned}
&\mathcal{S}_{\mathscr{B}}(\tau_1,\ldots,\tau_{B_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1,\ldots,\tau_{B_a},\tau_1^0,\ldots,\tau_{r-1}^0,\tau_r^0-\rho_T,\tau_r^0,\tau_r^0+\rho_T,\tau_{r+1}^0,\ldots,\tau_{B_a^*}^0)\\
\ge&(B_a^*+2)Q_{\mathscr{B}} + \frac{N_{\tau_r^0-\rho_T,\tau_r^0}N_{\tau_r^0,\tau_r^0+\rho_T}}{N}\sum_{j\in\mathscr{B}}(q_{rj}-q_{r-1,j})^2 + o_p(\eta_{p,N})\\
\ge&(B_a^*+2)Q_{\mathscr{B}} + \frac{(\rho_T\underline{n})^2}{N}\min_{1\le r\le L^*}\sum_{j\in\mathscr{B}}(q_{rj}-q_{r-1,j})^2 + o_p(\eta_{p,N}),
\end{aligned}
$$

where $N_{t_1,t_2} = \sum_{t=t_1+1}^{t_2} n_t$. Hence, the claim (S.4) follows.

Similarly, it is straightforward to show that (S.5) holds. □

LEMMA S.4. *Assume the conditions in Theorem 5 hold, then for any* $1 \le i, k \le L^*$,

$$\begin{aligned}
&\sup_{\substack{\tau_i^* < t_1 < t_2 \le \tau_{i+1}^* \\ \tau_k^* < t_3 < t_4 \le \tau_{k+1}^*}} \sum_{j \in \mathscr{B}} \frac{N_{t_1,t_2} N_{t_3,t_4}}{N_{t_1,t_2} + N_{t_3,t_4}} \{\bar{X}_j(t_1,t_2) - \bar{X}_j(t_3,t_4)\}^2 \\
=& \frac{N_{t_1,t_2} N_{t_3,t_4}}{N_{t_1,t_2} + N_{t_3,t_4}} \left(\frac{Q_{i,\mathscr{B}}}{N_{t_1,t_2}} + \frac{Q_{k,\mathscr{B}}}{N_{t_3,t_4}} \right) + \frac{N_{t_1,t_2} N_{t_3,t_4}}{N_{t_1,t_2} + N_{t_3,t_4}} \|\mathbf{q}_{i,\mathscr{B}} - \mathbf{q}_{k,\mathscr{B}}\|^2 \{1 + o_p(1)\} \\
&+ o_p(\eta_{p,N}).
\end{aligned}$$

where $Q_{i,\mathscr{B}} = \sum_{j \in \mathscr{B}} q_{ij}$ *with* q_{ij} *being the* j*-th component of* $\mathbf{q}_i$*,*

PROOF. For ease of notation, we write $\mathcal{N}_i = N_{t_1,t_2}, \mathcal{N}_k = N_{t_3,t_4}$, $\mathcal{N}_{ik} = \mathcal{N}_i + \mathcal{N}_k$ and $c_{ik} = \mathcal{N}_i \mathcal{N}_k / \mathcal{N}$. We introduce $\mathbf{Y}_{i1}, \ldots, \mathbf{Y}_{i\mathcal{N}_i} \overset{\text{iid}}{\sim} \text{Multi}(1, \mathbf{q}_i)$, where $\mathbf{Y}_{ih} = (Y_{ih,1}, \ldots, Y_{ih,p})^\top$ for $h = 1, \ldots, \mathcal{N}_i$. Let $\mathbf{Y}_{ih,\mathscr{B}}$ and $\mathbf{q}_{i,\mathscr{B}}$ be the collections of $Y_{ih,j}$ and q_{ij} with $j \in \mathscr{B}$, respectively. Denote $\widetilde{\mathbf{Y}}_{ih,\mathscr{B}} = \mathbf{Y}_{ih,\mathscr{B}} - \mathbf{q}_{i,\mathscr{B}}$. Following similar arguments in the proof of Theorem 3, we can show that

$$\begin{aligned}
&\sum_{j \in \mathscr{B}} \frac{N_{t_1,t_2} N_{t_3,t_4}}{N_{t_1,t_2} + N_{t_3,t_4}} \{\bar{X}_j(t_1,t_2) - \bar{X}_j(t_3,t_4)\}^2 \\
=& c_{ik}(Q_{i,\mathscr{B}}/\mathcal{N}_i + Q_{k,\mathscr{B}}/\mathcal{N}_k) - \frac{\mathcal{N}_k}{\mathcal{N}_{ik}} \mathbf{q}_{i,\mathscr{B}}^\top \mathbf{q}_{i,\mathscr{B}} - \frac{\mathcal{N}_i}{\mathcal{N}_{ik}} \mathbf{q}_{k,\mathscr{B}}^\top \mathbf{q}_{k,\mathscr{B}} + \frac{\mathcal{N}_k}{\mathcal{N}_{ik}} d_{i,\mathscr{B}} + \frac{\mathcal{N}_i}{\mathcal{N}_{ik}} d_{k,\mathscr{B}} \\
&+ c_{ik} \boldsymbol{\delta}_{ik,\mathscr{B}}^\top \boldsymbol{\delta}_{ik,\mathscr{B}} + 2\sqrt{\frac{\mathcal{N}_k}{\mathcal{N}_{ik}}} \sqrt{c_{ik}} S_{i,\boldsymbol{\delta}_{ik,\mathscr{B}}} - 2\sqrt{\frac{\mathcal{N}_i}{\mathcal{N}_{ik}}} \sqrt{c_{ik}} S_{k,\boldsymbol{\delta}_{ik,\mathscr{B}}} \\
&+ \frac{\mathcal{N}_k}{\mathcal{N}_{ik}} U_{i,\mathscr{B}} + \frac{\mathcal{N}_i}{\mathcal{N}_{ik}} U_{k,\mathscr{B}} - \frac{2\sqrt{\mathcal{N}_i \mathcal{N}_k}}{\mathcal{N}_{ik}} V_{ik,\mathscr{B}} - 2\frac{\mathcal{N}_k}{\mathcal{N}_{ik}} \frac{W_{i,\mathscr{B}}}{\sqrt{\mathcal{N}_i}} - 2\frac{\mathcal{N}_i}{\mathcal{N}_{ik}} \frac{W_{k,\mathscr{B}}}{\sqrt{\mathcal{N}_k}},
\end{aligned}$$

