Biometrika (2018), **xx**, x, *pp*. 1–22
Printed in Great Britain

Supplementary material for 'Homogeneity tests of covariance matrices with high-dimensional longitudinal data'

BY PING-SHOU ZHONG

Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, Illinois 60607, U.S.A.

pszhong@uic.edu

RUNZE LI

Department of Statistics and The Methodology Center, Pennsylvania State University University Park, Pennsylvania 16802-2111, U.S.A.

rzli@psu.edu

AND SHAWN SANTO

Department of Statistics and Probability, Michigan State University, Michigan 48824, U.S.A.

santosha@stt.msu.edu

1. PROOF OF LEMMAS

In this section, we present the proofs to some lemmas used in the proofs of the main theorems. Without loss of generality, assume that $\mu_t = 0$ in our proofs for each $t \in \{1, \ldots, T\}$ because the test statistic, $\hat{D}_{nt}$, is invariant with respect to μ_t.

LEMMA 1. *(i) For any symmetric matrices A and B with appropriate dimensions, we have* $\mathrm{tr}^2(AB) \le \mathrm{tr}(A^2)\mathrm{tr}(B^2)$; *(ii) for any square matrix A, $|\mathrm{tr}(A^2)| \le \mathrm{tr}(AA^{\mathrm{T}})$; and (iii) for any square matrix A, $\|A^2\|_F^2 \le \|A^{\mathrm{T}}A\|_F^2$ where $\|B\|_F^2 = \mathrm{tr}(B^{\mathrm{T}}B)$ is the Frobenius norm of B.*

Proof. (i) Let $A = (a_{ij})$ and $B = (b_{ij})$. By the Cauchy-swartz inequality,

$$\mathrm{tr}(AB) = \sum_i \sum_j a_{ij} b_{ij} \le \sum_i \Big(\sum_j a_{ij}^2\Big)^{1/2} \Big(\sum_j b_{ij}^2\Big)^{1/2} \le \Big(\sum_i \sum_j a_{ij}^2\Big)^{1/2} \Big(\sum_i \sum_j b_{ij}^2\Big)^{1/2}.$$

Since A and B are symmetric, the right-hand side of the above inequality is the square root of $\mathrm{tr}(A^2)\mathrm{tr}(B^2)$.

(ii) Assume that $A = (a_{ij})$ is any $p \times p$ matrix. If $\mathrm{tr}(A^2) \ge 0$, because $\mathrm{tr}\{(A^{\mathrm{T}} - A)(A^{\mathrm{T}} - A)^{\mathrm{T}}\} \ge 0$ and $\mathrm{tr}\{(A^{\mathrm{T}} - A)(A^{\mathrm{T}} - A)^{\mathrm{T}}\} = 2\mathrm{tr}(A^{\mathrm{T}}A) - 2\mathrm{tr}(A^2)$, we have $|\mathrm{tr}(A^2)| \le \mathrm{tr}(AA^{\mathrm{T}})$. If $\mathrm{tr}(A^2) < 0$, because $\mathrm{tr}\{(A^{\mathrm{T}} + A)(A^{\mathrm{T}} + A)^{\mathrm{T}}\} \ge 0$ and $\mathrm{tr}\{(A^{\mathrm{T}} + A)(A^{\mathrm{T}} + A)^{\mathrm{T}}\} = 2\mathrm{tr}(A^{\mathrm{T}}A) + 2\mathrm{tr}(A^2) = 2\mathrm{tr}(A^{\mathrm{T}}A) - 2|\mathrm{tr}(A^2)|$, we have $|\mathrm{tr}(A^2)| \le \mathrm{tr}(AA^{\mathrm{T}})$.

(iii) By definition, $\|A^2\|_F^2 = \mathrm{tr}(A^{\mathrm{T}}A^{\mathrm{T}}AA) = \mathrm{tr}(A^{\mathrm{T}}AAA^{\mathrm{T}})$. Since $A^{\mathrm{T}}A$ and AA^{T} are symmetric matrices, it follows by using part (i) that

$$\mathrm{tr}(A^{\mathrm{T}}AAA^{\mathrm{T}}) \le \|A^{\mathrm{T}}A\|_F \|AA^{\mathrm{T}}\|_F = \|A^{\mathrm{T}}A\|_F^2,$$

and this completes the proof. □

LEMMA 2. *Define $U_{s_1 s_2,0} = \{n(n-1)\}^{-1} \sum_{i \neq j=1}^n (Y_{is_1}^{\mathrm{T}} Y_{js_2})^2$ for any $s_1, s_2 \in \{1, \ldots, T\}$. Under Condition 1, the leading order term of the covariance between $U_{s_1 s_2,0}$ and $U_{h_1 h_2,0}$ is*

© 2016 Biometrika Trust

$G_n(s_1,s_2,h_1,h_2)=\mathrm{cov}(U_{s_1s_2,0},U_{h_1h_2,0})$, *where*

$$G_n(s_1,s_2,h_1,h_2)=\frac{2}{n(n-1)}\mathrm{tr}^2(C_{s_1h_1}C^{\mathrm{T}}_{s_2h_2})+\frac{2}{n(n-1)}\mathrm{tr}^2(C_{s_1h_2}C^{\mathrm{T}}_{s_2h_1})$$
$$+\frac{2(n-2)}{n(n-1)}\sum_{u,v\in\{1,2\}}\mathrm{tr}(\Sigma_{s_{u^c}}C_{s_uh_v}\Sigma_{h_{v^c}}C^{\mathrm{T}}_{s_uh_v}).$$

Denote u^c as the complement set of $\{u\}$. That is, $u^c=\{1,2\}/\{u\}$.

Proof. Using the notation $\tilde{\sum}$ defined in §3 of the main paper, we define

$$L_1=\frac{1}{(P_n^2)^2}\sum_{i\neq j=1}^{n}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{ih_1}Y_{jh_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big],$$
$$L_2=\frac{1}{(P_n^2)^2}\sum_{i\neq j=1}^{n}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{jh_1}Y_{ih_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big],$$
$$L_3=\frac{1}{(P_n^2)^2}\tilde{\sum}_{i,j,l}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{ih_1}Y_{lh_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big],$$
$$L_4=\frac{1}{(P_n^2)^2}\tilde{\sum}_{i,j,l}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{lh_1}Y_{ih_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big],$$
$$L_5=\frac{1}{(P_n^2)^2}\tilde{\sum}_{i,j,l}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{jh_1}Y_{lh_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big],$$
$$L_6=\frac{1}{(P_n^2)^2}\tilde{\sum}_{i,j,l}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{lh_1}Y_{jh_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big],$$
$$L_7=\frac{1}{(P_n^2)^2}\tilde{\sum}_{i,j,k,l}E\big[\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2-\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})\}\{(Y^{\mathrm{T}}_{kh_1}Y_{lh_2})^2-\mathrm{tr}(\Sigma_{h_1}\Sigma_{h_2})\}\big].$$

Then $\mathrm{cov}(U_{s_1s_2,0},U_{h_1h_2,0})=L_1+\cdots+L_7$ since $E\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2\}=\mathrm{tr}(\Sigma_{s_1}\Sigma_{s_2})$. Applying standard results in multivariate analysis, we obtain

$$E\{(Y^{\mathrm{T}}_{is_1}Y_{js_2})^2(Y^{\mathrm{T}}_{ih_1}Y_{jh_2})^2\}=2\mathrm{tr}(C_{h_1s_1}C_{s_2h_2}C_{h_1s_1}C_{s_2h_2})+2\mathrm{tr}(\Sigma_{s_1}C^{\mathrm{T}}_{s_2h_2}\Sigma_{h_1}C_{s_2h_2})$$
$$+2\mathrm{tr}^2(C_{s_2h_2}C_{h_1s_1})+2\mathrm{tr}(\Sigma_{s_2}C_{s_1h_1}\Sigma_{h_2}C^{\mathrm{T}}_{s_1h_1})+\mathrm{tr}(\Sigma_{s_2}\Sigma_{s_1})\mathrm{tr}(\Sigma_{h_2}\Sigma_{h_1}).$$

This implies that

$$L_1+L_2=\frac{2}{n(n-1)}\Big[\mathrm{tr}\{(C_{s_2h_2}C_{h_1s_1})^2\}+\mathrm{tr}\{(C_{s_2h_1}C_{h_2s_1})^2\}+\mathrm{tr}^2(C_{s_2h_2}C_{h_1s_1})$$
$$+\mathrm{tr}^2(C_{s_2h_1}C_{h_2s_1})+\mathrm{tr}(\Sigma_{s_1}C_{s_2h_2}\Sigma_{h_1}C^{\mathrm{T}}_{s_2h_2})+\mathrm{tr}(\Sigma_{s_2}C_{s_1h_1}\Sigma_{h_2}C^{\mathrm{T}}_{s_1h_1})$$
$$+\mathrm{tr}(\Sigma_{s_1}C_{s_2h_1}\Sigma_{h_2}C^{\mathrm{T}}_{s_2h_1})+\mathrm{tr}(\Sigma_{s_2}C_{s_1h_2}\Sigma_{h_1}C^{\mathrm{T}}_{s_1h_2})\Big].$$

Furthermore, $L_7=0$ and

$$\sum_{i=3}^{6}L_i=\frac{2(n-2)}{n(n-1)}\sum_{u,v\in\{1,2\}}\mathrm{tr}(\Sigma_{s_{u^c}}C_{s_uh_v}\Sigma_{h_{v^c}}C^{\mathrm{T}}_{s_uh_v}),$$

This with Condition 1 implies that Lemma 2 is valid. □

LEMMA 3. *Define* $U_{s_1s_2,1} = (1/P_n^3)\,\tilde{\sum}_{i,j,k}\, Y_{is_1}^{\mathrm{T}} Y_{js_2} Y_{js_2}^{\mathrm{T}} Y_{ks_1}$. *The leading term in the covariance between* $U_{s_1s_2,1}$ *and* $U_{h_1h_2,1}$ *is*

$$
\begin{aligned}
\operatorname{cov}(U_{s_1s_2,1}, U_{h_1h_2,1}) = {} & \frac{4}{n^3} \sum_{u,v\in\{1,2\}} \operatorname{tr}^2(C_{s_uh_v} C_{s_{u^c}h_{v^c}}^{\mathrm{T}}) \\
& + \frac{2}{n^2} \sum_{u,v\in\{1,2\}} \operatorname{tr}(\Sigma_{s_{u^c}} C_{s_uh_v} \Sigma_{h_{v^c}} C_{s_uh_v}^{\mathrm{T}}),
\end{aligned}
$$

where u^c *is the complement set of* $\{u\}$. *That is,* $u^c = \{1,2\}/\{u\}$. *In addition,* $\operatorname{var}(\hat{D}_{nt,1}) = o\{\operatorname{var}(\hat{D}_{nt,0})\}$.

Proof. Because $E(U_{s_1s_2,1}) = 0$, $\operatorname{cov}(U_{s_1s_2,1}, U_{h_1h_2,1}) = E(U_{s_1s_2,1}U_{h_1h_2,1})$. By definition,

$$
\begin{aligned}
U_{s_1s_2,1}U_{h_1h_2,1} = \frac{1}{(P_n^3)^2} \tilde{\sum}_{i,j,k} \tilde{\sum}_{i_1,j_1,k_1} & (Y_{is_1}^{\mathrm{T}} Y_{js_2} Y_{js_2}^{\mathrm{T}} Y_{ks_1} + Y_{is_2}^{\mathrm{T}} Y_{js_1} Y_{js_1}^{\mathrm{T}} Y_{ks_2}) \\
& \times (Y_{i_1h_1}^{\mathrm{T}} Y_{j_1h_2} Y_{j_1h_2}^{\mathrm{T}} Y_{k_1h_1} + Y_{i_1h_2}^{\mathrm{T}} Y_{j_1h_1} Y_{j_1h_1}^{\mathrm{T}} Y_{k_1h_2}).
\end{aligned}
$$

According to the number of equivalent indices among two sets $\{i,j,k\}$ and $\{i_1,j_1,k_1\}$, we decompose $U_{s_1s_2,1}U_{h_1h_2,1}$ into three terms. Let $I_c = \{i,j,k\} \cup \{i_1,j_1,k_1\}$ where c represents the number of indices that are equivalent to each other in two sets $\{i,j,k\}$ and $\{i_1,j_1,k_1\}$. If there is one index equivalent,

$$
\begin{aligned}
I_1 = \{ & (i=i_1, j, k, j_1, k_1), (i=j_1, j, k, i_1, k_1), (i, j, k=i_1, i_1, j_1), \\
& (i, j=i_1, k, j_1, k_1), (i, j=j_1, k, i_1, k_1), (i, j=k_1, k, i_1, j_1), \\
& (i, j, k=i_1, j_1, k_1), (i, j, k=j_1, i_1, k_1), (i, j, k=k_1, i_1, j_1)\}.
\end{aligned}
$$

For each case within I_1, the expectation of corresponding summand in $U_{s_1s_2,1}U_{h_1h_2,1}$ is 0. If there are two indices equivalent,

$$
\begin{aligned}
I_2 = \{ & (i=i_1, j=j_1, k, k_1), (i=j_1, j=i_1, k, k_1), (i=i_1, k=k_1, j, j_1), \\
& (i=k_1, k=i_1, j, j_1), (j=j_1, k=k_1, i, i_1), (j=k_1, k=j_1, i, i_1), \\
& (i=i_1, j=k_1, k, j_1), (i=j_1, j=k_1, k, i_1), (i=i_1, k=j_1, j, k_1), \\
& (i=k_1, k=j_1, j, i_1), (j=j_1, k=i_1, i, k_1), (j=k_1, k=i_1, i, j_1)\}.
\end{aligned}
$$

Among all the cases in I_2, there exist two cases $\{(i=i_1, k=k_1, j, j_1), (i=k_1, k=i_1, j, j_1)\}$ whose expectations of the summand in $U_{s_1s_2,1}U_{h_1h_2,1}$ are not zero. Similarly, if there are three indices equivalent,

$$
\begin{aligned}
I_3 = \{ & (i=i_1, j=j_1, k=k_1), (i=j_1, j=i_1, k=k_1), (i=k_1, j=j_1, k=i_1), \\
& (i=k_1, j=i_1, k=j_1), (i=i_1, j=k_1, k=j_1), (i=j_1, j=k_1, k=i_1)\}.
\end{aligned}
$$

Among all the cases in I_3, there are two cases $(i=i_1, j=j_1, k=k_1)$ and $(i=k_1, j=j_1, k=i_1)$ that have non-zero expectation.

