TESTING CONSTANCY OF CONDITIONAL VARIANCE IN HIGH DIMENSION

Lu Deng, Changliang Zou, Zhaojun Wang and Xin Chen

Nankai University and South University of Science and Technology of China

Supplementary Material

The supplementary material includes all the proofs of Theorems 1–2 and Propositions 1–4, as well as a necessary lemma. Some additional simulation results are also presented.

We first give a useful lemma as follows.

Lemma 1. *Under Condition (C4), given* $\mathbf{Y}$*, let* $\mathbf{A}(\mathbf{Y}) = \Gamma^{\mathrm{T}}(\mathbf{Y})\Gamma(\mathbf{Y})$*, we have*

(i) for any positive integer k*,*

$$E\{(\mathbf{Z}^{\mathrm{T}}(\mathbf{Y})\mathbf{A}^k(\mathbf{Y})\mathbf{Z}(\mathbf{Y}))^2 \mid \mathbf{Y}\} = \mathrm{tr}^2(\boldsymbol{\Sigma}_0^k)+2\mathrm{tr}(\boldsymbol{\Sigma}_0^{2k})+\Delta(\mathbf{Y})\mathrm{tr}(\mathbf{A}^k(\mathbf{Y})\circ\mathbf{A}^k(\mathbf{Y})),$$

here we define $\mathbf{F} \circ \mathbf{G} = (f_{kl}g_{kl})$*, where* $\mathbf{F} = (f_{kl})$ *and* $\mathbf{G} = (g_{kl})$*;*

(ii) for independent variables $\mathbf{Z}_1(\mathbf{Y})$ *and* $\mathbf{Z}_2(\mathbf{Y})$*,*

$$E\{(\mathbf{Z}_1^{\mathrm{T}}(\mathbf{Y})\mathbf{A}(\mathbf{Y})\mathbf{Z}_2(\mathbf{Y}))^4 \mid \mathbf{Y}\} = 3\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + 6\mathrm{tr}(\boldsymbol{\Sigma}_0^4) + 6\Delta(\mathbf{Y})\mathrm{tr}(\mathbf{A}^2(\mathbf{Y}) \circ \mathbf{A}^2(\mathbf{Y}))$$

$$+\Delta^2(\mathbf{Y})\sum_{k,l=1}^{m}(\mathbf{A}_{kl}(\mathbf{Y}))^4,$$

(iii) $\sum_{k,l=1}^{m}(\mathbf{A}_{kl}(\mathbf{Y}))^4 \le \mathrm{tr}(\mathbf{\Sigma}_0^4)$ *and* $\mathrm{tr}(\mathbf{A}^k(\mathbf{Y})\circ\mathbf{A}^k(\mathbf{Y})) \le \mathrm{tr}(\mathbf{\Sigma}_0^{2k})$.

(iv) for any $m\times m$ *positive definite matrix* F,

$$E\big\{(\mathbf{Z}^{\mathrm{T}}(\mathbf{Y})\mathrm{F}\mathbf{Z}(\mathbf{Y})-\mathrm{tr}(\mathrm{F}))^4\big\} \le C\mathrm{tr}^2(\mathrm{F}^2).$$

where C *is a constant which does't depend on* $\mathbf{Y}$.

This lemma is similar to Proposition A.1 in Chen, Zhang and Zhong (2010), replacing all the expectations by the conditional expectations.

Proof of Proposition 1

Proof. For notational convenience, let $A_i = \widehat{\mathrm{tr}(\mathbf{\Sigma}_i^2)}$ and $C_{ij} = \widehat{\mathrm{tr}(\mathbf{\Sigma}_i\mathbf{\Sigma}_j)}$; Denote $\boldsymbol{\phi}_i(\{s,t\})$ and $\boldsymbol{\phi}_i(\{s\})$ as the counterparts of $\boldsymbol{\theta}_i(\{s,t\})$ and $\boldsymbol{\theta}_i(\{s\})$ after replacing every $\mathbf{X}_{ik}$ by $\boldsymbol{\varepsilon}_{ik}$, respectively. Similarly, let $\boldsymbol{\omega}_i(\{s,t\})$ and $\boldsymbol{\omega}_i(\{s,t\})$ be the counterparts of $\boldsymbol{\theta}_i(\{s,t\})$ and $\boldsymbol{\theta}_i(\{s\})$ by replacing $\mathbf{X}_{ik}$ by $\boldsymbol{\mu}_{ik}$, respectively. The test statistic T_n can be rewritten as

$$\begin{aligned}
T_n =&(H-1)\sum_{i=1}^{H}A_i - 2\sum_{i<j}C_{ij}\\
=&(H-1)\sum_{i=1}^{H}\frac{1}{l(l-1)}\sum_{s\neq t}\{(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\varepsilon}_{it}-\boldsymbol{\phi}_i(\{s,t\}))\}^2\\
&-2\sum_{i<j}\frac{1}{l^2}\sum_{s,t}\{(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s\}))^{\mathrm{T}}(\boldsymbol{\varepsilon}_{jt}-\boldsymbol{\phi}_j(\{t\}))\}^2+R_1,
\end{aligned}$$

where R_1 denotes the remaining terms.

Expanding $\boldsymbol{\phi}_i(\{s,t\})$ and $\boldsymbol{\phi}_i(\{s\})$ as $(l-2)^{-1}\sum_{r\neq s,t}\boldsymbol{\varepsilon}_{ir}$ and $(l-1)^{-1}\sum_{r\neq s}\boldsymbol{\varepsilon}_{ir}$, respectively, we then have

$$
\begin{aligned}
T_n = (H-1)\sum_{i=1}^{H}\Bigg[&\left\{\frac{1}{l(l-1)}+\frac{2}{l(l-1)(l-2)}+\frac{2(l-3)}{l(l-1)(l-2)^3}\right\}\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2\\
&+\left\{\frac{-2}{l(l-1)(l-2)}+\frac{-8(l-3)}{l(l-1)(l-2)^3}+\frac{4(l-3)(l-4)}{l(l-1)(l-2)^4}\right\}\sum_{s,t,r}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\\
&+\frac{l+4}{l(l-2)^4}\sum_{s,t,r,q}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{iq}+\frac{-2l+8}{l(l-1)(l-2)^4}\sum_{s,t,r}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir}\\
&+\frac{-4}{l(l-1)(l-2)^3}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}+\frac{1}{l(l-2)^3}\sum_{s}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is})^2\\
&+\frac{l-3}{l(l-1)(l-2)^3}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\Bigg]\\
&-2\sum_{i<j}\left\{\frac{1}{(l-1)^2}\sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jt})^2-\frac{1}{(l-1)^3}\sum_{s}\sum_{t\neq r}\boldsymbol{\varepsilon}_{jt}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jr}\right.\\
&\left.-\frac{1}{(l-1)^3}\sum_{s}\sum_{t\neq r}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{js}\boldsymbol{\varepsilon}_{js}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir}+\frac{1}{(l-1)^4}\sum_{s\neq t}\sum_{r\neq q}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jr}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jq}\right\}+R_1
\end{aligned}
\tag{S.1}
$$

By Lemma 1 in the Supplementary Material, it can be shown that

$$
\begin{aligned}
E(T_n-R_1)=&\left\{\frac{2}{l-2}+\frac{2}{(l-2)^2}-\frac{2}{l-1}-\frac{1}{(l-1)^2}\right\}H(H-1)\mathrm{tr}(\boldsymbol{\Sigma}_0^2)\\
&+\frac{1}{(l-2)^2}H(H-1)\mathrm{tr}^2(\boldsymbol{\Sigma}_0)+O\left\{\frac{H(H-1)}{(l-2)^3}\mathrm{tr}(\boldsymbol{\Sigma}_0^2)\right\}.
\end{aligned}
$$