where

$$
\begin{aligned}
d_{i,\mathscr{B}} =& \mathcal{N}_i^{-1} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{1}_{\mathscr{B}} \\
S_{i,\boldsymbol{\delta}_{ik,\mathscr{B}}} =& \mathcal{N}_i^{-1/2} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \boldsymbol{\delta}_{ik,\mathscr{B}} \\
U_{i,\mathscr{B}} =& \mathcal{N}_i^{-1} \sum_{1 \le h_1 \ne h_2 \le \mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih_1,\mathscr{B}}^\top \widetilde{\mathbf{Y}}_{ih_2,\mathscr{B}} \\
V_{ik,\mathscr{B}} =& (\mathcal{N}_i \mathcal{N}_k)^{-1/2} \sum_{h_1=1}^{\mathcal{N}_i} \sum_{h_2=1}^{\mathcal{N}_k} \widetilde{\mathbf{Y}}_{ih_1,\mathscr{B}}^\top \widetilde{\mathbf{Y}}_{kh_2,\mathscr{B}} \\
W_{i,\mathscr{B}} =& \mathcal{N}_i^{-1/2} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{q}_{i,\mathscr{B}}
\end{aligned}
$$

and $\boldsymbol{\delta}_{ik,\mathscr{B}} = \mathbf{q}_{i,\mathscr{B}} - \mathbf{q}_{k,\mathscr{B}}$.

By Lemmas S.5–S.9, we conclude that

$$
\begin{aligned}
&\sum_{j \in \mathscr{B}} \frac{N_{t_1,t_2} N_{t_3,t_4}}{N_{t_1,t_2} + N_{t_3,t_4}} \{\bar{X}_j(t_1,t_2) - \bar{X}_j(t_3,t_4)\}^2 \\
=& c_{ik}(Q_{i,\mathscr{B}}/\mathcal{N}_i + Q_{k,\mathscr{B}}/\mathcal{N}_k) + c_{ik} \boldsymbol{\delta}_{ik,\mathscr{B}}^\top \boldsymbol{\delta}_{ik,\mathscr{B}} \{1 + o_p(1)\} + o_p(\eta_{p,N}).
\end{aligned}
$$

□

Proposition S.2 shows that the global minimum of $\mathcal{S}_{B_a}^{\text{Pen}}$ includes no more than B_a^* change-point estimators asymptotically.

PROPOSITION S.2. *Assume the conditions in Theorem 5 hold, then* $\Pr(\hat{B}_a > B_a^*) \to 0$.

PROOF. Let $\overline{L}$ be the upper bound on L^*. First assume that $B_a^* > 0$. It suffices to show that, for all $B_a^* < B_a \le \overline{L}$, $\Pr(\mathcal{S}_{B_a}^{\text{Pen}} > \mathcal{S}_{B_a^*}^{\text{Pen}}) \to 1$. Based on the similar argument in the proof of Proposition S.1, we have

$$
\begin{aligned}
& \Pr(\mathcal{S}_{B_a}^{\text{Pen}} > \mathcal{S}_{B_a^*}^{\text{Pen}}) \ge \Pr(\mathcal{S}_{B_a}^{\text{Pen}} > \mathcal{S}_{B_a^*}^{0\text{Pen}}) \\
= \;& \Pr\{\mathcal{S}_{\mathscr{B}}(\hat{\tau}_{B_a,1}, \ldots, \hat{\tau}_{B_a,B_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1^0, \ldots, \tau_{B_a^*}^0) + (B_a - B_a^*)\hat{Q}_{\mathscr{B}}(\tau_0, \tau_{B_a+1}) \\
& \qquad + (B_a - B_a^*)\eta_{p,N} > 0\} \\
\ge \;& \Pr\left\{o_p(\eta_{p,N}) + (B_a - B_a^*)\eta_{p,N} > 0\right\} \\
\to \;& 1.
\end{aligned}
$$

If $B_a^* = 0$, and similarly, for any $B_a > 0$, we have $\Pr(\mathcal{S}_{B_a}^{\text{Pen}} > \mathcal{S}_0^{\text{Pen}}) \to 1$. Thus, the result follows. $\square$

PROPOSITION S.3. *Assume the conditions in Theorem 5 hold and if $B_a^* > 0$, then*

$$\Pr\left\{|\hat{\tau}_{\hat{B}_a,b} - \tau_b^0| \leq \delta_{\mathscr{B},T}, b = 1, \ldots, B_a^*\right\} \to 1.$$

PROOF. If there exists some $r = 1, \ldots, B_a^*$ such that $|\hat{\tau}_{\hat{B}_a,r} - \tau_r^0| > \delta_{\mathscr{B},T}$, following similar arguments in the proof of Proposition S.1, we can show that, with probability approaching one,

$$\mathcal{S}_{\mathscr{B}}(\hat{\tau}_{\hat{B}_a,1}, \ldots, \hat{\tau}_{\hat{B}_a,\hat{B}_a}) - \mathcal{S}_{\mathscr{B}}(\tau_1^0, \ldots, \tau_{B_a^*}^0) \geq \frac{\delta_{\mathscr{B},T}^2 \underline{n}^2}{N} \Delta_{\mathscr{B}} + o_p(\eta_{p,N}).$$

Hence,

$$\begin{aligned}\Pr(\mathcal{S}_{\hat{B}_a}^{\text{Pen}} > \mathcal{S}_{B_a^*}^{\text{Pen}}) \geq \Pr(\mathcal{S}_{\hat{B}_a}^{\text{Pen}} > \mathcal{S}_{B_a^*}^{0\text{Pen}}) &\geq \Pr\left\{\frac{\delta_{\mathscr{B},T}^2 \underline{n}^2}{N} \Delta_{\mathscr{B}} + o_p(\eta_{p,N}) > 0\right\} \\ &\to 1,\end{aligned}$$

provided that $\delta_{\mathscr{B},T}^2 \underline{n}^2 N^{-1} \Delta_{\mathscr{B}} / \eta_{p,N} \to \infty$. This contradicts the fact that $\Pr(\hat{B}_a = B_a^*) \to 1$ as we have demonstrated before. $\square$

Lemmas S.5–S.9 and their proofs.