In summary,

$$
\begin{aligned}
E(U_{s_1s_2,1}U_{h_1h_2,1}) &= \frac{2}{(P_n^3)^2}E\Big\{ \widetilde{\sum}_{i,k,j,j_1} (Y_{is_1}^{\mathrm{T}}Y_{js_2}Y_{js_2}^{\mathrm{T}}Y_{ks_1} + Y_{is_2}^{\mathrm{T}}Y_{js_1}Y_{js_1}^{\mathrm{T}}Y_{ks_2}) \\
&\qquad\qquad \times (Y_{ih_1}^{\mathrm{T}}Y_{j_1h_2}Y_{j_1h_2}^{\mathrm{T}}Y_{kh_1} + Y_{ih_2}^{\mathrm{T}}Y_{j_1h_1}Y_{j_1h_1}^{\mathrm{T}}Y_{kh_2})\Big\} \\
&\quad + \frac{2}{(P_n^3)^2}E\Big\{ \widetilde{\sum}_{i,k,j} (Y_{is_1}^{\mathrm{T}}Y_{js_2}Y_{js_2}^{\mathrm{T}}Y_{ks_1} + Y_{is_2}^{\mathrm{T}}Y_{js_1}Y_{js_1}^{\mathrm{T}}Y_{ks_2}) \\
&\qquad\qquad \times (Y_{ih_1}^{\mathrm{T}}Y_{jh_2}Y_{jh_2}^{\mathrm{T}}Y_{kh_1} + Y_{ih_2}^{\mathrm{T}}Y_{jh_1}Y_{jh_1}^{\mathrm{T}}Y_{kh_2})\Big\} \\
&= \frac{2}{P_n^3}\sum_{u,v\in\{1,2\}}\Big[(n-3)\mathrm{tr}(\Sigma_{s_{u^c}}C_{s_uh_v}\Sigma_{h_{v^c}}C_{s_uh_v}^{\mathrm{T}}) + \mathrm{tr}^2(C_{s_uh_v}C_{s_{u^c}h_{v^c}}^{\mathrm{T}}) \\
&\quad + \mathrm{tr}\{(C_{s_uh_v}C_{s_{u^c}h_{v^c}}^{\mathrm{T}})^2\} + \mathrm{tr}(\Sigma_{s_{u^c}}C_{s_uh_v}\Sigma_{h_{v^c}}C_{s_uh_v}^{\mathrm{T}})\Big].
\end{aligned}
$$

This completes the proof. □

LEMMA 4. *Define* $U_{s_1s_2,2} = (1/P_n^4)\widetilde{\sum}_{i,j,k,l} (Y_{is_1}^{\mathrm{T}}Y_{js_2})(Y_{ks_1}^{\mathrm{T}}Y_{ls_2})$. *For any fixed* $u,v,k,l \in \{1,2\}$*, the covariance between* $U_{s_us_v,2}$ *and* $U_{h_kh_l,2}$ *is*

$$
\begin{aligned}
\mathrm{cov}(U_{s_us_v,2},U_{h_kh_l,2}) &= \frac{2}{P_n^4}\{\mathrm{tr}^2(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) + \mathrm{tr}(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) \\
&\quad + 2\mathrm{tr}(C_{s_vh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_l}C_{s_uh_k}^{\mathrm{T}}) + 2\mathrm{tr}(C_{s_vh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_k}C_{s_uh_l}^{\mathrm{T}}) \\
&\quad + 3\mathrm{tr}(C_{s_vh_k}C_{s_uh_l}^{\mathrm{T}}C_{s_vh_l}C_{s_uh_k}^{\mathrm{T}}) + 3\mathrm{tr}(C_{s_vh_k}C_{s_uh_l}^{\mathrm{T}})\mathrm{tr}(C_{s_vh_l}C_{s_uh_k}^{\mathrm{T}})\}.
\end{aligned}
$$

Moreover, $\mathrm{var}(\hat{D}_{nt,2}) = o\{\mathrm{var}(\hat{D}_{nt,0})\}$.

Proof. Because $E(U_{s_us_v,2}) = 0$, $\mathrm{cov}(U_{s_us_v,2},U_{h_kh_l,2}) = E(U_{s_us_v,2}U_{h_kh_l,2})$. Therefore,

$$
\begin{aligned}
\mathrm{cov}(U_{s_us_v,2},U_{h_kh_l,2}) &= \frac{1}{(P_n^4)^2}\widetilde{\sum}_{i,j,k,l}\widetilde{\sum}_{i_1,j_1,k_1,l_1} E\{(Y_{is_u}^{\mathrm{T}}Y_{js_v})(Y_{ks_u}^{\mathrm{T}}Y_{ls_v})(Y_{i_1h_k}^{\mathrm{T}}Y_{j_1h_l})(Y_{k_1h_k}^{\mathrm{T}}Y_{l_1h_l})\} \\
&= \frac{2}{P_n^4}\{\mathrm{tr}^2(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) + \mathrm{tr}(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) \\
&\quad + 2\mathrm{tr}(C_{s_vh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_l}C_{s_uh_k}^{\mathrm{T}}) + 2\mathrm{tr}(C_{s_vh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_k}C_{s_uh_l}^{\mathrm{T}}) \\
&\quad + 3\mathrm{tr}(C_{s_vh_k}C_{s_uh_l}^{\mathrm{T}}C_{s_vh_l}C_{s_uh_k}^{\mathrm{T}}) + 3\mathrm{tr}(C_{s_vh_k}C_{s_uh_l}^{\mathrm{T}})\mathrm{tr}(C_{s_vh_l}C_{s_uh_k}^{\mathrm{T}})\}.
\end{aligned}
$$

This completes the proof of the first part.

For the second part, we write $\hat{D}_{nt,2} = w^{-1}(t)\sum_{s_1=1}^{t}\sum_{s_2=t+1}^{\mathrm{T}}\sum_{u,v\in\{1,2\}}(-1)^{|u-v|}U_{s_us_v,2}$. It follows by the first part that

$$
\begin{aligned}
\mathrm{var}(\hat{D}_{nt,2}) &= \frac{1}{w^2(t)}\frac{2}{n^4}\sum_{s_1,s_2,h_1,h_2}^{*}\sum_{u,v,k,l\in\{1,2\}}(-1)^{|u-v|+|k-l|}\{\mathrm{tr}^2(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) \\
&\quad + \mathrm{tr}(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) + 2\mathrm{tr}(C_{s_vh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_l}C_{s_uh_k}^{\mathrm{T}}) \\
&\quad + 2\mathrm{tr}(C_{s_vh_k}C_{s_vh_l}^{\mathrm{T}}C_{s_uh_k}C_{s_uh_l}^{\mathrm{T}}) + 3\mathrm{tr}(C_{s_vh_k}C_{s_uh_l}^{\mathrm{T}}C_{s_vh_l}C_{s_uh_k}^{\mathrm{T}}) \\
&\quad + 3\mathrm{tr}(C_{s_vh_k}C_{s_uh_l}^{\mathrm{T}})\mathrm{tr}(C_{s_vh_l}C_{s_uh_k}^{\mathrm{T}})\}.
\end{aligned}
$$

Applying the inequalities given in Lemma 1, we can show that $\mathrm{var}(\hat{D}_{nt,2}) = o\{\mathrm{var}(\hat{D}_{nt,0})\}$. This completes the proof of this Lemma. □

LEMMA 5. *Let Z be an m-dimensional multivariate normally distributed random vector with mean 0 and covariance I_m. Define $M = ZZ^{\mathrm{T}} - I$. Assume A, B, C, D are matrices with appropriate dimensions. Then $E\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\} = \mathrm{tr}^2(A^{\mathrm{T}}B) + \mathrm{tr}\{(A^{\mathrm{T}}B)^2\}$ and*

$$\begin{aligned}
&\mathrm{cov}\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}}), \mathrm{tr}(CMC^{\mathrm{T}}DMD^{\mathrm{T}})\} \\
&= 2\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}(C^{\mathrm{T}}D)\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(C^{\mathrm{T}}D + D^{\mathrm{T}}C)\} \\
&\quad + \frac{1}{2}\mathrm{tr}^2\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(C^{\mathrm{T}}D + D^{\mathrm{T}}C)\} + \mathrm{tr}\big[\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(C^{\mathrm{T}}D + D^{\mathrm{T}}C)\}^2\big] \\
&\quad + 2\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(C^{\mathrm{T}}DC^{\mathrm{T}}D + D^{\mathrm{T}}CD^{\mathrm{T}}C)\} \\
&\quad + 2\mathrm{tr}(C^{\mathrm{T}}D)\mathrm{tr}\{(C^{\mathrm{T}}D + D^{\mathrm{T}}C)(A^{\mathrm{T}}BA^{\mathrm{T}}B + B^{\mathrm{T}}AB^{\mathrm{T}}A)\} \\
&\quad + 2\mathrm{tr}\{(A^{\mathrm{T}}BA^{\mathrm{T}}B + B^{\mathrm{T}}AB^{\mathrm{T}}A)(C^{\mathrm{T}}DC^{\mathrm{T}}D + D^{\mathrm{T}}CD^{\mathrm{T}}C)\}.
\end{aligned}$$

In particular,

$$\begin{aligned}
&\mathrm{var}\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\} = 2\mathrm{tr}^2(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)^2\} + \frac{1}{2}\mathrm{tr}^2\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)^2\} \\
&\quad + 4\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(A^{\mathrm{T}}BA^{\mathrm{T}}B + B^{\mathrm{T}}AB^{\mathrm{T}}A)\} \\
&\quad + 2\mathrm{tr}\{(A^{\mathrm{T}}BA^{\mathrm{T}}B + B^{\mathrm{T}}AB^{\mathrm{T}}A)^2\} + \mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)^4\}.
\end{aligned}$$

Moreover, $\mathrm{var}\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\} \le K\Big[\mathrm{tr}^4(A^{\mathrm{T}}B) + \mathrm{tr}^2\{(A^{\mathrm{T}}B)^{\otimes 2}\}\Big]$ *for a constant $K > 0$.*

Proof. We first consider $E\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\}$. Because $M = ZZ^{\mathrm{T}} - I$, we have

$$\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}}) = (Z^{\mathrm{T}}A^{\mathrm{T}}BZ)^2 - Z^{\mathrm{T}}A^{\mathrm{T}}BB^{\mathrm{T}}AZ - Z^{\mathrm{T}}B^{\mathrm{T}}AA^{\mathrm{T}}BZ + \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A). \tag{S.1}$$

Taking expectation of the both sides of equation (S.1), we have

$$\begin{aligned}
E\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\} &= \mathrm{tr}^2(A^{\mathrm{T}}B) + \mathrm{tr}\{(A^{\mathrm{T}}B)^2\} + \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A) - \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A) \\
&= \mathrm{tr}^2(A^{\mathrm{T}}B) + \mathrm{tr}\{(A^{\mathrm{T}}B)^2\}.
\end{aligned}$$