To derive the variance of T_n, consider its variants $\widetilde{T_n}$

$$
\widetilde{T_n}=(H-1)\sum_{i=1}^{H}\left\{\{l(l-1)\}^{-1}+2\{l(l-1)(l-2)\}^{-1}\right.
$$

$$+ 2(l-3)\{l(l-1)(l-2)^3\}^{-1}\} \sum_{s\neq t} (\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2 - 2\sum_{i<j}(l-1)^{-2}\sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jt})^2.$$

Tedious algebra yields that

$$\begin{aligned}
&\mathrm{var}(\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2) = l(l-1)\{(-4l+6)\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + 4(l-2)\delta_2 + 2\delta_3\},\\
&\mathrm{var}(\sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jt})^2) = l^2\{(-2l+1)\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + 2(l-1)\delta_2 + \delta_3\}, \qquad \text{(S.2)}\\
&\mathrm{cov}(\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2, \sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jt})^2) = 2l^2(l-1)\{-\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + \delta_2\},\\
&\mathrm{cov}(\sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jt})^2, \sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{j't})^2) = l^3\{-\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + \delta_2\},
\end{aligned}$$

where δ_2 and δ_3 are $E\{(\boldsymbol{\varepsilon}_{i1}^{\mathrm{T}}\boldsymbol{\varepsilon}_{i2})^2(\boldsymbol{\varepsilon}_{i1}^{\mathrm{T}}\boldsymbol{\varepsilon}_{i3})^2\}$ and $E\{(\boldsymbol{\varepsilon}_{i1}^{\mathrm{T}}\boldsymbol{\varepsilon}_{i2})^4\}$, respectively. By Lemma 1, we have $\delta_2 = \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + O(\mathrm{tr}(\boldsymbol{\Sigma}_0^4))$, $\delta_3 = 3\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + O(\mathrm{tr}(\boldsymbol{\Sigma}_0^4))$. By Condition (C3), we have

$$\mathrm{var}(\widetilde{T_n}) = 4l^{-2}H^2(H-1)\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2)(1+o(1)).$$

By the first part of Condition (C2), the last term in $E(T_n - R_1)$ is $o(\sqrt{\mathrm{var}(\widetilde{T_n})})$. By similar arguments, we also have that

$$\begin{aligned}
\mathrm{var}(\sum_{s,t,r}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}) =& 2l(l-1)(l-2)\{\delta_2 + (l-3)\mathrm{tr}(\boldsymbol{\Sigma}_0^4)\},\\
\mathrm{var}(\sum_{s,t,r,q}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{iq}) =& 8l(l-1)(l-2)(l-3)\{\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + 2\mathrm{tr}(\boldsymbol{\Sigma}_0^4)\},\\
\mathrm{var}(\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}) =& P_l^4\mathrm{tr}^4(\boldsymbol{\Sigma}_0) + 4P_l^3\{\mathrm{tr}^2(\boldsymbol{\Sigma}_0) + O(\mathrm{tr}(\boldsymbol{\Sigma}_0^4))\}\mathrm{tr}^2(\boldsymbol{\Sigma}_0)\\
&+ 2P_l^2\{\mathrm{tr}^2(\boldsymbol{\Sigma}_0) + O(\mathrm{tr}(\boldsymbol{\Sigma}_0^4))\}^2 - (P_l^2)^2\mathrm{tr}^4(\boldsymbol{\Sigma}_0)
\end{aligned}$$

$$
\begin{aligned}
&= l^3 O\{\mathrm{tr}(\boldsymbol{\Sigma}_0^2)\mathrm{tr}^2(\boldsymbol{\Sigma}_0)\},\\
\mathrm{var}(\sum_{s,t,r}^{*} \boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir}) &= l^4 O\{\mathrm{tr}(\boldsymbol{\Sigma}_0^2)\mathrm{tr}^2(\boldsymbol{\Sigma}_0)\},\\
\mathrm{var}(\sum_{s\neq t} \boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}) &= l^3 O\{\mathrm{tr}(\boldsymbol{\Sigma}_0^2)\mathrm{tr}^2(\boldsymbol{\Sigma}_0)\},\\
\mathrm{var}(\sum_{s} (\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is})^2) &= l O\{\mathrm{tr}(\boldsymbol{\Sigma}_0^2)\mathrm{tr}^2(\boldsymbol{\Sigma}_0)\},
\end{aligned}
$$

Under the condition $p = o(l^3)$, each term in $T_n - \widetilde{T_n} - R_1$ is of $o_p(\sqrt{\mathrm{var}(\widetilde{T_n})})$. Now it suffices to verity that $R_1 = o_p(\sqrt{\mathrm{var}(\widetilde{T_n})})$, which has the form

$$
\begin{aligned}
R_1 =&(H-1)\sum_i \frac{1}{l(l-1)}\sum_{s\neq t}\Big\{2(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\varepsilon}_{it}-\boldsymbol{\phi}_i(\{s,t\}))\\
&+(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))+(\boldsymbol{\varepsilon}_{it}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))\\
&+(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))\Big\}\Big\{(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))\\
&+(\boldsymbol{\varepsilon}_{it}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))+(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))\Big\}\\
&-2\sum_{i\leq j}\frac{1}{l^2}\sum_{s,t}\Big\{2(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s\}))^{\mathrm{T}}(\boldsymbol{\varepsilon}_{jt}-\boldsymbol{\phi}_j(\{t\}))\\
&+(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s\}))^{\mathrm{T}}(\boldsymbol{\mu}_{jt}-\boldsymbol{\omega}_j(\{t\}))+(\boldsymbol{\varepsilon}_{jt}-\boldsymbol{\phi}_j(\{t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s\}))\\
&+(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s\}))^{\mathrm{T}}(\boldsymbol{\mu}_{jt}-\boldsymbol{\omega}_j(\{t\}))\Big\}\Big\{(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s\}))^{\mathrm{T}}(\boldsymbol{\mu}_{jt}-\boldsymbol{\omega}_j(\{t\}))\\
&+(\boldsymbol{\varepsilon}_{jt}-\boldsymbol{\phi}_j(\{t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s\}))+(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s\}))^{\mathrm{T}}(\boldsymbol{\mu}_{jt}-\boldsymbol{\omega}_j(\{t\}))\Big\}\\
=&R_{1,1}+R_{1,2}.
\end{aligned}
$$

Expanding $\boldsymbol{\phi}_i$ and $\boldsymbol{\omega}_i$ as in (S.1), by Conditions (C1) and (C2), it can be seen that in $R_{1,1}$

$$\begin{aligned}
&(H-1)\sum_i \frac{1}{l(l-1)}\sum_{s\neq t}\big\{(\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))\big\}^2\\
&\quad\le H\sum_i \max_{s,t}\|\boldsymbol{\mu}_{is}-\boldsymbol{\mu}_{it}\|^4 \le HM^4p^2\sum_i \max_{s,t}\|\mathbf{Y}_{is}-\mathbf{Y}_{it}\|^{4\alpha}\\
&\quad\le HM^4p^2\sum_i r_{1i}^{4\alpha} = \sqrt{\mathrm{var}(\widetilde{T_n})}O_p\big(H^{-1/2}pl\sum_i r_{1i}^{4\alpha}\big) = o_p(\sqrt{\mathrm{var}(\widetilde{T_n})}).
\end{aligned}$$