LEMMA S.5. *Assume the conditions in Theorem 5 hold, then for any $0 \leq i \leq L^*$,*

$$\sup_{\tau_i^* < t_1 < t_2 \leq \tau_{i+1}^*} \left| N_{t_1,t_2}^{-1} \sum_{1 \leq h_1 \neq h_2 \leq N_{t_1,t_2}} \widetilde{\mathbf{Y}}_{ih_1,\mathscr{B}}^\top \widetilde{\mathbf{Y}}_{ih_2,\mathscr{B}} \right| = o_p(\eta_{p,N}).$$

PROOF. Let $\mathbb{U}_{\mathcal{N}_i} = \sum_{h_1 < h_2} \widetilde{\mathbf{Y}}_{ih_1,\mathscr{B}}^\top \widetilde{\mathbf{Y}}_{ih_2,\mathscr{B}}$. It can be verified that the parameters entering in the concentration inequality given by Lemma S.3 are $A = 2$, $B^2 = 2(\mathcal{N}_i - 1)\max_{j \in \mathscr{B}} q_{ij}$, $C^2 = \mathcal{N}_i(\mathcal{N}_i - 1)/2 \cdot \text{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}^2)$ and $D = \mathcal{N}_i/2 \cdot \sqrt{\text{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}^2)}$, respectively, where $\mathbf{\Sigma}_{i,\mathscr{B}} = \text{diag}(\mathbf{q}_{i,\mathscr{B}}) - \mathbf{q}_{i,\mathscr{B}} \mathbf{q}_{i,\mathscr{B}}^\top$.

Hence, we have

$$\Pr\left(\sup_{\tau_i^*<t_1<t_2\le\tau_{i+1}^*}\left|\mathcal{N}_i^{-1}\sum_{1\le h_1\ne h_2\le\mathcal{N}_i}\widetilde{\mathbf{Y}}_{ih_1,\mathscr{B}}^\top\widetilde{\mathbf{Y}}_{ih_2,\mathscr{B}}\right|>\eta_{p,N}\right)$$
$$\le\sum_{t_1<t_2}2\Pr(\mathbb{U}_{\mathcal{N}_i}>\mathcal{N}_i\eta_{p,N}/2)$$
$$\le\sum_{t_1<t_2}2\exp\left[1-C_0\min\left\{\left(\sqrt{\frac{\mathcal{N}_i}{\mathcal{N}_i-1}}\frac{\eta_{p,N}}{\sqrt{\mathrm{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}^2)}}\right)^2,\frac{\eta_{p,N}}{\sqrt{\mathrm{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}^2)}},\right.\right.$$
$$\left.\left.\left(\frac{\mathcal{N}_i\eta_{p,N}}{\sqrt{(\mathcal{N}_i-1)\max_{j\in\mathscr{B}}q_{ij}}}\right)^{2/3},(\mathcal{N}_i\eta_{p,N})^{1/2}\right\}\right]\to 0,$$

provided that $\{\mathrm{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}^2)\}^{-1/2}\eta_{p,N}(\log T)^{-1}\to\infty$ and $\underline{n}\eta_{p,N}(\log T)^{-2}\to\infty$, as $(p,N,T)\to\infty$. Note that $\underline{n}(\max_{j\in\mathscr{B}}q_{ij})^{-1}\eta_{p,N}^2(\log T)^{-3}\to\infty$ results from the former two since $(\mathbf{q}_{i,\mathscr{B}}^\top\mathbf{q}_{i,\mathscr{B}})^{1/2}/\max_{j\in\mathscr{B}}q_{ij}>1$ and $\mathrm{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}^2)=\mathbf{q}_{i,\mathscr{B}}^\top\mathbf{q}_{i,\mathscr{B}}\{1+o(1)\}$ by Assumption (B1). □

LEMMA S.6. *Assume the conditions in Theorem 5 hold, then for any* $0\le i<k\le L^*$,

$$\sup_{\substack{\tau_i^*<t_1<t_2\le\tau_{i+1}^*\\ \tau_k^*<t_3<t_4\le\tau_{k+1}^*}}\left|N_{t_1,t_2}^{-1/2}N_{t_3,t_4}^{-1/2}\sum_{h_1=1}^{N_{t_1,t_2}}\sum_{h_2=1}^{N_{t_3,t_4}}\widetilde{\mathbf{Y}}_{ih_1,\mathscr{B}}^\top\widetilde{\mathbf{Y}}_{kh_2,\mathscr{B}}\right|=o_p(\eta_{p,N}).$$

PROOF. Let

$$\tilde{\mathcal{Y}}_h=\begin{cases}\widetilde{\mathbf{Y}}_{ih,\mathscr{B}}, & \text{for } h=1,\ldots,\mathcal{N}_i,\\ \widetilde{\mathbf{Y}}_{k,h-\mathcal{N}_i,\mathscr{B}}, & \text{for } h=\mathcal{N}_i+1,\ldots,\mathcal{N}_{ik},\end{cases}$$

and

$$a_{h_1,h_2}=\begin{cases}1, & \text{if } h_1=1,\ldots,\mathcal{N}_i;\ h_2=\mathcal{N}_i+1,\ldots,\mathcal{N}_{ik},\\ 0, & \text{otherwise.}\end{cases}$$

We have

$$\sum_{h_1=1}^{\mathcal{N}_i}\sum_{h_2=1}^{\mathcal{N}_k}\widetilde{\mathbf{Y}}_{ih_1}^\top\widetilde{\mathbf{Y}}_{kh_2}=\sum_{1\le h_1<h_2\le\mathcal{N}_{ik}}a_{h_1,h_2}\tilde{\mathcal{Y}}_{h_1}^\top\tilde{\mathcal{Y}}_{h_2}.$$

Based on similar argument in the proof of Lemma S.5, this lemma follows. Note that the parameters entering in the concentration inequality (Lemma S.3) are $A=2$, $B^2=2\max\{\mathcal{N}_i,\mathcal{N}_k\}\max_{i;j\in\mathscr{B}}q_{ij}$, $C^2=\mathcal{N}_i\mathcal{N}_k\mathrm{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}\mathbf{\Sigma}_{k,\mathscr{B}})$ and $D=\sqrt{\mathcal{N}_i\mathcal{N}_k}\mathrm{tr}(\mathbf{\Sigma}_{i,\mathscr{B}}\mathbf{\Sigma}_{k,\mathscr{B}})$, respectively. □