Next, we consider the covariance part. Using equation (S.1), we have

$$\begin{aligned}
&\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\mathrm{tr}(CMC^{\mathrm{T}}DMD^{\mathrm{T}}) = (Z^{\mathrm{T}}A^{\mathrm{T}}BZ)^2(Z^{\mathrm{T}}C^{\mathrm{T}}DZ)^2 \\
&- (Z^{\mathrm{T}}A^{\mathrm{T}}BZ)^2Z^{\mathrm{T}}C^{\mathrm{T}}DD^{\mathrm{T}}CZ - (Z^{\mathrm{T}}A^{\mathrm{T}}BZ)^2Z^{\mathrm{T}}D^{\mathrm{T}}CC^{\mathrm{T}}DZ \\
&- (Z^{\mathrm{T}}A^{\mathrm{T}}BZ)^2\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C) - Z^{\mathrm{T}}B^{\mathrm{T}}AA^{\mathrm{T}}BZ(Z^{\mathrm{T}}C^{\mathrm{T}}DZ)^2 \\
&+ (Z^{\mathrm{T}}B^{\mathrm{T}}AA^{\mathrm{T}}BZ)(Z^{\mathrm{T}}C^{\mathrm{T}}DD^{\mathrm{T}}CZ) + (Z^{\mathrm{T}}B^{\mathrm{T}}AA^{\mathrm{T}}BZ)(Z^{\mathrm{T}}D^{\mathrm{T}}CC^{\mathrm{T}}DZ) \\
&- (Z^{\mathrm{T}}B^{\mathrm{T}}AA^{\mathrm{T}}BZ)\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C) - Z^{\mathrm{T}}A^{\mathrm{T}}BB^{\mathrm{T}}AZ(Z^{\mathrm{T}}C^{\mathrm{T}}DZ)^2 \\
&+ Z^{\mathrm{T}}A^{\mathrm{T}}BB^{\mathrm{T}}AZ(Z^{\mathrm{T}}C^{\mathrm{T}}DD^{\mathrm{T}}CZ) + Z^{\mathrm{T}}A^{\mathrm{T}}BB^{\mathrm{T}}AZ(Z^{\mathrm{T}}D^{\mathrm{T}}CC^{\mathrm{T}}DZ) \\
&- Z^{\mathrm{T}}A^{\mathrm{T}}BB^{\mathrm{T}}AZ\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C) + \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)(Z^{\mathrm{T}}C^{\mathrm{T}}DZ)^2 \\
&- \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)(Z^{\mathrm{T}}C^{\mathrm{T}}DD^{\mathrm{T}}CZ) - \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)(Z^{\mathrm{T}}D^{\mathrm{T}}CC^{\mathrm{T}}DZ) \\
&- \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C).
\end{aligned}$$

Define the terms in the above expression as $J_1, \ldots, J_{16}$. We consider the expectation of each J_i for $i = 1, \ldots, 16$. We have the following:

$$
\begin{aligned}
E(J_4) &= [\mathrm{tr}^2(A^{\mathrm{T}}B) + \mathrm{tr}\{(A^{\mathrm{T}}B)^2\} + \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)]\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C),\\
E(J_6) &= \mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}B)\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C) + 2\mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}BC^{\mathrm{T}}DD^{\mathrm{T}}C),\\
E(J_7) &= \mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}B)\mathrm{tr}(D^{\mathrm{T}}CC^{\mathrm{T}}D) + 2\mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}BD^{\mathrm{T}}CC^{\mathrm{T}}D),\\
E(J_8) &= E(J_{12}) = E(J_{14}) = -\mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}B)\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C),\\
E(J_{10}) &= \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C) + 2\mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}AC^{\mathrm{T}}DD^{\mathrm{T}}C),\\
E(J_{11}) &= \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)\mathrm{tr}(D^{\mathrm{T}}CC^{\mathrm{T}}D) + 2\mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}AD^{\mathrm{T}}CC^{\mathrm{T}}D),\\
E(J_{13}) &= \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)[\mathrm{tr}^2(C^{\mathrm{T}}D) + \mathrm{tr}\{(C^{\mathrm{T}}D)^2\} + \mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C)].
\end{aligned}
$$

In addition, we can show that, for any matrices A, B, C of appropriate dimensions,

$$
\begin{aligned}
E(Z^{\mathrm{T}}AZZ^{\mathrm{T}}BZZ^{\mathrm{T}}CZ) = {}& \mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}(C) + \mathrm{tr}(A)\{\mathrm{tr}(BC) + \mathrm{tr}(B^{\mathrm{T}}C)\}\\
&+ \mathrm{tr}(B)\{\mathrm{tr}(AC) + \mathrm{tr}(A^{\mathrm{T}}C)\} + \mathrm{tr}(C)\{\mathrm{tr}(AB) + \mathrm{tr}(A^{\mathrm{T}}B)\}\\
&+ \mathrm{tr}\{(A + A^{\mathrm{T}})(B + B^{\mathrm{T}})(C + C^{\mathrm{T}})\}.
\end{aligned}
$$

Applying the above formula to J_2, J_3, J_5 and J_9, we obtain

$$
\begin{aligned}
-E(J_2) = {}& \mathrm{tr}^2(A^{\mathrm{T}}B)\mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C) + \mathrm{tr}(A^{\mathrm{T}}B)\{\mathrm{tr}(A^{\mathrm{T}}BC^{\mathrm{T}}DD^{\mathrm{T}}C) + \mathrm{tr}(B^{\mathrm{T}}AC^{\mathrm{T}}DD^{\mathrm{T}}C)\}\\
&+ \mathrm{tr}(A^{\mathrm{T}}B)\{\mathrm{tr}(A^{\mathrm{T}}BC^{\mathrm{T}}DD^{\mathrm{T}}C) + \mathrm{tr}(B^{\mathrm{T}}A^{\mathrm{T}}C^{\mathrm{T}}DD^{\mathrm{T}}C)\}\\
&+ \mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C)\{\mathrm{tr}(A^{\mathrm{T}}BA^{\mathrm{T}}B) + \mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}B)\}\\
&+ 2\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)^2C^{\mathrm{T}}DD^{\mathrm{T}}C\}.
\end{aligned}
$$

The expectation of J_3 is the same as $E(J_2)$ above except for changing $C^{\mathrm{T}}DD^{\mathrm{T}}C$ to $D^{\mathrm{T}}CC^{\mathrm{T}}D$. Similarly,

$$
\begin{aligned}
-E(J_5) = {}& \mathrm{tr}^2(C^{\mathrm{T}}D)\mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}B) + \mathrm{tr}(C^{\mathrm{T}}D)\{\mathrm{tr}(C^{\mathrm{T}}DB^{\mathrm{T}}AA^{\mathrm{T}}B) + \mathrm{tr}(D^{\mathrm{T}}CB^{\mathrm{T}}AA^{\mathrm{T}}B)\}\\
&+ \mathrm{tr}(C^{\mathrm{T}}D)\{\mathrm{tr}(C^{\mathrm{T}}DB^{\mathrm{T}}AA^{\mathrm{T}}B) + \mathrm{tr}(D^{\mathrm{T}}C^{\mathrm{T}}B^{\mathrm{T}}AA^{\mathrm{T}}B)\}\\
&+ \mathrm{tr}(B^{\mathrm{T}}AA^{\mathrm{T}}B)\{\mathrm{tr}(C^{\mathrm{T}}DC^{\mathrm{T}}D) + \mathrm{tr}(D^{\mathrm{T}}CC^{\mathrm{T}}D)\}\\
&+ 2\mathrm{tr}\{(C^{\mathrm{T}}D + D^{\mathrm{T}}C)^2B^{\mathrm{T}}AA^{\mathrm{T}}B\},
\end{aligned}
$$

and $E(J_9)$ is the same as $E(J_5)$ with replacing $B^{\mathrm{T}}AA^{\mathrm{T}}B$ with $A^{\mathrm{T}}BB^{\mathrm{T}}A$. Finally, we can show that

$$
\begin{aligned}
E(J_1) = {}& \mathrm{tr}^2(A^{\mathrm{T}}B)\mathrm{tr}^2(C^{\mathrm{T}}D) + \mathrm{tr}^2(A^{\mathrm{T}}B)[\mathrm{tr}\{(C^{\mathrm{T}}D)^2\} + \mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C)]\\
&+ \mathrm{tr}^2(C^{\mathrm{T}}D)[\mathrm{tr}\{(A^{\mathrm{T}}B)^2\} + \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)]\\
&+ 4\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}(C^{\mathrm{T}}D)\{\mathrm{tr}(A^{\mathrm{T}}BC^{\mathrm{T}}D) + \mathrm{tr}(B^{\mathrm{T}}AC^{\mathrm{T}}D)\}\\
&+ [\mathrm{tr}\{(A^{\mathrm{T}}B)^2\} + \mathrm{tr}(A^{\mathrm{T}}BB^{\mathrm{T}}A)][\mathrm{tr}\{(C^{\mathrm{T}}D)^2\} + \mathrm{tr}(C^{\mathrm{T}}DD^{\mathrm{T}}C)]\\
&+ 2\{\mathrm{tr}(A^{\mathrm{T}}BC^{\mathrm{T}}D) + \mathrm{tr}(B^{\mathrm{T}}AC^{\mathrm{T}}D)\}^2\\
&+ 2\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(CD + D^{\mathrm{T}}C^{\mathrm{T}})^2\}\\
&+ 2\mathrm{tr}(C^{\mathrm{T}}D)\mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)^2(CD + D^{\mathrm{T}}C^{\mathrm{T}})\}\\
&+ \mathrm{tr}\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)^2(CD + D^{\mathrm{T}}C^{\mathrm{T}})^2\} + \mathrm{tr}[\{(A^{\mathrm{T}}B + B^{\mathrm{T}}A)(CD + D^{\mathrm{T}}C^{\mathrm{T}})\}^2].
\end{aligned}
$$

Summarizing the above $E(J_i)$'s, we obtain $E\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\mathrm{tr}(CMC^{\mathrm{T}}DMD^{\mathrm{T}})\}$. From this result and (S.1), we can obtain the covariance between $\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})$ and

$\mathrm{tr}(CMC^{\mathrm{T}}DMD^{\mathrm{T}})$. The variance is a special case of the covariance. This completes the first part of the Lemma.

Next, we prove the inequality given in the second part. Using the Cauchy-Schwarz inequality and Lemma 1,

$$2\mathrm{tr}^2(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B+B^{\mathrm{T}}A)^2\} \le K\mathrm{tr}^2(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B)^{\otimes 2}\}$$

and

$$\begin{aligned}
&2\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}\{(A^{\mathrm{T}}B+B^{\mathrm{T}}A)(A^{\mathrm{T}}BA^{\mathrm{T}}B+B^{\mathrm{T}}AB^{\mathrm{T}}A)\}\\
&\le 2\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}^{1/2}\{(A^{\mathrm{T}}B+B^{\mathrm{T}}A)^2\}\mathrm{tr}^{1/2}\{(A^{\mathrm{T}}BA^{\mathrm{T}}B+B^{\mathrm{T}}AB^{\mathrm{T}}A)^2\}\\
&\le K\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}^{1/2}\{(A^{\mathrm{T}}B)^{\otimes 2}\}\mathrm{tr}^{1/2}\{(A^{\mathrm{T}}BA^{\mathrm{T}}B)^{\otimes 2}\}\\
&\le K\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}^{1/2}\{(A^{\mathrm{T}}B)^{\otimes 2}\}\mathrm{tr}\big[\{(A^{\mathrm{T}}B)^{\mathrm{T}}(A^{\mathrm{T}}B)\}^2\big]\\
&\le K\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}^{3/2}\{(A^{\mathrm{T}}B)^{\otimes 2}\}.
\end{aligned}$$

Moreover, $\mathrm{tr}\{(A^{\mathrm{T}}B+B^{\mathrm{T}}A)^4\} \le \mathrm{tr}^2\{(A^{\mathrm{T}}B+B^{\mathrm{T}}A)^2\} \le K\mathrm{tr}^2\{(A^{\mathrm{T}}B)^{\otimes 2}\}$. In summary,

$$\begin{aligned}
\mathrm{var}\{\mathrm{tr}(AMA^{\mathrm{T}}BMB^{\mathrm{T}})\} &\le K\Big[\mathrm{tr}(A^{\mathrm{T}}B)\mathrm{tr}^{1/2}\{(A^{\mathrm{T}}B)^{\otimes 2}\}+\mathrm{tr}\{(A^{\mathrm{T}}B)^{\otimes 2}\}\Big]^2\\
&\le K\Big[\mathrm{tr}^4(A^{\mathrm{T}}B)+\mathrm{tr}^2\{(A^{\mathrm{T}}B)^{\otimes 2}\}\Big].
\end{aligned}$$

This finishes the proof of this Lemma. □

Define

$$V_{n0}(s_1,s_2,h_1,h_2)=\frac{4}{n(n-1)}\sum_{u,v,k,l\in\{1,2\}}(-1)^{-|u-v|-|k-l|}\mathrm{tr}^2(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}),$$

$$V_{n1}(s_1,s_2,h_1,h_2)=\frac{8(n-2)}{n(n-1)}\sum_{u,v\in\{1,2\}}(-1)^{|u-v|}\mathrm{tr}\{(\Sigma_{s_1}-\Sigma_{s_2})C_{s_uh_v}(\Sigma_{h_1}-\Sigma_{h_2})C_{s_uh_v}^{\mathrm{T}}\}.$$

LEMMA 6. *Let* $W_{s_1s_2}=U_{s_1s_1,0}+U_{s_2s_2,0}-U_{s_1s_2,0}-U_{s_2s_1,0}$. *The covariance between* $W_{s_1s_2}$ *and* $W_{h_1h_2}$ *is* $V_u(s_1,s_2,h_1,h_2)$, *where* $V_u(s_1,s_2,h_1,h_2)=V_{n0}(s_1,s_2,h_1,h_2)+V_{n1}(s_1,s_2,h_1,h_2)$ *and* $V_{n0}(s_1,s_2,h_1,h_2)$ *is the covariance between* $W_{s_1s_2}$ *and* $W_{h_1h_2}$ *under* H_0.