Based on the specific form of R_1, the above result also holds for corresponding term in $R_{1,2}$. Now we consider the next term in $R_{1,1}$.

$$\begin{aligned}
&(H-1)\sum_i \frac{1}{l(l-1)}\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\varepsilon}_{it}-\boldsymbol{\phi}_i(\{s,t\}))\\
&\qquad\qquad (\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))\\
&=(H-1)\sum_i \frac{1}{l(l-1)}\sum_{s\neq t}\big\{(\boldsymbol{\varepsilon}_{is}-\boldsymbol{\phi}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\varepsilon}_{it}-\boldsymbol{\phi}_i(\{s,t\}))-(l-2)^{-2}\mathrm{tr}(\boldsymbol{\Sigma}_0)\big\}\\
&\qquad\qquad (\boldsymbol{\mu}_{is}-\boldsymbol{\omega}_i(\{s,t\}))^{\mathrm{T}}(\boldsymbol{\mu}_{it}-\boldsymbol{\omega}_i(\{s,t\}))\\
&+(H-1)\sum_i \frac{\mathrm{tr}(\boldsymbol{\Sigma}_0)}{(l-2)^2}\sum_{s\neq t}\big(\frac{1}{l}\sum_k \boldsymbol{\mu}_{ik}^T\boldsymbol{\mu}_{ik}-\frac{1}{l(l-1)}\sum_{s\neq t}\boldsymbol{\mu}_{is}^T\boldsymbol{\mu}_{it}\big)\\
&=o_p(\sqrt{\mathrm{var}(\widetilde{T_n})})+O_p(Hp^2l^{-3}\sum_i r_{2i}) = o_p(\sqrt{\mathrm{var}(\widetilde{T_n})}).
\end{aligned}$$

Note that the expectation of the first term in the last equality is exact 0, and with Conditions (C1) and (C2), similar calculations as in (S.2) imply its variance is a higher-order term than $\mathrm{var}(\widetilde{T_n})$, so we need only consider the latter term. Denote $\Lambda_i = l^{-1}\sum_k \boldsymbol{\mu}_{ik}^{\mathrm{T}}\boldsymbol{\mu}_{ik} - \{l(l-1)\}^{-1}\sum_{s\neq t}\boldsymbol{\mu}_{is}^{\mathrm{T}}\boldsymbol{\mu}_{it}$. Let $\boldsymbol{\mu}_i^{\mathrm{T}} = (\boldsymbol{\mu}_{i1}^{\mathrm{T}},\ldots,\boldsymbol{\mu}_{il}^{\mathrm{T}})^{\mathrm{T}}$, $\widetilde{\boldsymbol{\mu}}_i^{\mathrm{T}} = (\boldsymbol{\mu}_{i1}^{\mathrm{T}}-\boldsymbol{\mu}_{i2}^{\mathrm{T}},\ldots,\boldsymbol{\mu}_{i,l-1}^{\mathrm{T}}-\boldsymbol{\mu}_{il}^{\mathrm{T}},\boldsymbol{\mu}_{il}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{\mu}_i^{\mathrm{T}}\mathbf{S}^{-1}$, then Λ_i can be written in quadratic form: $\Lambda_i = l^{-1}\boldsymbol{\mu}_i^{\mathrm{T}}\mathbf{D}\boldsymbol{\mu}_i = l^{-1}\widetilde{\boldsymbol{\mu}}_i^{\mathrm{T}}\mathbf{S}\mathbf{D}\mathbf{S}^{\mathrm{T}}\widetilde{\boldsymbol{\mu}}_i$, where $\mathbf{D}$ is the matrix whose diagonal elements are 1 and off-diagonal elements are $-(l-1)^{-1}$. Note that the last row and last column of $\mathbf{S}\mathbf{D}\mathbf{S}^{\mathrm{T}}$ are all 0, therefore $\Lambda_i \le C' l^{-1}\sum_{k=1}^{l-1}(\boldsymbol{\mu}_{ik} - \boldsymbol{\mu}_{i,k+1})^{\mathrm{T}}(\boldsymbol{\mu}_{ik} - \boldsymbol{\mu}_{i,k+1})$ for some constant C'. By Conditions (C1) and (C2), $\Lambda_i \le C' l^{-1} M^2 p \sum_{k=1}^{l-1}\|\mathbf{Y}_{ik} - \mathbf{Y}_{i,k+1}\|^{2\alpha} = O_p(l^{-1}pr_{2i})$.

Using similar arguments, we can shown the results above hold for all remaining terms in R_1, from which our assertion holds. □

Proof of Theorem 1

Proof. From the proof of Proposition 2, we need only to show the asymptotic normality of $\widetilde{T_n'}$

$$\widetilde{T_n'} = (H-1)\sum_{i=1}^{H}\{l(l-1)\}^{-1}\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2 - 2\sum_{i<j}(l-1)^{-2}\sum_{s,t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{jt})^2.$$

Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_k = \sigma\{\boldsymbol{\varepsilon}_{(1)}, \ldots, \boldsymbol{\varepsilon}_{(k)}\}$ with $k = 1, 2, \ldots, n$, and $E_k(.)$ denote the conditional expectation given $\mathcal{F}_k$, $E_0(.) = E(.)$, then $\{D_k,\ k = 1, \ldots, n\}$ is a martingale difference sequence with respect to the σ-fields $\{\mathcal{F}_k,\ k = 1, \ldots, n\}$, where $\widetilde{T_n'} = \sum_{k=1}^{n} D_k$, $D_k = (E_k - E_{k-1})\widetilde{T_n'}$. By noting the fact the $\boldsymbol{\varepsilon}_{(i)}$'s are conditional independent given $\mathbf{Y}_{(i)}$'s, D_k has the exact form

$$D_k = \frac{2(H-1)}{l(l-1)}\{\boldsymbol{\varepsilon}_{(k)}^{\mathrm{T}}\mathbf{Q}_{k-1}\boldsymbol{\varepsilon}_{(k)} - \mathrm{tr}(\mathbf{Q}_{k-1}\boldsymbol{\Sigma}_0)\} - \frac{2}{l^2}\{\boldsymbol{\varepsilon}_{(k)}^{\mathrm{T}}\mathbf{W}_{k-1}\boldsymbol{\varepsilon}_{(k)} - \mathrm{tr}(\mathbf{W}_{k-1}\boldsymbol{\Sigma}_0)\}, \tag{S.3}$$

where $\mathbf{Q}_{k-1} = \sum_{s=(i-1)*l+1}^{k-1}(\boldsymbol{\varepsilon}_{(s)}\boldsymbol{\varepsilon}_{(s)}^{\mathrm{T}} - \boldsymbol{\Sigma}_0)$, $\mathbf{W}_{k-1} = \sum_{s=1}^{(i-1)*l}(\boldsymbol{\varepsilon}_{(s)}\boldsymbol{\varepsilon}_{(s)}^{\mathrm{T}} - \boldsymbol{\Sigma}_0)$ and $k = (i-1)l + j$.