LEMMA S.7. *Assume the conditions in Theorem 5 hold, then for any* $0 \leq i \leq L^*$,

$$\sup_{\tau_i^* < t_1 < t_2 \leq \tau_{i+1}^*} \left| N_{t_1,t_2}^{-1/2} \sum_{h=1}^{N_{t_1,t_2}} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{q}_{i,\mathscr{B}} \right| = o_p(\eta_{p,N}).$$

PROOF. By Lemma S.2, we have

$$\begin{aligned} & \Pr\Big(\sup_{\tau_i^* < t_1 < t_2 \leq \tau_{i+1}^*} \Big| \mathcal{N}_i^{-1/2} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{q}_{i,\mathscr{B}} \Big| > \eta_{p,N} \Big) \\ \leq \;& 2 \sum_{t_1 < t_2} \exp\left(- \frac{\eta_{p,N}^2}{2\mathbf{q}_{i,\mathscr{B}}^\top \mathbf{\Sigma}_{i,\mathscr{B}} \mathbf{q}_{i,\mathscr{B}} + 4/3 \max_{j \in \mathscr{B}} q_{ij} \mathcal{N}_i^{-1/2} \eta_{p,N}} \right) \to 0, \end{aligned}$$

by noting that $\sup |\widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{q}_{i,\mathscr{B}}| = 2 \max_{j \in \mathscr{B}} q_{ij}$ and

$$\begin{aligned} \frac{\eta_{p,N}^2}{\mathbf{q}_{i,\mathscr{B}}^\top \mathbf{\Sigma}_{i,\mathscr{B}} \mathbf{q}_{i,\mathscr{B}}} \frac{1}{\log T} \;&\geq\; \frac{1}{2} \left(\frac{\eta_{p,N}}{\sqrt{\max_i \mathbf{q}_{i,\mathscr{B}}^\top \mathbf{q}_{i,\mathscr{B}}}} \frac{1}{\log T} \right)^2 \frac{\log T}{\max_{i;j \in \mathscr{B}} q_{ij}} \to \infty, \\ \frac{\mathcal{N}_i^{1/2} \eta_{p,N}}{\max_{j \in \mathscr{B}} q_{ij}} \frac{1}{\log T} \;&\geq\; \frac{\underline{n}^{1/2} \eta_{p,N}}{(\max_{i;j \in \mathscr{B}} q_{ij})^{1/2} (\log T)^{3/2}} \frac{(\log T)^{1/2}}{(\max_{i;j \in \mathscr{B}} q_{ij})^{1/2}} \to \infty, \end{aligned}$$

from which this lemma follows. □

LEMMA S.8. *Assume the conditions in Theorem 5 hold, then for any* $0 \leq i \leq L^*$,

$$\sup_{\tau_i^* < t_1 < t_2 \leq \tau_{i+1}^*} \left| N_{t_1,t_2}^{-1} \sum_{h=1}^{N_{t_1,t_2}} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{1}_{\mathscr{B}} \right| = o_p(\eta_{p,N}).$$

PROOF. When $\mathscr{B} = \{1, \ldots, p\}$, this lemma holds obviously since $\widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{1}_{\mathscr{B}} = 0$. Hence we assume that $Q_{i,\mathscr{B}} < 1$. Again by Lemma S.2, we have

$$\begin{aligned} & \Pr\Big(\sup_{\tau_i^* < t_1 < t_2 \leq \tau_{i+1}^*} \Big| \mathcal{N}_i^{-1} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{B}}^\top \mathbf{1}_{\mathscr{B}} \Big| > \eta_{p,N} \Big) \\ \leq \;& 2 \sum_{t_1 < t_2} \exp\Big(- \frac{\underline{n} \eta_{p,N}^2}{2 Q_{i,\mathscr{B}} Q_{i,\mathscr{A}} + 2/3 \eta_{p,N}} \Big) \to 0, \end{aligned}$$

since $\underline{n} \eta_{p,N}^2 (\log T)^{-1} \to \infty$ and $\underline{n} \eta_{p,N} (\log T)^{-1} \to \infty$. This lemma follows. □

LEMMA S.9. *Assume the conditions in Theorem 5 hold, then*

$$\sup_{\substack{i,k \\ c_{ik}\|\boldsymbol{\delta}_{ik,\mathscr{B}}\|^2/\eta_{p,N}\to\infty}} \frac{\sqrt{c_{ik}}S_{i,\boldsymbol{\delta}_{ik,\mathscr{B}}}}{c_{ik}\|\boldsymbol{\delta}_{ik,\mathscr{B}}\|^2} = o_p(1),$$

and

$$\sup_{\substack{i,k \\ c_{ik}\|\boldsymbol{\delta}_{ik,\mathscr{B}}\|^2/\eta_{p,N}=O(1)}} \sqrt{c_{ik}}S_{i,\boldsymbol{\delta}_{ik,\delta}} = o_p(\eta_{p,N}).$$

PROOF. We first assume that $c_{ik}\|\boldsymbol{\delta}_{ik,\delta}\|^2/\eta_{p,N}\to\infty$. For any $1\le l<l'\le L^*$, denote $\boldsymbol{\delta}\equiv\boldsymbol{\delta}_{ll',\mathscr{B}}=\mathbf{q}_{l,\mathscr{B}}-\mathbf{q}_{l',\mathscr{B}}$. It suffices to show that

$$\Pr\Big(\sup_{\substack{1\le t_1<t_2\le\mathcal{N}_1\\1\le t_3<t_4\le\mathcal{N}_2}}\sum_{h=1}^{\mathcal{N}_1}\widetilde{\mathbf{Y}}_{lh,\mathscr{B}}^\top\boldsymbol{\delta}>\mathcal{N}_1^{1/2}\sqrt{\mathcal{C}}\|\boldsymbol{\delta}\|^2\Big)\to 0,$$

where $\mathcal{N}_1=N_{t_1,t_2}$, $\mathcal{N}_2=N_{t_3,t_4}$ and $\mathcal{C}=\mathcal{N}_1\mathcal{N}_2/(\mathcal{N}_1+\mathcal{N}_2)$. By the Bernstein inequality (Lemma S.2), we have