Proof. Let $G_n(\cdot)$ be the function defined in Lemma 2. It then follows,

$$V_u(s_1,s_2,h_1,h_2)=\sum_{u,v,k,l\in\{1,2\}}(-1)^{-|u-v|-|k-l|}G_n(s_u,s_v,h_k,h_l).$$

Applying Lemma 2, we have

$$\begin{aligned}
V_u(s_1,s_2,h_1,h_2)&=\frac{2}{n(n-1)}\sum_{u,v,k,l\in\{1,2\}}(-1)^{-|u-v|-|k-l|}\big\{\mathrm{tr}^2(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}})+\mathrm{tr}^2(C_{s_uh_l}C_{s_vh_k}^{\mathrm{T}})\big\}\\
&+\frac{2(n-2)}{n(n-1)}\sum_{u,v,k,l\in\{1,2\}}(-1)^{-|u-v|-|k-l|}\big\{\mathrm{tr}(\Sigma_{s_u}C_{s_vh_l}\Sigma_{h_k}C_{s_vh_l}^{\mathrm{T}})+\mathrm{tr}(\Sigma_{s_v}C_{s_uh_k}\Sigma_{h_l}C_{s_uh_k}^{\mathrm{T}})\\
&+\mathrm{tr}(\Sigma_{s_u}C_{s_vh_k}\Sigma_{h_l}C_{s_vh_k}^{\mathrm{T}})+\mathrm{tr}(\Sigma_{s_v}C_{s_uh_l}\Sigma_{h_k}C_{s_uh_l}^{\mathrm{T}})\big\}.
\end{aligned}$$

Hence,

$$V_u(s_1, s_2, h_1, h_2) = \frac{4}{n(n-1)} \sum_{u,v,k,l \in \{1,2\}} (-1)^{-|u-v|-|k-l|} \mathrm{tr}^2(C_{s_u h_k} C_{s_v h_l}^{\mathrm{T}})$$
$$+ \frac{8(n-2)}{n(n-1)} \sum_{u,v,k,l \in \{1,2\}} (-1)^{-|u-v|-|k-l|} \mathrm{tr}(\Sigma_{s_u} C_{s_v h_l} \Sigma_{h_k} C_{s_v h_l}^{\mathrm{T}}).$$

After some algebra, one can show the second term in the above expression is equivalent to $V_{n1}(s, h, h_1, h_2)$.

Under H_0, $V_{n1}(s, h, h_1, h_2) = 0$. Therefore, $V_u(s, h, h_1, h_2) = V_0(s, h, h_1, h_2)$ is the covariance under H_0. This completes the proof of Lemma 6. □

2. PROOFS OF MAIN RESULTS

In this section, we present proofs for the main results of the paper. By definition, $\hat{D}_{nt}$ can be expressed as $\hat{D}_{nt} = \hat{D}_{nt,0} - 2\hat{D}_{nt,1} + \hat{D}_{nt,2}$, where for $k = 0, 1$ and 2,

$$\hat{D}_{nt,k} = \frac{1}{t(T-t)} \sum_{s_1=1}^{t} \sum_{s_2=t+1}^{T} (U_{s_1 s_1,k} + U_{s_2 s_2,k} - U_{s_1 s_2,k} - U_{s_2 s_1,k}). \tag{S.2}$$

Here $U_{s_1 s_2,k}$ was defined in § 3 of the main paper.

Proof of Theorem 1. Based on the definition of $\hat{D}_{nt}$, the expectation of $\hat{D}_{nt}$ is

$$E(\hat{D}_{nt}) = \frac{1}{t} \sum_{s=1}^{t} \mathrm{tr}(\Sigma_s^2) + \frac{1}{T-t} \sum_{h=t+1}^{T} \mathrm{tr}(\Sigma_h^2) - \frac{2}{t(T-t)} \sum_{s=1}^{t} \sum_{h=t+1}^{T} \mathrm{tr}(\Sigma_s \Sigma_h) = D_t.$$

We next calculate the order of the variance of $\hat{D}_{nt}$. By using the definition of $\hat{D}_{nt}$, write $\hat{D}_{nt}$ as $\hat{D}_{nt} = \hat{D}_{nt,0} - 2\hat{D}_{nt,1} + \hat{D}_{nt,2}$. By Lemmas 3 and 4, it follows that $\hat{D}_{nt,1} = o_p(\hat{D}_{nt,0})$ and $\hat{D}_{nt,2} = o_p(\hat{D}_{nt,0})$. Therefore, it suffices to compute the variance of $\hat{D}_{nt,0}$. Using Lemma 6,

$$\sigma_{nt}^2 = w^{-2}(t) \sum_{\substack{s_1, s_2, \\ h_1, h_2}}^{*} \mathrm{cov}(W_{s_1 s_2}, W_{h_1 h_2}) = w^{-2}(t) \sum_{\substack{s_1, s_2, \\ h_1, h_2}}^{*} V_u(s_1, s_2, h_1, h_2).$$

This completes the proof of Theorem 1. □

Proof of Theorem 2. By Theorem 1, it is sufficient to establish the asymptotic normality of $\hat{D}_{nt,0}$. We first write $\hat{D}_{nt,0}$ into a martingale. Define $A_{js_u} = Y_{js_u} Y_{js_u}^{\mathrm{T}} - \Sigma_{s_u}$, and

$$G_{nj} = \frac{1}{t(T-t)} \sum_{s_1=1}^{t} \sum_{s_2=t+1}^{T} \sum_{u,v \in \{1,2\}} (-1)^{|u-v|} \big\{ Y_{is_u}^{\mathrm{T}} A_{js_v} Y_{is_u} - \mathrm{tr}(\Sigma_{s_u} A_{js_v}) \big\},$$

$$Q_{ni} = \frac{1}{t(T-t)} \sum_{s_1=1}^{t} \sum_{s_2=t+1}^{T} \sum_{u,v \in \{1,2\}} (-1)^{|u-v|} \{ Y_{is_u}^{\mathrm{T}} \Sigma_{s_v} Y_{is_u} - \mathrm{tr}(\Sigma_{s_u} \Sigma_{s_v}) \}.$$

Let $Z_{ni} = Z_{ni}^{(1)} + Z_{ni}^{(2)}$, where $Z_{ni}^{(1)} = 2\sum_{j=1}^{i-1} G_{nj}/\{n(n-1)\}$ and $Z_{ni}^{(2)} = 4Q_{ni}/n$. Then, $\hat{D}_{nt,0} - D_t = \sum_{i=1}^{n} Z_{ni}$, Let $\mathscr{F}_k$ be the σ-algebra generated by $\sigma\{Y_1,\dots,Y_k\}$ where $Y_i=\{Y_{i1},\dots,Y_{iT}\}$ is the collection of Y for the i-th sample. It follows that $E(Z_{nk}|\mathscr{F}_{k-1})=0$. Therefore, Z_{nk} is a sequence of martingale difference with respect to $\mathscr{F}_k$.

Let $\sigma_{ni}^2=E(Z_{ni}^2|\mathscr{F}_{i-1})$. To prove the asymptotic normality, we check two following conditions (Hall and Hedye, 1980):

Condition (a) $\sum_{i=1}^n \sigma_{ni}^2/\mathrm{var}(\hat{D}_{nt}) \xrightarrow{p} 1$;
Condition (b) $\sum_{i=1}^n E(Z_{ni}^4)/\mathrm{var}^2(\hat{D}_{nt}) \to 0$.

We first prove Condition (a). Consider $E(\sum_{i=1}^n \sigma_{ni}^2)=\sum_{i=1}^n E(Z_{ni}^2)=\sum_{i=1}^n \mathrm{var}(Z_{ni})$. Furthermore, $\mathrm{var}(\hat{D}_{nt})=\sum_{i=1}^n E(Z_{ni}^2)+2E\{\sum_{i<j} Z_{ni}E(Z_{nj}|\mathscr{F}_{j-1})\}=\sum_{i=1}^n E(Z_{ni}^2)$. Thus, we have $E(\sum_{i=1}^n \sigma_{ni}^2)=\mathrm{var}(\hat{D}_{nt})$. Thus, it suffices to show $\mathrm{var}(\sum_{i=1}^n \sigma_{ni}^2)=o\{\mathrm{var}^2(\hat{D}_{nt})\}$.

Now we obtain σ_{ni}^2 as

$$\sigma_{ni}^2=E(Z_{ni}^2|\mathscr{F}_{i-1})=\binom{n}{2}^{-2}\sum_{j=1}^{i-1}\sum_{j_1=1}^{i-1}E(G_{nj}G_{nj_1}|\mathscr{F}_{i-1})+\frac{4}{n^2}E(Q_{ni}^2|\mathscr{F}_{i-1})$$
$$+\frac{1}{n}\frac{4}{\binom{n}{2}}E\big(Q_{ni}\sum_{j=1}^{i-1}G_{nj}|\mathscr{F}_{i-1}\big)=R_{ni,1}+R_{ni,2}+R_{ni,3}.$$

Recall that $A_{js_1}=Y_{js_1}Y_{js_1}^{\mathrm{T}}-\Sigma_{s_1}$. We can further show that $R_{ni,2}$ is a constant and has no impact on $\mathrm{var}(\sum_{i=1}^n \sigma_{ni}^2)$. Moreover,

$$\begin{aligned}R_{ni,1}&=\frac{2}{\binom{n}{2}^2 w^2(t)}\sum_{j=1}^{i-1}\sum_{j_1=1}^{i-1}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{u,v,k,l\in\{1,2\}}(-1)^{|u-v|+|k-l|}\mathrm{tr}(A_{js_v}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}})\\&=R_{ni,1}^{(0)}+R_{ni,1}^{(1)},\end{aligned}$$

where $R_{ni,1}^{(0)}$ corresponds to summation of the terms where $j=j_1$ and $R_{ni,1}^{(1)}$ is the summation of the terms where $j\neq j_1$. To prove Condition (a), it suffices to show that

(a1) $\mathrm{var}(\sum_{i=1}^n R_{ni,1}^{(0)})=o(\sigma_{nt}^4)$,
(a2) $\mathrm{var}(\sum_{i=1}^n R_{ni,1}^{(1)})=o(\sigma_{nt}^4)$ and
(a3) $\mathrm{var}(\sum_{i=1}^n R_{ni,3})=o(\sigma_{nt}^4)$.

We first show (a1). We have

$$\begin{aligned}\mathrm{var}\Big(\sum_{i=1}^n R_{ni,1}^{(0)}\Big)&\le\frac{C}{w^4(t)n^5}\mathrm{var}\Big\{\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}}(-1)^{|u-v|+|k-l|}\mathrm{tr}(A_{js_v}C_{s_uh_k}A_{jh_l}C_{s_uh_k}^{\mathrm{T}})\Big\}\\&\le Cw^{-4}(t)n^{-5}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}}\mathrm{var}\Big\{\mathrm{tr}(A_{js_v}C_{s_uh_k}A_{jh_l}C_{s_uh_k}^{\mathrm{T}})\Big\}.\end{aligned}$$

Applying Lemma 5 and using the fact

$$\mathrm{tr}(A_{js_v}C_{s_uh_k}A_{jh_l}C_{s_uh_k}^{\mathrm{T}})=\mathrm{tr}\{\Gamma_{s_u}^{\mathrm{T}}\Gamma_{s_v}(Z_jZ_j^{\mathrm{T}}-I)\Gamma_{s_v}^{\mathrm{T}}\Gamma_{s_u}\Gamma_{h_k}^{\mathrm{T}}\Gamma_{h_l}(Z_jZ_j^{\mathrm{T}}-I)\Gamma_{h_l}^{\mathrm{T}}\Gamma_{h_k}\},$$

we have

$$\mathrm{var}\{\mathrm{tr}(A_{js_v}C_{s_uh_k}A_{jh_l}C_{s_uh_k}^{\mathrm{T}})\}\le C\big[\mathrm{tr}^4(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}})+\mathrm{tr}^2\{(\Gamma_{s_v}^{\mathrm{T}}C_{s_uh_k}\Gamma_{h_l})^{\otimes 2}\}\big].$$

Under Condition 1, $\mathrm{var}(\sum_{i=1}^n R_{ni,1}^{(0)}) = o(\sigma_{nt}^4)$. This completes the proof of (a1).