To apply martingale central limit theorem (Hall and Hype, 1980) to establish the limiting distribution of $\widetilde{T_n'}$, we would further verify the following two conditions:

$$\frac{\sum_{k=1}^n \sigma_k^2}{\mathrm{var}(\widetilde{T_n'})} \xrightarrow{p} 1, \quad \text{and} \quad \frac{\sum_{k=1}^n E(D_k^4)}{\mathrm{var}^2(\widetilde{T_n'})} \xrightarrow{p} 0, \tag{S.4}$$

where $\sigma_k^2 = E_{k-1}(D_k^2)$. As it is true that $E(\sum_{k=1}^n \sigma_k^2) = \mathrm{var}(\widetilde{T_n'})$, it suffices to show $\mathrm{var}(\sum_{k=1}^n \sigma_k^2) = o(\mathrm{var}^2(\widetilde{T_n'}))$. By Lemma 1, we express $\sum_{k=1}^n \sigma_k^2$ in the following form:

$$\begin{aligned}
\sum_{k=1}^n \sigma_k^2 =& \frac{4(H-1)^2}{l^2(l-1)^2}\sum_{k=1}^n \Big\{2\mathrm{tr}\{(\mathbf{Q}_{k-1}\boldsymbol{\Sigma}_0)^2\} \\
&+ E_{k-1}\{\Delta(\mathbf{Y}_{(k)})\mathrm{tr}(\Gamma^{\mathrm{T}}(\mathbf{Y}_{(k)})\mathbf{Q}_{k-1}\Gamma(\mathbf{Y}_{(k)}) \circ \Gamma^{\mathrm{T}}(\mathbf{Y}_{(k)})\mathbf{Q}_{k-1}\Gamma(\mathbf{Y}_{(k)}))\}\Big\} \\
&+ \frac{4}{l^4}\sum_{k=1}^n \Big\{2\mathrm{tr}\{(\mathbf{W}_{k-1}\boldsymbol{\Sigma}_0)^2\} \\
&+ E_{k-1}\{\Delta(\mathbf{Y}_{(k)})\mathrm{tr}(\Gamma^{\mathrm{T}}(\mathbf{Y}_{(k)})\mathbf{W}_{k-1}\Gamma(\mathbf{Y}_{(k)}) \circ \Gamma^{\mathrm{T}}(\mathbf{Y}_{(k)})\mathbf{W}_{k-1}\Gamma(\mathbf{Y}_{(k)}))\}\Big\} \\
&- \frac{8(H-1)}{l^3(l-1)}\sum_{k=1}^n \Big\{2\mathrm{tr}(\mathbf{Q}_{k-1}\boldsymbol{\Sigma}_0)\mathrm{tr}(\mathbf{W}_{k-1}\boldsymbol{\Sigma}_0)
\end{aligned}$$

$$
\begin{aligned}
&+ E_{k-1}\{\Delta(\mathbf{Y}_{(k)})\mathrm{tr}(\Gamma^{\mathrm{T}}(\mathbf{Y}_{(k)})\mathbf{Q}_{k-1}\Gamma(\mathbf{Y}_{(k)}) \circ \Gamma^{\mathrm{T}}(\mathbf{Y}_{(k)})\mathbf{W}_{k-1}\Gamma(\mathbf{Y}_{(k)}))\}\Big\} \\
\equiv& D_{1,1} + D_{1,2} + D_{2,1} + D_{2,2} + D_{3,1} + D_{3,2}.
\end{aligned}
$$

Now we calculate the variance of $D_{i,j}$ $(i = 1, 2, 3; j = 1, 2)$. Consider firstly $D_{1,1}$. Using Lemma 1, for any positive integer $k \le r$, with $k = (i-1)l + j$, we have

$$
\begin{aligned}
&\mathrm{cov}(\mathrm{tr}\{(\mathbf{Q}_{k-1}\mathbf{\Sigma}_0)^2\}, \mathrm{tr}\{(\mathbf{Q}_{r-1}\mathbf{\Sigma}_0)^2\}) \\
&\quad= \mathrm{cov}(\mathrm{tr}\{(\mathbf{Q}_{k-1}\mathbf{\Sigma}_0)^2\}, \mathrm{tr}\{(\mathbf{Q}_{k-1}\mathbf{\Sigma}_0)^2\}) \\
&\quad= 2j(j-1)\mathrm{var}\big\{\mathrm{tr}\{(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^{\mathrm{T}} - \mathbf{\Sigma}_0)\mathbf{\Sigma}_0(\boldsymbol{\varepsilon}_2\boldsymbol{\varepsilon}_2^{\mathrm{T}} - \mathbf{\Sigma}_0)\mathbf{\Sigma}_0\}\big\} \\
&\qquad + j\mathrm{var}\big\{\mathrm{tr}\{(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^{\mathrm{T}} - \mathbf{\Sigma}_0)\mathbf{\Sigma}_0\}^2\big\} \\
&\quad= O(j^2)\mathrm{tr}^2(\mathbf{\Sigma}_0^2)\mathrm{tr}(\mathbf{\Sigma}_0^4). \qquad \text{(S.5)}
\end{aligned}
$$

Then we can rewrite $\sum_{k=1}^{n} \mathrm{tr}\{(\mathbf{Q}_{k-1}\mathbf{\Sigma}_0)^2\}$ as $\sum_{i=1}^{H}\sum_{j=1}^{l} \mathrm{tr}\{(\mathbf{Q}_{(i-1)l+j-1}\mathbf{\Sigma}_0)^2\}$ based on the specific form of Q_{k-1}. Thus, $\mathrm{var}(D_{1,1}) = O(H^5 l^{-4})\mathrm{tr}^2(\mathbf{\Sigma}_0^2)\mathrm{tr}(\mathbf{\Sigma}_0^4)$ and by Condition (C3),

$$
\mathrm{var}(D_{1,1})/\mathrm{var}^2(\widetilde{T_n'}) = O(\{H\mathrm{tr}^2(\mathbf{\Sigma}_0^2)\}^{-1}\mathrm{tr}(\mathbf{\Sigma}_0^4)) \to 0.
$$

Next, consider the part $D_{1,2}$. Since $\triangle(\mathbf{Y})$ is uniformly bounded by some constant $\triangle_0$ and as well as $\Gamma(\mathbf{Y})$, we have $D_{1,2} \le D_{1,3}$ where

$$
D_{1,3} = \frac{4(H-1)^2}{l^2(l-1)^2}\sum_{k=1}^{n}\Delta_0\mathrm{tr}(\Gamma^{\mathrm{T}}\mathbf{Q}_{k-1}\Gamma \circ \Gamma^{\mathrm{T}}\mathbf{Q}_{k-1}\Gamma).
$$

By similar argument, we can verify that the result for $D_{1,1}$ also holds for $D_{1,3}$ and for all $D_{i,j}$ with $i = 1, 2, 3; j = 1, 2$. Hence we complete the first part of (S.4).