$$\begin{aligned}&\Pr\Big(\sup_{\substack{1\le t_1<t_2\le\mathcal{N}_1\\1\le t_3<t_4\le\mathcal{N}_2}}\sum_{h=1}^{\mathcal{N}_1}\widetilde{\mathbf{Y}}_{lh,\mathscr{B}}^\top\boldsymbol{\delta}>\mathcal{N}_1^{1/2}\sqrt{\mathcal{C}}\|\boldsymbol{\delta}\|^2\Big)\\ \le\;&\sum_{t_1<t_2}\sum_{t_3<t_4}\exp\Big(-\frac{\mathcal{C}\|\boldsymbol{\delta}\|^4}{2\boldsymbol{\delta}^\top\boldsymbol{\Sigma}_{l,\mathscr{B}}\boldsymbol{\delta}+8/3\max_{l;j\in\mathscr{B}}q_{lj}\mathcal{N}_1^{-1/2}\sqrt{\mathcal{C}}\|\boldsymbol{\delta}\|^2}\Big)\to 0,\end{aligned}$$

by the fact that $\sup|\widetilde{\mathbf{Y}}_{lh,\mathscr{B}}^\top\boldsymbol{\delta}|=4\max_{l;j\in\mathscr{B}}q_{lj}$, $\max_{l;j\in\mathscr{B}}q_{lj}\le\sqrt{\max_l\mathbf{q}_{l,\mathscr{B}}^\top\mathbf{q}_{l,\mathscr{B}}}$ and

$$\begin{aligned}\frac{\mathcal{C}\|\boldsymbol{\delta}\|^4}{\boldsymbol{\delta}^\top\boldsymbol{\Sigma}_{l,\mathscr{B}}\boldsymbol{\delta}\log T}&\ge\frac{1}{2}\frac{\mathcal{C}\|\boldsymbol{\delta}\|^2}{\eta_{p,N}}\frac{\eta_{p,N}}{\sqrt{\max_l\mathbf{q}_{l,\mathscr{B}}^\top\mathbf{q}_{l,\mathscr{B}}}\log T}\to\infty,\\ \frac{\mathcal{N}_1^{1/2}\sqrt{\mathcal{C}}\|\boldsymbol{\delta}\|^2}{\max_{l;j\in\mathscr{B}}q_{lj}\log T}&\ge\frac{\mathcal{C}\|\boldsymbol{\delta}\|^2}{\eta_{p,N}}\frac{\eta_{p,N}}{\sqrt{\max_l\mathbf{q}_{l,\mathscr{B}}^\top\mathbf{q}_{l,\mathscr{B}}}\log T}\to\infty.\end{aligned}$$

Hence the first part of this lemma follows.

Suppose that $c_{ik}\|\boldsymbol{\delta}_{ik}\|^2/\eta_{p,N}=O(1)$. Similarly, we have

$$\begin{aligned}&\Pr\Big(\sup_{\substack{1\le t_1<t_2\le\mathcal{N}_1\\1\le t_3<t_4\le\mathcal{N}_2}}\sum_{h=1}^{\mathcal{N}_1}\widetilde{\mathbf{Y}}_{lh,\mathscr{B}}^\top\boldsymbol{\delta}>\eta_{p,N}\mathcal{N}_1^{1/2}\mathcal{C}^{-1/2}\Big)\\ \le\;&\sum_{t_1<t_2}\sum_{t_3<t_4}\exp\Big(-\frac{\eta_{p,N}^2}{2\mathcal{C}\boldsymbol{\delta}^\top\boldsymbol{\Sigma}_{l,\mathscr{B}}\boldsymbol{\delta}+8/3\max_{l;j\in\mathscr{B}}q_{lj}\mathcal{N}_1^{-1/2}\sqrt{\mathcal{C}}\eta_{p,N}}\Big)\to 0,\end{aligned}$$

since

$$\frac{\eta_{p,N}^2}{\mathcal{C}\boldsymbol{\delta}^\top\boldsymbol{\Sigma}_{l,\mathscr{B}}\boldsymbol{\delta}\log T} \geq \frac{1}{2}\frac{\eta_{p,N}}{\mathcal{C}\|\boldsymbol{\delta}\|^2}\frac{\eta_{p,N}}{\sqrt{\max_l \mathbf{q}_{l,\mathscr{B}}^\top\mathbf{q}_{l,\mathscr{B}}}\log T} \to \infty,$$

$$\frac{\eta_{p,N}\mathcal{N}_1^{1/2}}{\max_{l;j\in\mathscr{B}} q_{lj}\sqrt{\mathcal{C}}\log T} \geq \frac{\eta_{p,N}}{\sqrt{\max_l \mathbf{q}_{l,\mathscr{B}}^\top\mathbf{q}_{l,\mathscr{B}}}\log T} \to \infty,$$

which completes the proof. □

Proof of Theorem 4. The proof of Theorem 4 is similar to that of Theorem 5, based on a parallel version of Lemma S.4 summarized below. Thus we omit the proof.

LEMMA S.10. *Assume the conditions in Theorem 4 hold, then for any* $1 \leq i, k \leq L^*$,

$$\sup_{\substack{\tau_i^*<t_1<t_2\leq\tau_{i+1}^*\\ \tau_k^*<t_3<t_4\leq\tau_{k+1}^*}} \sum_{j\in\mathscr{A}} \frac{N_{t_1,t_2}N_{t_3,t_4}}{N_{t_1,t_2}+N_{t_3,t_4}}\{\bar{X}_j(t_1,t_2)-\bar{X}_j(t_3,t_4)\}^2/\hat{q}_j$$
$$=\frac{N_{t_1,t_2}N_{t_3,t_4}}{N_{t_1,t_2}+N_{t_3,t_4}}\sum_{j\in\mathscr{A}}\frac{(q_{ij}-q_{kj})^2}{q_{*j}}\{1+o_p(1)\}+o_p(\xi_{p,N}),$$

where $q_{*j} = \mathrm{E}(\hat{q}_j) = \sum_{l=1}^{L^*} N_{\tau_{l-1}^*,\tau_l^*}/N q_{lj}$.