We next show (a2). Because of $j \neq j_1$, $E\{\mathrm{tr}(A_{js_v} C_{s_u h_k} A_{j_1 h_l} C_{s_u h_k}^{\mathrm{T}})\} = 0$. It follows that

$$\begin{aligned}\mathrm{var}(\sum_{i=1}^n R_{ni,1}^{(1)}) &\leq \frac{C}{w^4(t) n^4} \mathrm{var}\Big\{ \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{\substack{u,v,\\ k,l\in\{1,2\}}} (-1)^{|u-v|+|k-l|} \mathrm{tr}(A_{js_v} C_{s_u h_k} A_{j_1 h_l} C_{s_u h_k}^{\mathrm{T}}) \Big\} \\ &\leq C n^{-4} w^{-4}(t) \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{\substack{u,v,\\ k,l\in\{1,2\}}} \mathrm{var}\big\{\mathrm{tr}(A_{js_v} C_{s_u h_k} A_{j_1 h_l} C_{s_u h_k}^{\mathrm{T}})\big\}.\end{aligned}$$

Similar to Lemma 5, we obtain

$$\begin{aligned}&\mathrm{var}\{\mathrm{tr}(A_{js_v} C_{s_u h_k} A_{j_1 h_l} C_{s_u h_k}^{\mathrm{T}})\} = 2\mathrm{tr}^2\{(\Gamma_{s_v}^{\mathrm{T}} C_{s_u h_k} \Gamma_{h_l})^{\otimes 2}\} + 6\mathrm{tr}\{(\Gamma_{s_v}^{\mathrm{T}} C_{s_u h_k} \Gamma_{h_l})^{\otimes 4}\} \\ &\quad - 4\mathrm{tr}\{(\Gamma_{s_v}^{\mathrm{T}} C_{s_u h_k} \Gamma_{h_l} \Gamma_{s_v}^{\mathrm{T}} C_{s_u h_k} \Gamma_{h_l})^{\otimes 2}\} \leq C\mathrm{tr}^2\{(\Gamma_{s_v}^{\mathrm{T}} C_{s_u h_k} \Gamma_{h_l})^{\otimes 2}\},\end{aligned}$$

where $Q^{\otimes 2} = QQ^{\mathrm{T}}$ and $Q^{\otimes 4} = QQ^{\mathrm{T}}QQ^{\mathrm{T}}$. By Condition 1, we have $\mathrm{var}(\sum_{i=1}^n R_{ni,1}^{(1)}) = o(\sigma_{nt}^4)$, which proves (a2).

For proving (a3), a direct computation shows that

$$\begin{aligned}\sum_{i=1}^n R_{ni,3} &= \frac{16}{n\binom{n}{2}} \sum_{j=1}^n \frac{(n-j)}{w^2(t)} \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{\substack{u,v,\\ k,l\in\{1,2\}}} (-1)^{|u-v|+|k-l|} \mathrm{tr}(C_{s_u h_k} \Sigma_{h_l} C_{s_u h_k}^{\mathrm{T}} A_{js_v}) \\ &= \frac{16}{n\binom{n}{2}} \sum_{j=1}^n \frac{(n-j)}{w^2(t)} \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{u,k\in\{1,2\}} (-1)^{|u-k|} \mathrm{tr}\{(A_{js_1} - A_{js_2}) C_{s_u h_k} (\Sigma_{h_1} - \Sigma_{h_2}) C_{s_u h_k}^{\mathrm{T}}\}.\end{aligned}$$

Then,

$$\begin{aligned}\mathrm{var}(\sum_{i=1}^n R_{ni,3}) &\leq C n^{-3} w^{-4}(t) \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{u,k\in\{1,2\}} E\big[\mathrm{tr}^2\{(A_{js_1} - A_{js_2}) C_{s_u h_k} (\Sigma_{h_1} - \Sigma_{h_2}) C_{s_u h_k}^{\mathrm{T}}\}\big] \\ &= C n^{-3} w^{-4}(t) \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{u,k\in\{1,2\}} \mathrm{tr}\Big\{(\sum_{v=1}^2 (-1)^v \Gamma_{s_v}^{\mathrm{T}} C_{s_u h_k} (\Sigma_{h_1} - \Sigma_{h_2}) C_{s_u h_k}^{\mathrm{T}} \Gamma_{s_v})^2\Big\}.\end{aligned}$$

Therefore, by Condition 2, $\mathrm{var}(\sum_{i=1}^n R_{ni}) = o(\sigma_{nt}^4)$. Thus, Condition (a) is valid.

To check Condition (b), we compute

$$E(Z_{ni}^4) = E\{E(Z_{ni}^4 | \mathscr{F}_{i-1})\} \leq C n^{-8} E\Big[E\Big\{\Big(\sum_{j=1}^{i-1} G_{nj}\Big)^4 | \mathscr{F}_{i-1}\Big\}\Big] + C n^{-4} E(Q_{ni}^4) = J_{1i} + J_{2i},$$

where $E\big\{\big(\sum_{j=1}^{i-1} G_{nj}\big)^4 | \mathscr{F}_{i-1}\big\} = \sum_{j=1}^{i-1} E(G_{nj}^4 | \mathscr{F}_{i-1}) + \sum_{j\neq j_1}^{i-1} E(G_{nj}^2 G_{nj_1}^2 | \mathscr{F}_{i-1})$. Using the definition of G_{nj}, one obtain

$$G_{nj}^2 = \frac{1}{w^2(t)} \sum_{\substack{s_1,s_2,\\ h_1,h_2}}^{*} \sum_{u,v,k,l\in\{1,2\}} (-1)^{|u-v|+|k-l|} \mathrm{tr}(A_{is_u} A_{js_v}) \mathrm{tr}(A_{ih_k} A_{jh_l}).$$

For any two symmetric matrices A and B, $E[\{Z^{\mathrm{T}}AZ - \mathrm{tr}(A)\}^2 \{Z^{\mathrm{T}}BZ - \mathrm{tr}(B)\}^2] = 4\{\mathrm{tr}(A^2)\mathrm{tr}(B^2) + 2\mathrm{tr}^2(AB)\} + 16\{2\mathrm{tr}(A^2B^2) + \mathrm{tr}(ABAB)\}$ and $\mathrm{tr}[\{Z^{\mathrm{T}}AZ - \mathrm{tr}(A)\}^4] \leq$

$C\text{tr}^2(A^2)$. Accordingly,

$$J_{1i} \le \frac{C}{w^4(t)n^8}\sum_{j=1}^{i-1}\sum_{s_1,s_2}^{*}\sum_{u,v\in\{1,2\}} E\big\{\text{tr}^4(A_{is_u}A_{js_v})\big\}$$
$$+\frac{C}{w^4(t)n^8}\sum_{j\neq j_1}^{i-1}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}} E\big\{\text{tr}^2(A_{is_u}A_{js_v})\text{tr}^2(A_{ih_k}A_{j_1h_l})\big\} \le \sum_{k=1}^{5} J_{1i}^{(k)},$$

where

$$J_{1i}^{(1)} = \frac{C}{w^4(t)n^8}\sum_{j=1}^{i-1}\sum_{s_1,s_2}^{*}\sum_{u,v\in\{1,2\}} E\big\{\text{tr}^2(\Sigma_{s_u}A_{js_v}\Sigma_{s_u}A_{js_v})\big\},$$

$$J_{1i}^{(2)} = \frac{C}{w^4(t)n^8}\sum_{j\neq j_1}^{i-1}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}} E\big\{\text{tr}(A_{js_v}\Sigma_{s_u}A_{js_v}\Sigma_{s_u})\text{tr}(A_{j_1h_l}\Sigma_{h_k}A_{j_1h_l}\Sigma_{h_k})\big\},$$

$$J_{1i}^{(3)} = \frac{C}{w^4(t)n^8}\sum_{j\neq j_1}^{i-1}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}} E\big\{\text{tr}^2(A_{js_v}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}})\big\},$$

$$J_{1i}^{(4)} = \frac{C}{w^4(t)n^8}\sum_{j\neq j_1}^{i-1}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}} E\big\{\text{tr}(\Gamma_{s_u}^{\mathrm{T}}A_{js_v}\Gamma_{s_u}\Gamma_{s_u}^{\mathrm{T}}A_{js_v}\Gamma_{s_u}\Gamma_{h_k}^{\mathrm{T}}A_{j_1h_l}\Gamma_{h_k}\Gamma_{h_k}^{\mathrm{T}}A_{j_1h_l}\Gamma_{h_k})\big\},$$

$$J_{1i}^{(5)} = \frac{C}{w^4(t)n^8}\sum_{j\neq j_1}^{i-1}\sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*}\sum_{\substack{u,v,\\k,l\in\{1,2\}}} E\big\{\text{tr}(A_{js_v}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}}A_{js_v}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}})\big\}.$$

Consider the first term, $J_{1i}^{(1)}$, in the above inequality. By Lemma 5,

$$\text{var}\{\text{tr}(A_{js_v}\Sigma_{s_u}A_{js_v}\Sigma_{s_u})\} \le C\Big[\text{tr}^4(\Sigma_{s_u}\Sigma_{s_v}) + 4\text{tr}^2(\Sigma_{s_u}\Sigma_{s_v}\Sigma_{s_u}\Sigma_{s_v})$$
$$+\text{tr}\{(\Gamma_{s_v}^{\mathrm{T}}\Sigma_{s_u}\Gamma_{s_v})^4\} + \text{tr}\{(\Gamma_{s_v}^{\mathrm{T}}\Sigma_{s_u}\Gamma_{s_v}\Gamma_{s_v}^{\mathrm{T}}\Sigma_{s_u}\Gamma_{s_v})^{\otimes 2}\}\Big],$$

and $E\Big[\text{tr}(\Sigma_{s_u}A_{js_v}\Sigma_{s_u}A_{js_v})\Big] = \text{tr}^2(\Sigma_{s_u}\Sigma_{s_v}) + \text{tr}(\Sigma_{s_u}\Sigma_{s_v}\Sigma_{s_u}\Sigma_{s_v})$. Thus,

$$E\{\text{tr}^2(A_{js_v}\Sigma_{s_u}A_{js_v}\Sigma_{s_u})\} \le C\Big[\text{tr}^4(\Sigma_{s_u}\Sigma_{s_v}) + \text{tr}^2\{(\Sigma_{s_u}\Sigma_{s_v})^2\} + \text{tr}\{(\Sigma_{s_v}\Sigma_{s_u})^4\}\Big]$$
$$\le C\text{tr}^4(\Sigma_{s_u}\Sigma_{s_v}).$$

Therefore,

$$\sum_{i=1}^{n} J_{1i}^{(1)} \le \frac{C}{w^4(t)n^8}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{s_1,s_2}^{*}\sum_{u,v\in\{1,2\}} \text{tr}^4(\Sigma_{s_u}\Sigma_{s_v}) = o(\sigma_{nt}^4). \tag{S.3}$$

For the second and third terms, $J_{1i}^{(2)}$ and $J_{1i}^{(3)}$, consider the following inequality:

$$\text{tr}^2(A_{js_v}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}}) \le \text{tr}(A_{js_v}\Sigma_{s_u}A_{js_v}\Sigma_{s_u})\text{tr}(A_{j_1h_l}\Sigma_{h_k}A_{j_1h_l}\Sigma_{h_k}).$$

Thus, we have

$$\begin{aligned}\sum_{i=1}^{n}(J_{1i}^{(2)}+J_{1i}^{(3)}) &\le \frac{C}{w^4(t)n^8}\sum_{i=1}^{n}\sum_{j\neq j_1}^{i-1}\Big[\sum_{s_1,s_2}^{*}\sum_{u,v\in\{1,2\}}\{\mathrm{tr}^2(\Sigma_{s_u}\Sigma_{s_v})+\mathrm{tr}\{(\Sigma_{s_u}\Sigma_{s_v})^2\}\Big]^2\\ &\le \frac{C}{w^4(t)n^8}\sum_{i=1}^{n}\sum_{j\neq j_1}^{i-1}\Big\{\sum_{s_1,s_2}^{*}\sum_{u,v\in\{1,2\}}\mathrm{tr}^2(\Sigma_{s_u}\Sigma_{s_v})\Big\}^2=o(\sigma_{nt}^4). \qquad \text{(S.4)}\end{aligned}$$