Now we show the second part of (S.4) holds. Note that $E_{k-1}\{\boldsymbol{\varepsilon}_{(k)}^{\mathrm{T}}\mathbf{Q}_{k-1}\boldsymbol{\varepsilon}_{(k)}\} = \mathrm{tr}(\mathbf{Q}_{k-1}\boldsymbol{\Sigma}_0)$ and $E_{k-1}\{\boldsymbol{\varepsilon}_{(k)}^{\mathrm{T}}\mathbf{W}_{k-1}\boldsymbol{\varepsilon}_{(k)}\} = \mathrm{tr}(\mathbf{W}_{k-1}\boldsymbol{\Sigma}_0)$. By the part (iv) of Lemma 1, we have

$$\begin{aligned}\sum_{k=1}^{n} E(D_k^4) \leq & 8\Big[\frac{16(H-1)^4}{l^4(l-1)^4}\sum_{k=1}^{n} E\big\{(\boldsymbol{\varepsilon}_{(k)}^{\mathrm{T}}\mathbf{Q}_{k-1}\boldsymbol{\varepsilon}_{(k)} - \mathrm{tr}(\mathbf{Q}_{k-1}\boldsymbol{\Sigma}_0))^4\big\} \\ & + \frac{16}{l^8}\sum_{k=1}^{n} E\big\{(\boldsymbol{\varepsilon}_{(k)}^{\mathrm{T}}\mathbf{W}_{k-1}\boldsymbol{\varepsilon}_{(k)} - \mathrm{tr}(\mathbf{W}_{k-1}\boldsymbol{\Sigma}_0))^4\big\}\Big] \\ \leq & M(H^5 l^{-5}\mathrm{tr}^2(\boldsymbol{\Sigma}_0^4) + H^3 l^{-5}\mathrm{tr}^2(\boldsymbol{\Sigma}_0^4)) \\ = & o(\mathrm{var}^2(\widetilde{T_n'})),\end{aligned}$$

where the last two inequalities come from (S.5) and Condition (C3). The second part of (S.4) is proved and thus the proof is completed. □

Proof of Proposition 4

Proof. Similar to the proof of Proposition 2, we can rewrite T_n' as

$$T_n' = T_n^* + \{(J_n - U_n) + (U_n - \widetilde{W}) + (\widetilde{W} - W)\}.$$

By (S.1), T_n^* can be represented as

$$T_n^* = (H-1)\sum_{i=1}^{H}\Big[\Big\{\frac{1}{l(l-1)} + \frac{2}{l(l-1)(l-2)} + \frac{2(l-3)}{l(l-1)(l-2)^3}\Big\}\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2$$

$$
\begin{aligned}
&+ \left\{ \frac{-2}{l(l-1)(l-2)} + \frac{-8(l-3)}{l(l-1)(l-2)^3} + \frac{4(l-3)(l-4)}{l(l-1)(l-2)^4} \right\} \sum_{s,t,r}^{*} \boldsymbol{\varepsilon}_{is}^{\mathrm{T}} \boldsymbol{\varepsilon}_{ir} \boldsymbol{\varepsilon}_{ir}^{\mathrm{T}} \boldsymbol{\varepsilon}_{it} \\
&+ \frac{l+4}{l(l-2)^4} \sum_{s,t,r,q}^{*} \boldsymbol{\varepsilon}_{is}^{\mathrm{T}} \boldsymbol{\varepsilon}_{it} \boldsymbol{\varepsilon}_{ir}^{\mathrm{T}} \boldsymbol{\varepsilon}_{iq} \Bigg] \\
&- 2 \sum_{i<j} \left\{ \frac{1}{(l-1)^2} \sum_{s,t} (\boldsymbol{\varepsilon}_{is}^{\mathrm{T}} \boldsymbol{\varepsilon}_{jt})^2 - \frac{1}{(l-1)^3} \sum_{s} \sum_{t \neq r} \boldsymbol{\varepsilon}_{jt}^{\mathrm{T}} \boldsymbol{\varepsilon}_{is} \boldsymbol{\varepsilon}_{is}^{\mathrm{T}} \boldsymbol{\varepsilon}_{jr} \right. \\
&\left. - \frac{1}{(l-1)^3} \sum_{s} \sum_{t \neq r} \boldsymbol{\varepsilon}_{it}^{\mathrm{T}} \boldsymbol{\varepsilon}_{js} \boldsymbol{\varepsilon}_{js}^{\mathrm{T}} \boldsymbol{\varepsilon}_{ir} + \frac{1}{(l-1)^4} \sum_{s \neq t} \sum_{r \neq q} \boldsymbol{\varepsilon}_{is}^{\mathrm{T}} \boldsymbol{\varepsilon}_{jr} \boldsymbol{\varepsilon}_{it}^{\mathrm{T}} \boldsymbol{\varepsilon}_{jq} \right\} + R_1 \\
=&(H-1) \sum_{i=1}^{H} A_i^* - 2 \sum_{i<j} C_{ij}^* + R_1 \equiv \widetilde{T_n^*} + R_1,
\end{aligned}
$$

Now we elaborately study the variance of $\widetilde{T_n^*}$. It is easy to see $\mathrm{var}(\widetilde{T_n^*}) = H(H-1)^2 \mathrm{var}(A_i^*) - 4H(H-1)^2 \mathrm{cov}(A_i^*, C_{ij}^*) + 4H(H-1)(H-2)\mathrm{cov}(C_{ij}^*, C_{ij'}^*) + 2H(H-1)\mathrm{var}(C_{ij}^*)$. By Lemma 1 again and similar arguments for (S.2) in Proposition 1, under H_0,

$$
\begin{aligned}
\mathrm{var}(\widetilde{T_n^*}) = 4H^2(H-1) \Bigg[&\left\{ \frac{l}{(l-1)(l-2)^2} + \frac{2}{l(l-1)(l-2)} \right\} \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) \\
&+ O(l^{-4}) \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + O(l^{-3}) \mathrm{tr}(\boldsymbol{\Sigma}_0^4) + O(l^{-2}H^{-1}) \mathrm{tr}(\boldsymbol{\Sigma}_0^4) \\
&+ O(l^{-2})(\delta_2 - \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2)) + O(l^{-2})(\delta_3 - 3\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2)) + O(l^{-3})\delta_4 \Bigg] \\
= 4H^2(H-1) \Bigg[&\left\{ \frac{l}{(l-1)(l-2)^2} + \frac{2}{l(l-1)(l-2)} \right\} \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) \\
&+ o(l^{-3}) \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + O(l^{-2}) \mathrm{tr}(\boldsymbol{\Sigma}_0^4) \Bigg] \\
= 4H^2(H-1) \Bigg[&\left\{ \frac{l}{(l-1)(l-2)^2} + \frac{2}{l(l-1)(l-2)} \right\} \mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) \Bigg] (1 + o(1)),
\end{aligned}
$$

where $\delta_4 = E\{(\boldsymbol{\varepsilon}_{i1}^{\mathrm{T}} \boldsymbol{\varepsilon}_{i2})^2 \boldsymbol{\varepsilon}_{i1}^{\mathrm{T}} \boldsymbol{\varepsilon}_{i3} \boldsymbol{\varepsilon}_{i3}^{\mathrm{T}} \boldsymbol{\varepsilon}_{i2}\}$. The second equation holds because

$\delta_2 = \mathrm{tr}^2(\mathbf{\Sigma}_0^2) + O(\mathrm{tr}(\mathbf{\Sigma}_0^4))$, $\delta_3 = 3\mathrm{tr}^2(\mathbf{\Sigma}_0^2) + O(\mathrm{tr}(\mathbf{\Sigma}_0^4))$, and

$$\delta_4 \leq \{\mathrm{var}((\boldsymbol{\varepsilon}_{i1}^{\mathrm{T}}\boldsymbol{\varepsilon}_{i2})^2)\mathrm{var}(\boldsymbol{\varepsilon}_{i1}^{\mathrm{T}}\boldsymbol{\varepsilon}_{i3}\boldsymbol{\varepsilon}_{i3}^{\mathrm{T}}\boldsymbol{\varepsilon}_{i2})\}^{1/2} = \mathrm{tr}(\mathbf{\Sigma}_0^2)\mathrm{tr}^{1/2}(\mathbf{\Sigma}_0^4) = o(\mathrm{tr}^2(\mathbf{\Sigma}_0^2)).$$

The last equation is due to the condition $\mathrm{tr}(\mathbf{\Sigma}_0^4) = o(l^{-1})\mathrm{tr}^2(\mathbf{\Sigma}_0^2)$.