PROOF. Use the notation in the proof of Lemma S.4, i.e., $\mathcal{N}_i = N_{t_1,t_2}, \mathcal{N}_k = N_{t_3,t_4}, \mathcal{N}_{ik} = \mathcal{N}_i + \mathcal{N}_k$ and $c_{ik} = \mathcal{N}_i\mathcal{N}_k/\mathcal{N}$. We introduce $\mathbf{Y}_{i1},\ldots,\mathbf{Y}_{i\mathcal{N}_i} \overset{\text{iid}}{\sim} \text{Multi}(1,\mathbf{q}_i)$, where $\mathbf{Y}_{ih} = (Y_{ih,1},\ldots,Y_{ih,p})^\top$ for $h = 1,\ldots,\mathcal{N}_i$. Let $\mathbf{Y}_{ih,\mathscr{A}}$ and $\mathbf{q}_{i,\mathscr{A}}$ be the collections of $Y_{ih,j}$ and q_{ij} with $j \in \mathscr{A}$, respectively. Denote $\widetilde{\mathbf{Y}}_{ih,\mathscr{A}} = \mathbf{Y}_{ih,\mathscr{A}} - \mathbf{q}_{i,\mathscr{A}}$. Following similar arguments as those in the proof of Theorem 3, we can show that

$$\sum_{j\in\mathscr{A}} \frac{N_{t_1,t_2}N_{t_3,t_4}}{N_{t_1,t_2}+N_{t_3,t_4}}\{\bar{X}_j(t_1,t_2)-\bar{X}_j(t_3,t_4)\}^2/q_{*j}$$
$$=c_{ik}\sum_{j\in\mathscr{A}}(\sigma_{ij}^2/\mathcal{N}_i+\sigma_{kj}^2/\mathcal{N}_k)/q_{*j}+\frac{\mathcal{N}_k}{\mathcal{N}_{ik}}d_{i,\mathscr{A}}+\frac{\mathcal{N}_i}{\mathcal{N}_{ik}}d_{k,\mathscr{A}}$$
$$+c_{ik}\boldsymbol{\delta}_{ik,\mathscr{A}}^\top\mathbf{D}_{*,\mathscr{A}}^{-1}\boldsymbol{\delta}_{ik,\mathscr{A}}+2\sqrt{\frac{\mathcal{N}_k}{\mathcal{N}_{ik}}}\sqrt{c_{ik}}S_{i,\boldsymbol{\delta}_{ik,\mathscr{A}}}-2\sqrt{\frac{\mathcal{N}_i}{\mathcal{N}_{ik}}}\sqrt{c_{ik}}S_{k,\boldsymbol{\delta}_{ik,\mathscr{A}}}$$
$$+\frac{\mathcal{N}_k}{\mathcal{N}_{ik}}U_{i,\mathscr{A}}+\frac{\mathcal{N}_i}{\mathcal{N}_{ik}}U_{k,\mathscr{A}}-\frac{2\sqrt{\mathcal{N}_i\mathcal{N}_k}}{\mathcal{N}_{ik}}V_{ik,\mathscr{A}}-2\frac{\mathcal{N}_k}{\mathcal{N}_{ik}}\frac{W_{i,\mathscr{A}}}{\sqrt{\mathcal{N}_i}}-2\frac{\mathcal{N}_i}{\mathcal{N}_{ik}}\frac{W_{k,\mathscr{A}}}{\sqrt{\mathcal{N}_k}},$$

where

$$
\begin{aligned}
d_{i,\mathscr{A}} =& \mathcal{N}_i^{-1} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{A}}^\top \mathbf{D}_{*,\mathscr{A}}^{-1} \mathbf{1}_{\mathscr{A}} \\
S_{i,\boldsymbol{\delta}_{ik,\mathscr{A}}} =& \mathcal{N}_i^{-1/2} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{A}}^\top \mathbf{D}_{*,\mathscr{A}}^{-1} \boldsymbol{\delta}_{ik,\mathscr{A}} \\
U_{i,\mathscr{A}} =& \mathcal{N}_i^{-1} \sum_{1 \le h_1 \ne h_2 \le \mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih_1,\mathscr{A}}^\top \mathbf{D}_{*,\mathscr{A}}^{-1} \widetilde{\mathbf{Y}}_{ih_2,\mathscr{A}} \\
V_{ik,\mathscr{A}} =& (\mathcal{N}_i \mathcal{N}_k)^{-1/2} \sum_{h_1=1}^{\mathcal{N}_i} \sum_{h_2=1}^{\mathcal{N}_k} \widetilde{\mathbf{Y}}_{ih_1,\mathscr{A}}^\top \mathbf{D}_{*,\mathscr{A}}^{-1} \widetilde{\mathbf{Y}}_{kh_2,\mathscr{A}} \\
W_{i,\mathscr{A}} =& \mathcal{N}_i^{-1/2} \sum_{h=1}^{\mathcal{N}_i} \widetilde{\mathbf{Y}}_{ih,\mathscr{A}}^\top \mathbf{D}_{*,\mathscr{A}}^{-1} \mathbf{q}_{i,\mathscr{A}}
\end{aligned}
$$

and $\boldsymbol{\delta}_{ik,\mathscr{A}} = \mathbf{q}_{i,\mathscr{A}} - \mathbf{q}_{i,\mathscr{A}}$. Noting the fact that $q_{*j} > C^* \varepsilon$ and the conditions that $\xi_{p,N}(\log T)^{-1} \to \infty$ and $\underline{n}\xi_{p,N}(\log T)^{-2} \to \infty$, the conclusion follows. □

Additional simulation results. Following Section 4.2 in the manuscript, we present additional numerical comparisons of the empirical test sizes between our $Q_{p,\hat{\mathscr{A}}}$ test, and $W_{\tilde{p}}^{(\mathrm{SW})}$ proposed by **?** and $W_{\tilde{p}}^{(\mathrm{HS})}$ by **?**. Tables S1–S6 correspond to the significance levels of 1%, 5%, and 10% with $T = 100$ and 10. Similar conclusions can be drawn as follows. The empirical sizes of our test are reasonably close to the nominal levels, while those of $W_{\tilde{p}}^{(\mathrm{HS})}$ experience the most serious distortion, and $W_{\tilde{p}}^{(\mathrm{SW})}$ performs poorly for small n.

For power comparison, we examine more alternatives which allow signals to occur on both $\mathscr{A}$ and $\mathscr{B}$. Recall that $\mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$. We consider two alternatives:

(S–i) $\mathbf{q}_1 = \{\tilde{\omega}/d\mathbf{1}_d^\top, (1-\tilde{\omega})/(p-d)\mathbf{1}_{p-d}^\top\}^\top$,

(S–ii) $\mathbf{q}_1 = \{\omega/\tilde{d}\mathbf{1}_{\tilde{d}}^\top, (1-\omega)/(p-\tilde{d})\mathbf{1}_{p-\tilde{d}}^\top\}^\top$.