Next, we consider the fourth term $J_{1i}^{(4)}$. We first note the following results

$$\begin{aligned}&E[\mathrm{tr}\{A(Z_1Z_1^{\mathrm{T}}-I)A^{\mathrm{T}}C(Z_1Z_1^{\mathrm{T}}-I)C^{\mathrm{T}}D(Z_2Z_2^{\mathrm{T}}-I)D^{\mathrm{T}}F(Z_2Z_2-I)F^{\mathrm{T}}\}]\\ &=\mathrm{tr}(A^{\mathrm{T}}C)\mathrm{tr}(D^{\mathrm{T}}F)\mathrm{tr}(DF^{\mathrm{T}}AC^{\mathrm{T}})+\mathrm{tr}(A^{\mathrm{T}}C)\mathrm{tr}(FD^{\mathrm{T}}FD^{\mathrm{T}}AC^{\mathrm{T}})\\ &\quad+\mathrm{tr}(D^{\mathrm{T}}F)\mathrm{tr}(CA^{\mathrm{T}}CA^{\mathrm{T}}DF^{\mathrm{T}})+\mathrm{tr}(FD^{\mathrm{T}}FD^{\mathrm{T}}CA^{\mathrm{T}}CA^{\mathrm{T}})\end{aligned}$$

for matrices A, C, D and F with appropriate dimensions. Then we can obtain

$$\begin{aligned}&E\big\{\mathrm{tr}(\Gamma_{s_u}^{\mathrm{T}}A_{js_v}\Gamma_{s_u}\Gamma_{s_u}^{\mathrm{T}}A_{js_v}\Gamma_{s_u}\Gamma_{h_k}^{\mathrm{T}}A_{j_1h_l}\Gamma_{h_k}\Gamma_{h_k}^{\mathrm{T}}A_{j_1h_l}\Gamma_{h_k})\big\}\\ &=\mathrm{tr}(\Sigma_{s_v}\Sigma_{s_u})\{\mathrm{tr}(\Sigma_{h_l}\Sigma_{h_k})\mathrm{tr}(C_{s_uh_k}\Sigma_{h_l}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v})+\mathrm{tr}(\Sigma_{h_l}\Sigma_{h_k}\Sigma_{h_l}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v}C_{s_uh_k})\}\\ &+\mathrm{tr}(\Sigma_{h_l}\Sigma_{h_k})\mathrm{tr}(\Sigma_{s_v}\Sigma_{s_u}\Sigma_{s_v}C_{s_uh_k}\Sigma_{h_l}C_{s_uh_k}^{\mathrm{T}})+\mathrm{tr}(\Sigma_{h_l}\Sigma_{h_k}\Sigma_{h_l}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v}\Sigma_{s_u}\Sigma_{s_v}C_{s_uh_k})\\ &\le \mathrm{tr}^{3/2}(\Sigma_{s_v}\Sigma_{s_u})\mathrm{tr}^{3/2}(\Sigma_{h_l}\Sigma_{h_k})+\mathrm{tr}^{3/2}(\Sigma_{s_v}\Sigma_{s_u})\mathrm{tr}^{1/2}\{(\Sigma_{h_k}\Sigma_{h_l})^4\}\\ &+\mathrm{tr}^{3/2}(\Sigma_{h_l}\Sigma_{h_k})\mathrm{tr}^{1/2}\{(\Sigma_{s_u}\Sigma_{s_v})^4\}+\mathrm{tr}^{1/2}\{(\Sigma_{h_k}\Sigma_{h_l})^4\}\mathrm{tr}^{1/2}\{(\Sigma_{s_u}\Sigma_{s_v})^4\}.\end{aligned}$$

It follows that

$$\sum_{i=1}^{n}J_{1i}^{(4)}\le\frac{C}{w^4(t)n^8}\sum_{i=1}^{n}\sum_{j\neq j_1}^{i-1}\Big\{\sum_{s_1,s_2}^{*}\sum_{u,v\in\{1,2\}}\mathrm{tr}^2(\Sigma_{s_u}\Sigma_{s_v})\Big\}^2=o(\sigma_{nt}^4). \qquad \text{(S.5)}$$

By Lemma 5,

$$\begin{aligned}E\big[\mathrm{tr}\{(C_{s_uh_k}^{\mathrm{T}}A_{js_v}C_{s_uh_k}A_{j_1h_l})^2\}\big]&=E\big[\mathrm{tr}^2(\Gamma_{s_v}^{\mathrm{T}}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}}\Gamma_{s_v})+\mathrm{tr}\{(\Gamma_{s_v}^{\mathrm{T}}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}}\Gamma_{s_v})^2\}\big]\\ &=\{Z_j^{\mathrm{T}}\Gamma_{h_l}^{\mathrm{T}}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v}C_{s_uh_k}\Gamma_{h_l}Z_j-\mathrm{tr}(\Sigma_{h_l}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v}C_{s_uh_k})\}^2\\ &\quad+\mathrm{tr}(C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v}C_{s_uh_k}A_{j_1h_l}C_{s_uh_k}^{\mathrm{T}}\Sigma_{s_v})\\ &=3\mathrm{tr}\{(\Gamma_{h_l}^{\mathrm{T}}C_{s_uh_k}^{\mathrm{T}}\Gamma_{s_v})^{\otimes 4}\}+\mathrm{tr}^2\{(\Gamma_{h_l}^{\mathrm{T}}C_{s_uh_k}^{\mathrm{T}}\Gamma_{s_v})^{\otimes 2}\}.\end{aligned}$$

This along with Condition 1 together implies $\sum_{k=1}^{n}J_{1k}^{(5)}=o(\sigma_{nt}^4)$, and further with equations (S.3) and (S.4) implies $\sum_{k=1}^{n}J_{1k}=o(\sigma_{nt}^4)$.

Finally, we consider J_{2i}. We write Q_{ni} as

$$Q_{ni}=\frac{1}{w(t)}\sum_{s_1,s_2}^{*}\Big[Z_i^{\mathrm{T}}\Big\{\sum_{u=1}^{2}(-1)^{u-1}\Gamma_{s_u}^{\mathrm{T}}(\Sigma_{s_1}-\Sigma_{s_2})\Gamma_{s_u}\Big\}Z_i^{\mathrm{T}}-\mathrm{tr}\{(\Sigma_{s_1}-\Sigma_{s_2})(\Sigma_{s_1}+\Sigma_{s_2})\}\Big],$$

where $\sum_{u=1}^{2}(-1)^{u-1}\Gamma_{s_u}^{\mathrm{T}}(\Sigma_{s_1}-\Sigma_{s_2})\Gamma_{s_u}=(\Gamma_{s_1}+\Gamma_{s_2})^{\mathrm{T}}(\Sigma_{s_1}-\Sigma_{s_2})(\Gamma_{s_1}-\Gamma_{s_2})$. Using Proposition A.1 in Chen et al. (2010),

$$J_{2i}\le\frac{C}{n^4w^2(t)}\sum_{s_1=1}^{t}\sum_{s_2=t+1}^{\mathrm{T}}\mathrm{tr}^2\Big[\{(\Gamma_{s_1}+\Gamma_{s_2})^{\mathrm{T}}(\Sigma_{s_1}-\Sigma_{s_2})(\Gamma_{s_1}-\Gamma_{s_2})\}^{\otimes 2}\Big].$$

As a result, $\sum_{i=1}^{n} J_{2i} = o(\sigma_{nt}^4)$. Condition (b) is valid. This completes the proof of the asymptotic normality of $\hat{D}_{nt}$. $\square$

Proof of Theorem 3. Using the continuous mapping theorem, we only need to prove the joint multivariate normality of $\{\hat{D}_{nt}\}_{t=1}^{T-1}$. Let $a = (a_1, \ldots, a_{T-1})^{\mathrm{T}}$ be any non-zero constant vector of length $T-1$. By the Cramér-Wold device, it suffices to show that $\sum_{t=1}^{T-1} a_t \hat{D}_{nt}$ is asymptotically normal under H_0. By Lemma 6, the variance of $\sum_{t=1}^{T-1} a_t \hat{D}_{nt}$ is $\sigma_{0T}^2 = \mathrm{var}(\sum_{t=1}^{T-1} a_t \hat{D}_{nt}) = \sum_{t=1}^{T-1}\sum_{q=1}^{T-1} a_t a_q Q_{n,tq}$. Then we wish to show $\sigma_{0T}^{-1}\sum_{t=1}^{T-1} a_t \hat{D}_{nt} \to N(0,1)$ in distribution. The asymptotic normality of $\sum_{t=1}^{T-1} a_t \hat{D}_{nt}$ can be shown by using the martingale central limit theory, which is very similar to the proof of Theorem 2. Therefore, we omit the details. $\square$

Proof of Theorem 4. Assume that the alternative H_1^* is true. First, for $t \in \{1, \ldots k_1 - 1\}$,

$$
\begin{aligned}
E(\hat{D}_{nt}) &= \mathrm{tr}(\Sigma_1^2) + \frac{1}{T-t}\sum_{h=t+1}^{k_1}\mathrm{tr}(\Sigma_h^2) + \frac{1}{T-t}\sum_{h=k_1+1}^{T}\mathrm{tr}(\Sigma_h^2) \\
&\quad - \frac{2}{t(T-t)}\sum_{s=1}^{t}\Big\{\sum_{h=t+1}^{k_1}\mathrm{tr}(\Sigma_s\Sigma_h) + \sum_{h=k_1+1}^{T}\mathrm{tr}(\Sigma_s\Sigma_h)\Big\} \\
&= \Big\{1 + \frac{k_1-t}{T-t} - \frac{2t(k_1-t)}{t(T-t)}\Big\}\mathrm{tr}(\Sigma_1^2) + \frac{T-k_1}{T-t}\mathrm{tr}(\Sigma_T^2) - \frac{2t(T-k_1)}{t(T-t)}\mathrm{tr}(\Sigma_1\Sigma_T) \\
&= \frac{T-k_1}{T-t}\mathrm{tr}\big\{(\Sigma_1-\Sigma_T)^2\big\}.
\end{aligned}
$$

Similarly, if $k_1 \le t$, then $E(\hat{D}_{nt}) = k_1\mathrm{tr}\big\{(\Sigma_1-\Sigma_T)^2\big\}/t$.

Define $B(C) = \{t \in \{1, \ldots, T-1\} : |t-k_1| \ge C\beta_n/(n\Delta_n)\}$ for some constant C. To establish the rate of convergence of the change point estimator $\hat{k}_1$, we need to show, for any $\epsilon > 0$, there exist a constant C such that $\mathrm{pr}\{|\hat{k}_1 - k_1| > C\beta_n/(n\Delta_n)\} < \epsilon$. This is equivalent to show that $\mathrm{pr}\{\hat{k}_1 \in B(C)\} < \epsilon$. Since the event $\{\hat{k}_1 \in B(C)\} \subset \{\max_{t\in B(C)}\hat{D}_{nt} > \hat{D}_{nk_1}\}$, $\mathrm{pr}\{\hat{k}_1 \in B(C)\} \le \mathrm{pr}\{\max_{t\in B(C)}\hat{D}_{nt} > \hat{D}_{k_1}\}$. Thus, it suffices to show, for any $\epsilon > 0$, there exist a constant C such that

$$
\mathrm{pr}\{\max_{t\in B(C)}\hat{D}_{nt} > \hat{D}_{k_1}\} \le \sum_{t\in B(C)}\mathrm{pr}(\hat{D}_{nt} - D_{k_1} > \hat{D}_{nk_1} - D_{k_1}) < \epsilon. \tag{S.6}
$$

Under H_1^*, we have

$$
\begin{aligned}
\hat{D}_{nt} - D_{k_1} &= \hat{D}_{nt} - D_t + D_t - D_{k_1} = \hat{D}_{nt} - D_t + \{r(t;k_1) - 1\}\mathrm{tr}\{(\Sigma_1-\Sigma_T)^2\} \\
&= \hat{D}_{nt} - D_t - |t-k_1|G(t;k_1)\mathrm{tr}\{(\Sigma_1-\Sigma_T)^2\},
\end{aligned}
$$

where $G(t;k_1) = \{1/(T-t)\}I(1 \le t \le k_1) + (1/t)I(k_1+1 \le t \le T-1)$. Then, for $t \in B(C)$,

$$
\begin{aligned}
\mathrm{pr}(\hat{D}_{nt} > \hat{D}_{nk_1}) &\le \mathrm{pr}\{|\hat{D}_{nt} - D_t| > |t-k_1|G(t;k_1)\Delta_n/2\} \\
&\qquad + \mathrm{pr}\{|\hat{D}_{nk_1} - D_{k_1}| > |t-k_1|G(t;k_1)\Delta_n/2\} \\
&\le \mathrm{pr}\{|\sigma_{nt}^{-1}(\hat{D}_{nt} - D_t)| > C\beta_n G(t;k_1)/\surd(4V_{0t} + 8nV_{1t})\} \\
&\qquad + \mathrm{pr}\{|\sigma_{nk_1}^{-1}(\hat{D}_{nk_1} - D_{k_1})| > C\beta_n G(t;k_1)/\surd(4V_{0k_1} + 8nV_{1k_1})\}.
\end{aligned}
$$