Now we consider the remaining terms. Since $p = o(\{\sum_i r_{1i}^{4\alpha}\}^{-1}n^{1/2}l^{-5/2})$ and $p = o(\{\sum_i r_{2i}\}^{-1}n^{1/2}l^{1/2})$, by the same arguments used in Proposition 1 and 2, it can be seen that $R_1 = o_p(l^{-1}\sqrt{\mathrm{var}(\widetilde{T_n^*})})$, which still holds for $U_n - \widetilde{W}$ and $\widetilde{W} - W$. Combining these results with the fact $J_n - U_n = o_p(l^{-1}\sqrt{\mathrm{var}(\widetilde{T_n^*})})$ when $p = o(l^5)$, we can complete the proof of our Proposition 4. □

Proof of Proposition 2

Proof. From the proof of Proposition 1, the restriction $p = o(l^3)$ mainly comes from the term in the T_n

$$\begin{aligned} J_n =& (H-1)\sum_{i=1}^{H}\left\{\frac{-2l+8}{l(l-1)(l-2)^4}\sum_{s,t,r}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir} + \frac{-4}{l(l-1)(l-2)^3}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\right. \\ &\left. + \frac{1}{l(l-2)^3}\sum_{s}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is})^2 + \frac{l-3}{l(l-1)(l-2)^3}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\right\}. \end{aligned}$$

It can be shown that $\mathrm{var}(J_n)/\mathrm{var}(\widetilde{T_n}) = O(l^{-3}\mathrm{tr}^{-1}(\mathbf{\Sigma}_0^2)\mathrm{tr}^2(\mathbf{\Sigma}_0))$, from which we can observe the requirement on l.

Consider the modified test statistic T_n'. Let $W = H(H-1)\{\widehat{\mathrm{tr}(\mathbf{\Sigma}_0)}\}^2/(l-$

$2)^2$, and T_n' can be rewritten as

$$\begin{aligned} T_n' =&(H-1)\sum_{i=1}^{H} A_i - 2\sum_{i<j} C_{ij} - W \\ =&\left\{\left((H-1)\sum_{i=1}^{H} A_i - J_n\right) - 2\sum_{i<j} C_{ij}\right\} + (J_n - W) \\ =&T_n^* + (J_n - W). \end{aligned}$$

By similar arguments in Proposition 1, it can be verified that without the condition $p = o(l^3)$,

$$E(T_n^*) = \left\{2/(l-2)+2/(l-2)^2-2/(l-1)-1/(l-1)^2\right\}H(H-1)\mathrm{tr}(\mathbf{\Sigma}_0^2)+o(\sqrt{\mathrm{var}(\widetilde{T_n})})$$

and $\mathrm{var}(T_n^*) = \mathrm{var}(\widetilde{T_n})(1+o(1))$. Then we need to deal with the terms $J_n - W$. Define $U_n = (l-2)^{-2}(H-1)\sum_{i=1}^{H}\left(l^{-1}\sum_s \boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is} - \{l(l-1)\}^{-1}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\right)^2$. Some calculations yield that

$$\begin{aligned} J_n - U_n =&(l-2)^{-2}(H-1)\sum_{i=1}^{H}\left\{\frac{8}{l^2(l-1)(l-2)^2}\sum_{s,t,r}^{*}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{ir}\right. \\ &-\frac{8}{l^2(l-1)(l-2)}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it} + \frac{2}{l^2(l-2)}\sum_{s}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is})^2 \\ &\left.-\frac{2}{l^2(l-1)(l-2)}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{is}\boldsymbol{\varepsilon}_{it}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it} - \frac{1}{l^2(l-1)^2}\sum_{s\neq t}\sum_{r\neq q}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{ir}^{\mathrm{T}}\boldsymbol{\varepsilon}_{iq}\right\}. \end{aligned}$$

Taking similar procedures as for (A.2) and (A.3), we can show that $E(J_n - U_n) = O(H^2 l^{-3}\mathrm{tr}(\mathbf{\Sigma}_0^2)) = o(\sqrt{\mathrm{var}(\widetilde{T_n})})$ and $\mathrm{var}(J_n - U_n)/\mathrm{var}(\widetilde{T_n}) = O(l^{-7}\mathrm{tr}^{-1}(\mathbf{\Sigma}_0^2)\mathrm{tr}^2(\mathbf{\Sigma}_0))$. Consequently, it is only required that $p = o(l^7)$.

It remains to verify that $W - U_n = o_p(\sqrt{\mathrm{var}(\widetilde{T_n})})$. Let $\widetilde{W} = (l -$

$2)^2H(H-1)(H^{-1}\sum_i B_i)^2$, where $B_i = l^{-1}\sum_s \boldsymbol{\varepsilon}_{is}^T\boldsymbol{\varepsilon}_{is} - \{l(l-1)\}^{-1}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it}$. Note that $E(B_i) = \mathrm{tr}(\boldsymbol{\Sigma}_0)$ and $\mathrm{var}(B_i) = O(l^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_0^2))$. Observe

$$\begin{aligned}
\widetilde{W} - U_n =& \frac{(H-1)}{(l-2)^2}\sum_i \big\{(B_i - \mathrm{tr}(\boldsymbol{\Sigma}_0))(B_i - H^{-1}\sum_j B_j)\big\} \\
=& O(Hl^{-5/2}\mathrm{tr}^{1/2}(\boldsymbol{\Sigma}_0^2))\sum_i\{B_i - \mathrm{tr}(\boldsymbol{\Sigma}_0)\} \\
=& O_p(H^{3/2}l^{-3}\mathrm{tr}(\boldsymbol{\Sigma}_0^2)) = O_p(l^{-2})\sqrt{\mathrm{var}(\widetilde{T_n})} \\
=& o_p(\sqrt{\mathrm{var}(\widetilde{T_n})}).
\end{aligned}$$

Also,

$$\begin{aligned}
W - \widetilde{W} =& \frac{H(H-1)}{(l-2)^2}\big\{(H^{-1}\sum_i R_i)^2 + 2H^{-2}\sum_i B_i \sum_i R_i\big\} \\
R_i =& 2l^{-1}\sum_s \boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\mu}_{is} - 2\{l(l-1)\}^{-1}\sum_{s\neq t}\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\mu}_{it} \\
&+ l^{-1}\sum_s \boldsymbol{\mu}_{is}^{\mathrm{T}}\boldsymbol{\mu}_{is} - \{l(l-1)\}^{-1}\sum_{s\neq t}\boldsymbol{\mu}_{is}^{\mathrm{T}}\boldsymbol{\mu}_{it} \\
\equiv& R_{i1} + R_{i2}.
\end{aligned}$$