Figure S1 depicts the power curves of the above three tests against $\tilde{\omega}$ and $\tilde{d}$, respectively. In most cases, $Q_{p,\hat{\mathscr{A}}}$ performs better than the other two competitors in terms of attaining high power while maintaining the test size. Also, Figure S2 presents the size-corrected power comparison regarding the performance of $Q_{p,\hat{\mathscr{A}}}$, M_p, $H_p^{(\text{sum})}$ and $H_p^{(\text{max})}$. Overall, our procedure performs more satisfactorily under both alternatives.

TABLE S1

Comparison of empirical sizes (%) at a 1% significance level for the change-point test under $H_0 : \mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$ *and different* (p, N)*-settings when* $T = 100$*.*

		$\omega = 0.3$			$\omega = 0.5$			$\omega = 0.7$		
p	n	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$
500	1	1.44	0.00	0.00	1.68	0.14	0.00	1.46	0.78	0.02
	10	1.34	0.34	0.00	1.34	1.18	0.00	1.54	2.42	0.00
	20	1.14	0.50	0.20	1.32	0.98	0.06	1.42	2.76	0.00
	50	1.38	0.58	1.76	1.16	0.66	0.96	1.12	1.32	0.42
1000	1	1.56	0.00	0.00	2.10	0.08	0.00	1.36	0.82	0.00
	10	1.32	0.24	0.00	1.30	0.98	0.00	1.60	2.62	0.00
	20	0.94	0.48	0.00	1.26	1.32	0.00	1.38	2.92	0.00
	50	1.42	0.58	0.60	1.30	1.20	0.10	1.04	2.32	0.00

TABLE S2

Comparison of empirical sizes (%) at a 5% significance level for the change-point test under $H_0 : \mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$ *and different* (p, N)*-settings when* $T = 100$*.*

		$\omega = 0.3$			$\omega = 0.5$			$\omega = 0.7$		
p	n	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$
500	1	5.16	0.00	0.00	6.00	0.84	0.06	4.14	2.76	0.52
	10	5.44	1.84	0.66	5.50	3.24	0.22	5.88	5.82	0.04
	20	5.54	2.14	3.48	5.38	3.50	1.74	6.06	6.14	0.40
	50	6.20	2.60	9.14	5.12	3.04	6.56	5.36	4.56	3.58
1000	1	5.58	0.00	0.00	4.72	0.66	0.10	2.54	3.10	0.50
	10	5.64	1.26	0.04	5.56	3.26	0.00	5.54	5.94	0.00
	20	5.18	1.82	0.38	5.24	4.06	0.18	5.90	6.30	0.00
	50	5.98	2.22	4.46	5.72	3.32	2.60	5.54	6.44	0.46

Table S3
Comparison of empirical sizes (%) at a 10% significance level for the change-point test under $H_0 : \mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$ *and different* (p, N)*-settings when* $T = 100$.

		$\omega = 0.3$			$\omega = 0.5$			$\omega = 0.7$		
p	n	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$
500	1	8.68	0.00	0.04	9.68	1.88	0.60	6.62	5.40	2.20
	10	11.02	2.86	3.56	11.28	5.84	1.42	11.16	9.32	0.38
	20	10.68	3.98	9.76	11.18	5.90	6.12	10.92	9.24	2.22
	50	10.48	4.86	17.14	10.94	5.82	15.86	10.70	7.62	10.56
1000	1	8.36	0.02	0.00	7.54	1.86	0.56	3.94	5.48	1.76
	10	10.70	2.08	0.20	11.28	5.24	0.08	10.60	9.02	0.00
	20	10.80	3.18	2.44	9.98	5.36	1.02	11.10	10.44	0.20
	50	10.54	4.08	11.20	12.14	6.56	7.24	11.32	8.72	3.16

Table S4
Comparison of empirical sizes (%) at a 1% significance level for the change-point test under $H_0 : \mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$ *and different* (p, N)*-settings when* $T = 10$.

		$\omega = 0.3$			$\omega = 0.5$			$\omega = 0.7$		
p	n	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$
500	10	1.46	0.00	0.00	1.98	0.62	0.00	1.40	1.68	0.00
	20	1.60	0.26	0.00	1.82	0.90	0.00	1.80	3.34	0.00
	50	1.20	0.52	0.00	1.56	1.78	0.00	1.72	4.12	0.00
	100	1.30	0.94	0.00	1.26	2.02	0.00	1.64	4.06	0.00
1000	10	2.22	0.00	0.00	1.76	0.46	0.00	0.90	1.76	0.00
	20	1.64	0.06	0.00	1.74	1.26	0.00	1.90	2.66	0.00
	50	1.32	0.40	0.00	1.72	1.90	0.00	1.46	3.96	0.00
	100	1.44	0.66	0.00	1.46	2.36	0.00	1.68	4.12	0.00

TABLE S5

Comparison of empirical sizes (%) at a 5% significance level for the change-point test under $H_0 : \mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$ *and different* (p, N)*-settings when* $T = 10$*.*

		$\omega = 0.3$			$\omega = 0.5$			$\omega = 0.7$		
p	n	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$
500	10	4.70	0.14	0.00	6.00	2.74	0.04	4.42	6.32	0.10
	20	4.76	0.98	0.00	5.98	4.42	0.00	7.14	8.20	0.04
	50	5.44	2.36	0.02	5.70	6.34	0.00	5.70	9.90	0.00
	100	5.68	3.92	0.24	5.80	6.26	0.00	6.12	9.92	0.00
1000	10	4.88	0.10	0.00	4.40	2.74	0.00	2.92	6.26	0.18
	20	5.24	0.70	0.00	5.66	4.26	0.00	5.32	7.72	0.02
	50	5.66	1.72	0.00	6.34	5.30	0.00	5.74	9.12	0.00
	100	6.06	2.70	0.06	5.10	5.88	0.00	5.68	9.72	0.00

TABLE S6

Comparison of empirical sizes (%) at a 10% significance level for the change-point test under $H_0 : \mathbf{q}_0 = \{\omega/d\mathbf{1}_d^\top, (1-\omega)/(p-d)\mathbf{1}_{p-d}^\top\}^\top$ *and different* (p, N)*-settings when* $T = 10$*.*