For any t and some constant C_1, $\beta_n > C_1\sqrt{(4V_{0t}+8nV_{1t})}$. Furthermore, $w(t)$ and $G(t;k_1)$ are bounded away from zero for $t \in B(C)$. Thus, by Chebyshev's inequality,

$$\mathrm{pr}(\hat{D}_{nt} > \hat{D}_{nk_1}) \le \mathrm{pr}\{|\sigma_{nt}^{-1}(\hat{D}_{nt}-D_t)| > C\} + \mathrm{pr}\{|\sigma_{nk_1}^{-1}(\hat{D}_{nk_1}-D_{k_1})| > C\} \le \frac{2}{C^2} < \frac{\epsilon}{T},$$

for large enough C. Therefore, (S.6) is true. This finishes the proof of Theorem 4. □

Proof of Theorem 5. Let $k_0 = 0$ and $k_{q+1} = T$. Denote the common covariances between the change points k_j and k_{j+1} as $\tilde{\Sigma}_j$ for $j = 0,\ldots,q$. To show that $\max_t D_t$ is at one of the change points, it is enough to show that $\max_t D_t$ cannot be attained at any time points except change points $k_1,\ldots,k_q$. Thus, we need to show that the maximum of D_t is not attainable for t in the following sets: (1) $t \in \{1,\ldots,k_1-1\}$; (2) $t \in \{k_q+1,\ldots,T-1\}$; and (3) $t \in \{k_l+1,\ldots,k_{l+1}-1\}$ for some $l \in \{1,\ldots,q-1\}$. We do not need to consider case (1) if $k_1 = 1$ or case (3) if $k_q = T-1$. Without loss of generality, we assume $k_1 > 1$ and $k_q < T-1$ in the following proof.

First, if $t \in \{1,\ldots,k_1-1\}$, then using the definition of D_t, we have

$$\begin{aligned} D_t &= \frac{1}{t(T-t)}\sum_{s_1=1}^{t}\sum_{s_2=t+1}^{k_1}\|\Sigma_{s_1}-\Sigma_{s_2}\|_F^2 + \frac{1}{t(T-t)}\sum_{s_1=1}^{t}\sum_{s_2=k_1+1}^{T}\|\Sigma_{s_1}-\Sigma_{s_2}\|_F^2 \\ &= \frac{1}{T-t}\sum_{s_2=k_1+1}^{T}\|\tilde{\Sigma}_0-\Sigma_{s_2}\|_F^2 \end{aligned}$$

which is an increasing function of t in this scenario. Therefore, the maximum value of D_t will not be at any $t \in \{1,\ldots,k_1-1\}$.

Second, if $t \in \{k_q+1,\ldots,T-1\}$, then

$$\begin{aligned} D_t &= \frac{1}{t(T-t)}\sum_{s_1=1}^{k_q}\sum_{s_2=t+1}^{T}\|\Sigma_{s_1}-\Sigma_{s_2}\|_F^2 + \frac{1}{t(T-t)}\sum_{s_1=k_q+1}^{t}\sum_{s_2=t+1}^{T}\|\Sigma_{s_1}-\Sigma_{s_2}\|_F^2 \\ &= \frac{1}{t}\sum_{s_1=1}^{k_q}\|\tilde{\Sigma}_q-\Sigma_{s_1}\|_F^2 \end{aligned}$$

which is a decreasing function of t. Therefore, the maximum value of D_t will not be at any $t \in \{k_q+1,\ldots,T-1\}$.

At last, let us consider the third case with $t \in \{k_l+1,\ldots,k_{l+1}-1\}$ for some $l \in \{1,\ldots,q-1\}$. We rewrite D_t as

$$\begin{aligned} D_t = \frac{1}{t(T-t)}\Big\{&\sum_{i=0}^{l-1}\sum_{j=l+1}^{q}(k_{i+1}-k_i)(k_{j+1}-k_j)\|\tilde{\Sigma}_i-\tilde{\Sigma}_j\|_F^2 \\ &+ (t-k_l)\sum_{j=l+1}^{q}(k_{j+1}-k_j)\|\tilde{\Sigma}_l-\tilde{\Sigma}_j\|_F^2 + (k_{l+1}-t)\sum_{i=0}^{l-1}(k_{i+1}-k_i)\|\tilde{\Sigma}_i-\tilde{\Sigma}_l\|_F^2\Big\}. \end{aligned}$$

Since $\|\tilde{\Sigma}_i-\tilde{\Sigma}_j\|_F^2 = \|\tilde{\Sigma}_i-\tilde{\Sigma}_l\|_F^2 + \|\tilde{\Sigma}_l-\tilde{\Sigma}_j\|_F^2 + 2\mathrm{tr}\{(\tilde{\Sigma}_i-\tilde{\Sigma}_l)(\tilde{\Sigma}_l-\tilde{\Sigma}_j)\}$, we further write D_t as

$$D_t = \frac{1}{t(T-t)}\big\{2\Delta + tA + (T-t)B\big\},$$

where

$$\Delta = \sum_{i=0}^{l-1} \sum_{j=l+1}^{q} (k_{i+1} - k_i)(k_{j+1} - k_j)\mathrm{tr}\{(\tilde{\Sigma}_i - \tilde{\Sigma}_l)(\tilde{\Sigma}_l - \tilde{\Sigma}_j)\},$$

$A = \sum_{j=l+1}^{q}(k_{j+1} - k_j)\|\tilde{\Sigma}_l - \tilde{\Sigma}_j\|_F^2$ and $B = \sum_{i=0}^{l-1}(k_{i+1} - k_i)\|\tilde{\Sigma}_i - \tilde{\Sigma}_l\|_F^2$. Then we can use the fact that $1/\{t(T-t)\} = (1/T)\{1/t + 1/(T-t)\}$ to further write D_t as

$$D_t = \frac{1}{t}\Big(A + \frac{2\Delta}{T}\Big) + \frac{1}{T-t}\Big(B + \frac{2\Delta}{T}\Big).$$

We will consider four cases, (a)-(d), according to the signs of $A + 2\Delta/T$ and $B + 2\Delta/T$.

(a) If $A + 2\Delta/T \geq 0$ and $B + 2\Delta/T \leq 0$, then D_t is a decreasing function of t. In this case, the maximum of D_t will not be at any t for $t \in \{k_l + 1, \ldots, k_{l+1} - 1\}$.

(b) If $A + 2\Delta/T \leq 0$ and $B + 2\Delta/T \geq 0$, then D_t is an increasing function of t. In this case, the maximum of D_t will not be at any t for $t \in \{k_l + 1, \ldots, k_{l+1} - 1\}$.

(c) If $A + 2\Delta/T > 0$ and $B + 2\Delta/T > 0$, then the derivative of D_t with respect to t is

$$D_t' = \frac{1}{t^2(T-t)^2}\{(B-A)t^2 + 2(A + \frac{2\Delta}{T})Tt - (A + \frac{2\Delta}{T})T^2\}.$$

The denominator of D_t' is always positive for $t \in \{k_l + 1, \ldots, k_{l+1} - 1\}$. Thus, to determine the sign of D_t', we only need to know the sign of the numerator of D_t'.

The numerator of D_t' is a quadratic form of t. To know the sign of the numerator of D_t', we consider two cases: (i) $B > A$ and (ii) $B < A$. In the case (i) with $B > A$, one of the solution of $t^2(T-t)^2 D_t' = 0$ is less than 0, another solution t_0 is greater than 0. If $t_0 \in (k_l, k_{l+1})$, then D_t' is negative for $k_l < t < t_0$ and positive for $t_0 < t < k_{l+1}$. This implies that the function D_t decreases for $k_l < t < t_0$ and increases for $t_0 < t < k_{l+1}$. Therefore, D_t attains its minimum at t_0 and the maximum of D_t will not be attained within (k_l, k_{l+1}). If $t_0 \notin (k_l, k_{l+1})$, then D_t' is either always negative or always positive for $t \in (k_l, k_{l+1})$. In this case, D_t is a monotonic function of t and hence the maximum of D_t will not be attained within (k_l, k_{l+1}).

In the case (ii) with $B < A$, it can be shown that $t^2(T-t)^2 D_t' = 0$ has two solutions, $t_1, t_2 = T\big[(A + 2\Delta/T)/(A-B) \pm \sqrt{\{(A + 2\Delta/T)/(A-B) - 1/2\}^2 - 1/4}\big]$. Here, t_1, t_2 corresponds to the positive and negative sign, respectively. Because $B + 2\Delta/T > 0$, $(A + 2\Delta/T)/(A-B) > 1$. It follows that $t_2 > T$. Similar to the case of $B > A$, if $t_1 \in (k_l, k_{l+1})$, the function D_t decreases for $k_l < t < t_1$ and increases for $t_1 < t < k_{l+1}$. Therefore, D_t attains its minimum at t_0 and the maximum of D_t will not be attained within (k_l, k_{l+1}). If $t_1 \notin (k_l, k_{l+1})$, D_t is a monotone function of t and hence the maximum of D_t will not be attained within (k_l, k_{l+1}). In summary, the maximum of D_t will not be attained within (k_l, k_{l+1}) if $A + 2\Delta/T > 0$ and $B + 2\Delta/T > 0$.

(d) If $A + 2\Delta/T < 0$ and $B + 2\Delta/T < 0$, then $2\Delta/T < 0$ because $A > 0$ and $B > 0$. Thus, $A - 2|\Delta|/T < 0$. Using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} A < 2|\Delta|/T &\leq \sum_{i=0}^{l-1} \sum_{j=l+1}^{q} (k_{i+1} - k_i)(k_{j+1} - k_j)(\|\tilde{\Sigma}_i - \tilde{\Sigma}_l\|_F^2 + \|\tilde{\Sigma}_l - \tilde{\Sigma}_j\|_F^2)/T \\ &= \{(T - k_{l+1})A + k_l B\}/T. \end{aligned}$$

The above inequality implies that $A/B < k_l/k_{l+1} < 1$. On the other hand, $B < 2|\Delta|/T \leq \{(T - k_{l+1})A + k_l B\}/T$, which implies that $A/B > (T - k_l)/(T - k_{l+1}) > 1$. This is a contradiction. Therefore, case (d) is not possible.

By the results of (a)-(d), the maximum of D_t will not attain within $\{k_l+1,\ldots,k_{l+1}-1\}$ for case (3). Thus, the proof is completed. $\square$

Proof of Theorem 6. At the beginning of the binary segmentation algorithm, we have $\mathcal{M}_n[1,T] > W_{\alpha_n}[1,T]$ with probability one because, for any $t \in \{1,\ldots,T-1\}$,

$$
\begin{aligned}
&\mathrm{pr}(\mathcal{M}_n[1,T] > W_{\alpha_n}[1,T]) \geq \mathrm{pr}(\sigma_{nt,0}^{-1}[1,T]\hat{D}_{nt}[1,T] > W_{\alpha_n}[1,T]) \\
&= \mathrm{pr}\big\{\sigma_{nt}^{-1}[1,T](\hat{D}_{nt}[1,T] - D_t[1,T]) > \sigma_{nt}^{-1}[1,T](\sigma_{nt,0}[1,T]W_{\alpha_n}[1,T] - D_t[1,T])\big\} \\
&= 1 - \Phi\big\{\sigma_{nt}^{-1}[1,T](\sigma_{nt,0}[1,T]W_{\alpha_n}[1,T] - D_t[1,T])\big\} \to 1,
\end{aligned}
$$

where we used the condition $W_{\alpha_n} = o(\mathrm{mSNR})$ in Theorem 6. Therefore, using Theorems 4 and 5, one change point in $\{k_1,\ldots,k_q\}$ will be detected and estimated with probability 1 because $\beta_n[1,T] = o(nD_{k_s}[1,T])$ for some $s \in \{1,\ldots,q\}$. Each subsequence satisfies the condition $W_{\alpha_n} = o(\mathrm{mSNR})$ in Theorem 6 and hence the detection continues.