As mentioned in Proposition 1, $R_{i2} = \Lambda_i = O_p(l^{-1}pr_{2i})$, we need only consider the terms R_{i1}. By the similar argument, $\sum_i R_{i1}$ is a higher-order term than $\sum_i R_{i2}$. Combine all these results, we have

$$\begin{aligned}
W - \widetilde{W} =& H(H-1)/(l-2)^2 O_p(p^2/n\sum_i r_{2i}) \\
=& \sqrt{\mathrm{var}(\widetilde{T_n})}O_p(H^{-1/2}l^{-2}p\sum_i r_{2i})
\end{aligned}$$

$$= o_p(\sqrt{\mathrm{var}(\widetilde{T_n})}),$$

from which we complete the proof of Proposition 2. □

Proof of Proposition 3

Proof. Rewrite $\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_0^2)}$ as

$$\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_0^2)} = \left\{1 + \frac{2}{l-2} + \frac{2}{(l-2)^2}\right\}^{-1} \left\{\left(\frac{1}{H}\sum_{i=1}^{H} A_i - \frac{J_n}{H(H-1)}\right) + \frac{J_n - W}{H(H-1)}\right\}$$

$$\equiv e_1 + e_2.$$

Similar to the proof of Proposition 1, we can show that

$$E(e_1) = \mathrm{tr}(\boldsymbol{\Sigma}_0^2) + O(l^{-3}\mathrm{tr}(\boldsymbol{\Sigma}_0^2)),$$

$$\mathrm{var}(e_1) = H^{-4}\mathrm{var}\{(H-1)\sum_i A_i - J_n\}(1+o(1)) = O\big(n^{-2}\mathrm{tr}^2(\boldsymbol{\Sigma}_0^2) + (H^2 l)^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_0^4)\big).$$

Accordingly, $\big\{e_1 - \mathrm{tr}(\boldsymbol{\Sigma}_0^2)\big\}/\mathrm{tr}(\boldsymbol{\Sigma}_0^2) \to 0$. By Proposition 2 and taking similar procedures, we have $J_n - W = o_p\{H^{3/2} l^{-1} \mathrm{tr}(\boldsymbol{\Sigma}_0^2)\}$. Thus, $e_2/\mathrm{tr}(\boldsymbol{\Sigma}_0^2) \to 0$. □

Proof of Theorem 2

Proof. For technical convenience, we assume that the n data can be exactly divided to H equal slices. Denote E_0 and E_1 represent the expectation under H_0 and under H_1 respectively. If a slice A_i falls into the area $\{\mathbf{Y} < a\}$, then

$E_1(A_i) = E_0(A_i)$. On the contrary, if it falls into the area $\{\mathbf{Y} > a\}$, we have $E_1(A_i) = (1+\theta_n)^2 E_0(A_i)$. Also note that the number of slices which partly falls into the former and rest falls into the latter is at most one, there is little need to consider this situation as long as H is sufficiently large. So we further assume H is an even number. Similar to the proof of Proposition, by Condition (C2) that $n/l^5 \to 0$ and the fact that $b_l/a_l = 1 + O(l^{-2})$, we have

$$\begin{aligned}
E_1\{T_n'\} =&(H-1)\sum_i E_1(A_i) - 2\sum_{i<j} E_1(C_{ij}) - E_1(W)\\
=&\frac{H(H-1)}{2}\Big\{a_l \mathrm{tr}(\mathbf{\Sigma}_0^2) + (l-2)^{-2}\mathrm{tr}^2(\mathbf{\Sigma}_0) + a_l(1+\theta_n)^2\mathrm{tr}(\mathbf{\Sigma}_0^2)\\
&+(l-2)^{-2}(1+\theta_n)^2\mathrm{tr}^2(\mathbf{\Sigma}_0) + O((l-2)^{-3}\mathrm{tr}(\mathbf{\Sigma}_0^2))\Big\} - (\frac{H}{2}-1)\frac{H}{2}b_l\mathrm{tr}(\mathbf{\Sigma}_0^2)\\
&-(\frac{H}{2}-1)\frac{H}{2}b_l(1+\theta_n)^2\mathrm{tr}(\mathbf{\Sigma}_0^2) - \frac{H^2}{2}b_l(1+\theta_n)\mathrm{tr}(\mathbf{\Sigma}_0^2)\\
&-H(H-1)(l-2)^{-2}\{1+\theta_n/2\}^2\mathrm{tr}^2(\mathbf{\Sigma}_0) + o(\sigma_{T_n',0})\\
=&b_l H^2\theta_n^2\mathrm{tr}(\mathbf{\Sigma}_0^2)/4 + H(H-1)(l-2)^{-2}\theta_n^2\mathrm{tr}^2(\mathbf{\Sigma}_0)/4\\
&+(a_l - b_l)H(H-1)\{1+(1+\theta_n)^2\}\mathrm{tr}(\mathbf{\Sigma}_0^2)/2 + o(\sigma_{T_n',0})\\
\equiv&\mu_{T_n',1} + o(\sigma_{T_n',0}).
\end{aligned}$$

If a slice A_i falls into the area $\{\mathbf{Y} > a\}$, then

$$\mathrm{var}(\sum_{s\neq t}(\boldsymbol{\varepsilon}_{is}^{\mathrm{T}}\boldsymbol{\varepsilon}_{it})^2) = l(l-1)\{4(1+\theta_n)^4\mathrm{tr}^2(\mathbf{\Sigma}_0^2) + O(\mathrm{tr}(\mathbf{\Sigma}_0^4))\}.$$

Tedious calculations yield

$$\sigma^2_{T_n,1} = 2l^{-2}H(H-1)^2\mathrm{tr}^2(\mathbf{\Sigma}_0^2)\{1+(1+\theta_n)^4\}(1+o(1)) = \sigma^2_{T_n,0}\{1+(1+\theta_n)^4\}/2.$$

Under the alternative H_1,

$$\widehat{\sigma}_{T'_n,0} = 2l^{-1}H^{3/2}\{\frac{1+(1+\theta_n)^2}{2}\mathrm{tr}(\mathbf{\Sigma}_0^2) + \frac{\theta_n^2\mathrm{tr}^2(\mathbf{\Sigma}_0)}{4l^2}\}(1+o_p(1)).$$

Similar to the proof of Theorem 1, we can establish the asymptotic normality of $\widetilde{T'_n}$. Here we omit the details. □

Further simulation results

An application of our proposed test lies in sufficient dimension reduction (SDR) which tries to reduce the dimension by replacing original predictors with a minimal set of their linear combinations without loss of information in regression. Many SDR methods were developed based on the paradigm of inverse regression (Li, 1991) and they usually rely on the validity of the constant variance condition (Cook and Weisberg, 1991), i.e., $\mathrm{cov}(\mathbf{X} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}) = \mathbf{\Sigma}_0$, where $\boldsymbol{\beta} \in \mathbb{R}^{p\times d}$ is a basis matrix of the central subspace and $\mathbf{X}$ stands for the predictor vector. In reality, we can test if $\mathrm{cov}(\mathbf{X} \mid \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X})$ is approximately a constant matrix at the value $\hat{\boldsymbol{\beta}}$ that is close to the true $\boldsymbol{\beta}$. In order not to affect the validity of detection, we can divide data into two parts, the one for estimating $\hat{\boldsymbol{\beta}}$, and the other for testing.