		$\omega = 0.3$			$\omega = 0.5$			$\omega = 0.7$		
p	n	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$	$Q_{p,\hat{\mathscr{A}}}$	$W_{\tilde{p}}^{(\mathrm{SW})}$	$W_{\tilde{p}}^{(\mathrm{HS})}$
500	10	8.12	0.78	0.06	8.82	5.82	0.26	5.74	10.50	0.68
	20	10.12	2.60	0.02	11.26	7.02	0.06	9.86	12.40	0.30
	50	9.88	4.34	0.14	9.58	9.38	0.06	10.80	14.28	0.02
	100	10.16	7.00	1.24	10.62	9.72	0.44	11.00	15.00	0.16
1000	10	7.18	0.70	0.00	7.08	5.62	0.10	4.28	10.50	0.70
	20	9.94	1.84	0.04	10.36	7.02	0.06	8.10	12.58	0.38
	50	11.58	3.86	0.02	10.68	9.02	0.00	10.62	14.00	0.00
	100	10.26	5.24	0.24	10.60	9.70	0.02	10.84	14.20	0.02

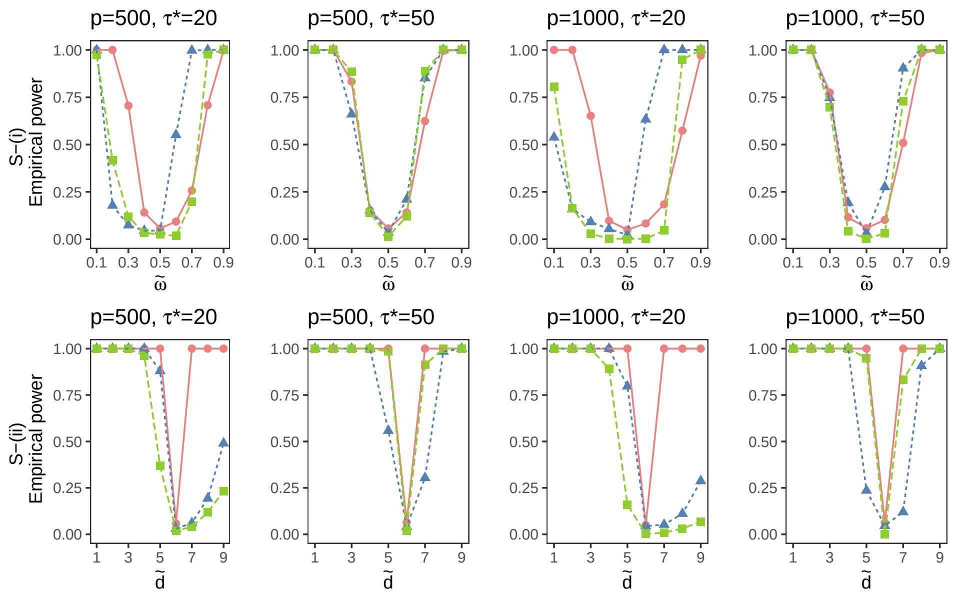

Fig S1. Comparison of empirical power for the proposed $Q_{p,\hat{\mathscr{A}}}$ test, $W_{\tilde{p}}^{(\mathrm{SW})}$ by **?** *and $W_{\tilde{p}}^{(\mathrm{HS})}$ by* **?** *under the alternative hypotheses (S–i) and (S–ii) with $\omega = 0.5$ and $d = 6$.*

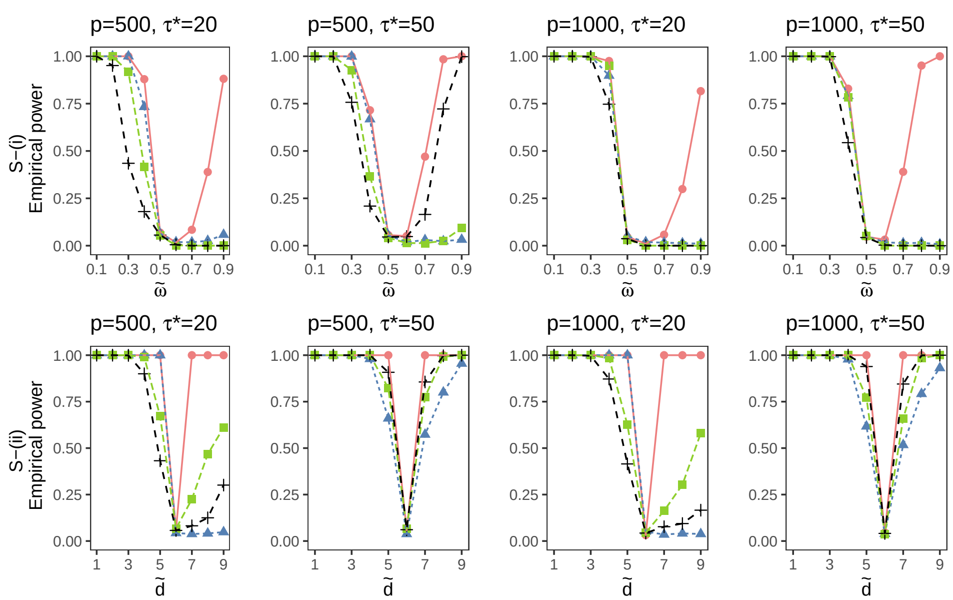

Fig S2. Comparison of size-corrected empirical power under the alternative hypotheses (S–i) and (S–ii).

References.

GINÉ, E. and NICKL, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge University Press.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic Press.

HORVÁTH, L. and SERBINOWSKA, M. (1995). Testing for Changes in Multinomial Observations: The Lindisfarne Scribes Problem. *Scandinavian Journal of Statistics* **22** 371–384.

SRIVASTAVA, M. S. and WORSLEY, K. J. (1986). Likelihood Ratio Tests for a Change in the Multivariate Normal Mean. *Journal of the American Statistical Association* **81** 199–204.

GUANGHUI WANG, CHANGLIANG ZOU
INSTITUTE OF STATISTICS AND LPMC
NANKAI UNIVERSITY
CHINA
E-MAIL: ghwang.nk@gmail.com
nk.chlzou@gmail.com

GUOSHENG YIN
DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE
THE UNIVERSITY OF HONG KONG
HONG KONG
E-MAIL: gyin@hku.hk