We added some numerical results regarding this strategy. To show the performance of empirical size, we use the simple linear model 1

$$x_{ik} \sim N(0, 0.2) \text{ independently for } k = 1, 2$$

$$x_{ik} \sim N(0, 1 + 0.6 \cdot I(x_{i2} > 0)) \text{ for } k = 3, \ldots, p$$

$$y_i = 4x_{i1} + \delta\epsilon_i$$

to generate $n = 320$ data points. ϵ_i is a standard normal distribution, independent of x_i's. Clearly noting that the central subspace contains only one SDR direction, standardized as $\beta = (1, 0, \ldots, 0)^{\mathrm{T}}$ and $\mathrm{cov}(\mathbf{X} \mid x_k)$ is constant only for $k = 1$. This setting greatly diminishes the number of β which makes null hypothesis hold. Similarly, to show the performance of empirical power, we use model 2 followed

$$x_{ik} \sim N(0, 1) \text{ independently for } k = 1, 3, \ldots, p$$

$$x_{i2} = (p-2)^{-1/2} \sum_{k=3}^{p} (x_{ik}^2 - 1) + e_i$$

$$y_i = 4x_{i2} + \delta\epsilon_i$$

where the SDR direction is $\beta = (0, 1, 0, \ldots, 0)^{\mathrm{T}}$, e_i, ϵ_i are both standard normal distribution, independent of x_i's. Similar to model 1, $\mathrm{cov}(\mathbf{X} \mid x_k)$ depends on x_k only for $k = 2$, so that it greatly diminishes the number of β which violates H_0. Here we use the first $n_1 = 160$ data points to estimate β by SIR, then use the rest $n - n_1 = 160$ data points to test

H_{01} : $\mathrm{cov}(\mathbf{X} \mid \beta^{\mathrm{T}}\mathbf{X}) = \mathbf{\Sigma}_0$ and H_{02} : $\mathrm{cov}(\mathbf{X} \mid \hat{\beta}^{\mathrm{T}}\mathbf{X}) = \mathbf{\Sigma}_0$, respectively. All the simulation results are obtained based on 1,000 repetitions and the nominal level is fixed as 0.05.

For the first 160 data, the classical SIR doesn't work in the "large p, small n" cases. Actually, there are many statistical methods that estimating central subspace for high dimensional data, for example the SSIR method (Ni, Cook and Tsai, 2005) and the CISE method (Chen, Zou and Cook , 2010). Here for convenience, we simply consider the relative small dimension $p = 20$, and obtain the estimate by classical SIR procedure. Table 1 shows the simulation result and its performance is reasonably well.

References

Cook, R. D. and Weisberg, S. (1991). Discussion of "Sliced inverse regression for dimension reduction". *J. Amer. Statist. Assoc.* **86**, 328–332.

Chen, X., Zou, C. and Cook, R.D. (2010). Coordinate-independent sparse sufficient dimension reduction and variable selection. *Ann. Statist.* **38**, 3696–3723.

Chen, S. X., Zhang, L,-X. and Zhong, P,-S. (2010). Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **105**, 810–819.

Hall, P. and Hyde, C. C. (1980). Martingale limit theory and its applications. *Academic press, New York.*

Table 1: The performance based on our strategy, both in model 1 and in model 2.

	Model 1		Model 2	
l	β	$\hat{\beta}$	β	$\hat{\beta}$
		$\delta = 0.5$		
8	7.9	7.9	38.6	21.6
10	7.2	8.3	41.1	25.0
20	8.0	9.0	65.8	38.2
		$\delta = 1$		
8	6.9	8.3	35.9	22.0
10	7.5	8.7	41.5	25.5
20	8.1	9.3	62.9	39.7

Li, K. C. (1991). Sliced inverse regression for dimension reduction(with discussion). *J. Am. Statist. Assoc.* **86**, 316–342.

Ni, L., Cook, R.D. and Tsai, C.L. (2005). A note on shrinkage sliced inverse regression. *IEEE Trans. Signal Process.* **50**, 635–650.

Table 2: Empirical sizes at 5% significance under the model of $d = 1$ and $y \sim U(2, 4)$.

$\mathbf{X} \mid Y$			Normal				Gamma			
	$\boldsymbol{\mu}_Y$		Case (I)		Case (II)		Case (I)		Case (II)	
methods			T'_n	$\tilde{T}_n$	T'_n	$\tilde{T}_n$	T'_n	$\tilde{T}_n$	T'_n	$\tilde{T}_n$
l	n	p	$\rho = 0$							
10	200	20	5.8	8.5	6.0	8.6	6.3	8.9	6.6	9.2
		40	5.2	7.3	5.3	7.6	5.9	8.8	5.5	7.3
		100	5.4	8.5	5.8	9.0	4.3	7.2	4.6	8.7
		200	5.1	7.8	7.0	9.9	3.6	6.8	5.4	8.9
		1000	4.7	7.2	13.9	19.8	5.1	6.9	11.3	15.1
	600	20	6.7	9.7	6.9	9.3	8.3	10.5	8.3	10.8
		40	5.8	8.4	5.8	8.6	5.8	9.3	5.6	9.6
		100	5.5	8.5	5.6	8.8	5.1	7.3	4.9	7.1
		200	5.8	9.0	5.9	9.0	5.0	7.8	5.2	7.7
		1000	4.5	6.9	4.6	7.0	5.5	8.1	6.2	8.4
15	200	20	4.9	6.7	5.5	8.0	7.2	8.8	7.5	10.0
		40	4.9	6.2	5.2	6.9	6.3	8.3	7.6	8.9
		100	4.9	7.2	7.9	10.2	6.4	8.7	8.2	10.9
		200	4.3	5.8	10.1	13.4	4.9	6.4	10.1	13.1
		1000	4.5	6.3	42.4	46.5	5.0	6.9	41.1	45.8
	600	20	5.1	6.6	5.3	6.5	8.0	9.6	8.6	9.9
		40	5.0	6.6	4.9	6.6	4.9	6.8	5.3	6.4
		100	4.0	6.0	4.1	6.1	6.8	8.3	6.9	8.6
		200	4.1	6.3	4.1	5.9	5.0	6.8	5.6	7.3
		1000	4.7	7.2	6.1	7.7	4.8	6.7	5.2	7.4

Table 3: Empirical sizes at 5% significance under the model of $d=1$ and $y \sim N(3, 0.2)$.

$\mathbf{X} \mid Y$			Normal				Gamma			
	$\boldsymbol{\mu}_Y$		Case (I)		Case (II)		Case (I)		Case (II)	
methods			T_n'	$\tilde{T}_n$	T_n'	$\tilde{T}_n$	T_n'	$\tilde{T}_n$	T_n'	$\tilde{T}_n$
l	n	p				$\rho=0$				
10	200	20	4.4	8.2	5.0	8.4	7.2	10.5	6.4	8.8
		40	5.4	9.1	5.8	9.2	7.6	10.2	7.0	10.0
		100	4.0	6.0	5.4	8.7	7.1	9.1	5.8	8.2
		200	4.7	7.4	2.9	5.4	5.9	7.8	4.9	7.3
		1000	4.2	8.3	6.0	9.9	5.2	9.2	6.9	9.7
	600	20	6.2	9.4	4.9	6.5	7.4	10.2	8.9	11.3
		40	5.7	8.3	5.7	7.7	7.2	10.0	5.9	8.5
		100	5.8	9.7	5.2	8.9	5.5	7.5	5.5	8.5
		200	5.6	8.6	5.1	8.7	4.7	7.7	5.3	8.4
		1000	5.3	8.4	4.7	8.1	6.3	8.8	6.5	9.5
15	200	20	5.4	7.2	6.8	9.0	7.0	8.2	7.4	9.4
		40	4.8	6.8	5.2	7.3	5.4	7.6	5.7	8.0
		100	5.4	7.4	5.1	7.3	4.6	6.4	5.2	7.3
		200	6.0	7.4	6.7	8.8	5.6	7.1	6.4	8.9
		1000	5.9	7.3	7.9	11.0	4.8	6.4	9.5	12.1
	600	20	4.4	6.5	5.9	7.8	8.1	9.5	6.7	8.7
		40	5.5	7.4	6.3	8.1	5.8	7.7	5.7	7.9
		100	4.5	5.5	5.4	7.2	5.9	7.5	6.0	7.5
		200	5.7	7.3	5.5	7.6	5.0	6.7	6.4	8.2
		1000	4.9	7.4	7.3	8.6	4.6	6.8	7.2	8.3