Suppose we have detected less than q change points. By the assumptions in this theorem, there exists a segment, $\{l_1+1,\ldots,l_2\}$, that contains a change point, k_s, such that $W_{\alpha_n} = o(\mathrm{mSNR})$ and $\beta_n[(l_1+1),l_2] = o\{nD_{k_s}[(l_1+1),l_2]\}$ hold. Therefore, by similar arguments as above, a change point will be detected and estimated consistently in the segment. Thus, $\hat{q} \geq q$. Once $\hat{q}$ reaches q, all subsequent segments have end points at the change points and two boundary points $1,k_1,\ldots,k_q,T$. Then, by Theorem 3, $\mathcal{M}_n[l_1,l_2] < W_{\alpha_n}$ with probability one as $\alpha_n \to 0$. This implies that no additional change point will be detected. The proof is completed. $\square$

3. NON-GAUSSIAN RANDOM ERRORS

To relax the Gaussian assumption, we assume the following data generation model for $\varepsilon_i = (\varepsilon_{i1}^{\mathrm{T}},\ldots,\varepsilon_{iT}^{\mathrm{T}})^{\mathrm{T}}$ and $\varepsilon_i = \Gamma Z_i$ where $\Gamma = (\Gamma_1^{\mathrm{T}},\ldots,\Gamma_T^{\mathrm{T}})^{\mathrm{T}}$ is a $Tp \times m$ matrix with $m \geq Tp$ such that $\Sigma = \Gamma\Gamma^{\mathrm{T}}$ and $\Gamma_s\Gamma_t^{\mathrm{T}} = C_{st}$. We assume $Z_1,\ldots,Z_n$ are independent and identically distributed m-dimensional random vectors such that $E(Z_1) = 0$ and $\mathrm{var}(Z_1) = I_m$. Write $Z_1 = (Z_{11},...,Z_{1m})^{\mathrm{T}}$. We assume that each Z_{1l} has a uniformly bounded 8th moment. Also, we assume there exists a finite constant such that for $l = 1,\ldots,m$, $E(Z_{1l}^4) = 3 + \Delta$ and for any integers $l_v \geq 0$ with $\sum_{v=1}^q l_v = 8$, $E(Z_{1i_1}^{l_1}\cdots Z_{1i_q}^{l_q}) = E(Z_{1i_1}^{l_1})\cdots E(Z_{1i_q}^{l_q})$, whenever $i_1,\ldots,i_q$ are distinct indices.

Under Condition 1 and the above setup, it can be shown that the leading order of the variance of $\hat{D}_{nt}$ is

$$
\begin{aligned}
\mathrm{var}(\hat{D}_{nt}) = &\frac{4}{n^2w^2(t)} \sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*} \sum_{\substack{u,v,\\k,l\in\{1,2\}}} (-1)^{|u-v|+|k-l|}\mathrm{tr}^2(C_{s_uh_k}C_{s_vh_l}^{\mathrm{T}}) \\
&+ \frac{8}{nw^2(t)} \sum_{\substack{s_1,s_2,\\h_1,h_2}}^{*} \sum_{u,k\in\{1,2\}} (-1)^{|u-k|}\big[\mathrm{tr}\{(\Sigma_{s_1}-\Sigma_{s_2})C_{s_uh_k}(\Sigma_{h_1}-\Sigma_{h_2})C_{s_uh_k}^{\mathrm{T}}\} \\
&\qquad + \Delta\mathrm{tr}\{\Gamma_{s_u}^{\mathrm{T}}(\Sigma_{s_1}-\Sigma_{s_2})\Gamma_{s_u} \circ \Gamma_{h_k}^{\mathrm{T}}(\Sigma_{h_1}-\Sigma_{h_2})\Gamma_{h_k}\}\big].
\end{aligned}
$$

Under the null hypothesis, $\mathrm{var}(\hat{D}_{nt}) = 4V_{0t}/\{n^2w^2(t)\}$. The variance V_{0t} can be estimated using the formula given below equation (6) in the main paper. The results in Theorems 2 and 3 can be established in a similar way.

To illustrate the numerical performance of the proposed method under the non-Gaussian setup, we generated data from the linear process model $Y_{it} = \mu_t + \sum_{h=0}^{L} A_{t,h}\eta_{i(t-h)}$ for $i = 1,\ldots,n$

and $t = 1, \ldots, T$, where $A_{t,h}$ is a $p \times p$ matrix, $\mu_t = 0$ and η_{it} are p-dimensional random vectors with each element independently generated from a standardized Gamma distribution with shape parameter 4 and scale parameter 0·5.

Let $k_1 = [T/2]$ be the largest integer no greater than $T/2$. For $t \in \{1, \ldots, k_1\}$, we set $A_{t,h} = A^{(1)} = \{0{\cdot}6^{|i-j|} I(|i-j| < p/5)\}$. For $t \in \{k_1+1, \ldots, T\}$, we set $A_{t,h} = A^{(2)} = \{(0{\cdot}6 + \delta)^{|i-j|} I(|i-j| < p/5)\}$. If $\delta = 0$, $A^{(1)}$ and $A^{(2)}$ are the same. Hence, the covariances, Σ_t, are the same for all $t \in \{1, \ldots, T\}$ and H_0 is true. If $\delta \neq 0$, the null hypothesis is not true and k_1 is the underlying true covariance change point.

In the simulation studies, we set $n = 40, 50$ and 60, with $p = 500, 750$ and 1000. The number of repeated measurements, T, was set to be 5 and 8 and set $L = 3$. The simulation results reported in Tables S1 and S2 were based on 500 simulation replications.

Table S1. *Empirical size and power of the proposed test, percentages of simulation replications that reject the null hypothesis for data generated from a standardized Gamma distribution under the nominal level 5%*

		$T = 5$			$T = 8$		
		p			p		
δ	n	500	750	1000	500	750	1000
	40	3·6	4·0	4·4	4·0	3·6	4·6
0(size)	50	4·2	5·2	4·4	5·2	4·8	4·8
	60	4·6	3·8	4·6	5·0	5·0	5·6
	40	23·4	21·4	28·2	35·2	38·6	31·2
0·05	50	38·2	36·2	33·4	47·8	50·4	47·8
	60	46·4	46·8	46·2	64·6	67·2	66·4
	40	99·8	99·8	100	100	100	100
0·10	50	100	100	100	100	100	100
	60	100	100	100	100	100	100

Table S1 reports the empirical size and power of the proposed test under the null and alternative hypotheses. We observe that Type I error is well controlled with the empirical sizes close to the nominal level of 5%. The results demonstrate the robustness of the proposed method for non-Gaussian distributed random vectors. When the differences between covariance matrices increase, the power of the proposed test increases accordingly. Table S2 reports the performance of the proposed change point identification procedure under the non-Gaussian distributed random vectors. We observe that the percentages of correct identification with non-Gaussian random vectors are similar to those under the Gaussian setup.

4. POWER ENHANCED TEST FOR SPARSE ALTERNATIVES

The proposed test statistic, $\mathcal{M}_n$, is powerful for alternatives with small absolute differences in many components of Σ_t. However, it might not be very powerful for sparse alternatives with the differences among Σ_t only residing in a few components. To enhance the power of the proposed test for sparse alternatives, we include an additional term with $\mathcal{M}_n$, as an idea in Fan et al. (2015).

Let $\bar{Y}_{s_1 v} = \sum_{i=1}^{n} Y_{is_1 v}/n$ be the sample mean of the vth component measured at time s_1, and define $\hat{\sigma}_{s_1,uv} = \sum_{i=1}^{n} (Y_{is_1 u} - \bar{Y}_{s_1 u})(Y_{is_1 v} - \bar{Y}_{s_1 v})/(n-1)$ as the sample covariance between components $u, v \in \{1, \ldots, p\}$ at time s_1. Define $\hat{D}_{nt,uv} = \sum_{s_1=1}^{t} \sum_{s_2=t+1}^{T} (\hat{\sigma}_{s_1,uv} - \hat{\sigma}_{s_2,uv})^2$

Table S2. *Percentages of correct change point identification among all rejected hypotheses for data generated from a standardized Gamma distribution*

		$T=5$			$T=8$		
			p			p	
δ	n	500	750	1000	500	750	1000
	40	32·48	42·99	35·46	26·70	30·05	30·13
0·05	50	50·26	52·49	46·71	40·17	42·06	46·86
	60	49·14	55·56	57·58	48·30	50·00	52·41
	40	93·79	96·79	95·80	95·20	94·60	95·00
0·10	50	98·80	99·40	99·00	98·60	97·00	97·20
	60	99·60	99·20	99·80	99·60	99·20	99·20

as an estimator of $D_{nt,uv} = \sum_{s_1=1}^{t}\sum_{s_2=t+1}^{T}(\sigma_{s_1,uv} - \sigma_{s_2,uv})^2$. The estimator $\hat{D}_{nt,uv}$ is a consistent estimator of $D_{nt,uv}$. Let $C_{s_k h_t}^{(uv)}$ be the (u,v) component of $C_{s_k h_t}$ and $\sigma_{h_t}^{(uv)}$ is the (u,v) component of Σ_{h_t}. To define the variance of $\hat{D}_{nt,uv}$, define the following notation:

$$
\begin{aligned}
F^{(uv)}_{s_k s_l h_t h_s} &= \sigma^{(uv)}_{s_l}\sigma^{(uv)}_{h_t}\{C^{(vu)}_{h_s s_k}C^{(uv)}_{s_k h_s} + C^{(vv)}_{h_s s_k}C^{(uu)}_{s_k h_s}\} + \sigma^{(uv)}_{s_l}\sigma^{(uv)}_{h_s}\{C^{(vu)}_{h_t s_k}C^{(uv)}_{s_k h_t} + C^{(vv)}_{h_t s_k}C^{(uu)}_{s_k h_t}\} \\
&\quad + \sigma^{(uv)}_{s_k}\sigma^{(uv)}_{h_t}\{C^{(vu)}_{h_s s_l}C^{(uv)}_{s_l h_s} + C^{(vv)}_{h_s s_l}C^{(uu)}_{s_l h_s}\} + \sigma^{(uv)}_{s_k}\sigma^{(uv)}_{h_s}\{C^{(vu)}_{h_t s_l}C^{(uv)}_{s_l h_t} + C^{(vv)}_{h_t s_l}C^{(uu)}_{s_l h_t}\}, \\
G^{(uv)}_{s_k s_l h_t h_s} &= \{C^{(vu)}_{s_k h_s}C^{(uv)}_{s_k h_s} + C^{(vv)}_{s_k h_s}C^{(uu)}_{s_k h_s}\}\{C^{(vu)}_{s_l h_t}C^{(uv)}_{s_l h_t} + C^{(vv)}_{s_l h_t}C^{(uu)}_{s_l h_t}\} \\
&\quad + \{C^{(vu)}_{s_k h_t}C^{(uv)}_{s_k h_t} + C^{(vv)}_{s_k h_t}C^{(uu)}_{s_k h_t}\}\{C^{(vu)}_{s_l h_s}C^{(uv)}_{s_l h_s} + C^{(vv)}_{s_l h_s}C^{(uu)}_{s_l h_s}\}.
\end{aligned}
$$

The leading order term of the variance of $\hat{D}_{nt,uv}$ is

$$
\sigma^2_{nt,uv} = \frac{1}{w^2(t)} \sum^{*}_{\substack{s_1,s_2,\\ h_1,h_2}} \sum_{\substack{k,l,\\ s,t\in\{1,2\}}} (-1)^{|k-l|+|s-t|}\{n^{-1}F^{(uv)}_{s_k s_l h_t h_s} + n^{-2}G^{(uv)}_{s_k s_l h_t h_s}\}. \tag{S.7}
$$

Under H_0, the first term in (S.7) is 0. Namely, $\sum^{*}_{\substack{s_1,s_2,\\ h_1,h_2}} \sum_{\substack{k,l,\\ s,t\in\{1,2\}}} (-1)^{|k-l|+|s-t|}F^{(uv)}_{s_k s_l h_t h_s} = 0$. The leading term in the variance of $\hat{D}_{nt,uv}$ under H_0 is

$$
\sigma^2_{nt,uv0} = \sum^{*}_{\substack{s_1,s_2,\\ h_1,h_2}} \sum_{\substack{k,l,\\ s,t\in\{1,2\}}} (-1)^{|k-l|+|s-t|}G^{(uv)}_{s_k s_l h_t h_s}/n^2.
$$

Let $\hat{G}^{(uv)}_{s_k s_l h_t h_s}$ be a sample plug-in estimate of $G^{(uv)}_{s_k s_l h_t h_s}$, and $\hat{\sigma}^2_{nt,uv0}$ be the corresponding sample estimate of $\sigma^2_{nt,uv0}$. Then, the power enhanced test statistic is

$$
\mathcal{M}^*_n = \max_{1\le t\le T-1}\Big\{\hat{\sigma}^{-1}_{nt,0}\hat{D}_{nt} + \lambda_n \sum_{u\le v} I(\hat{D}_{nt,uv} > \delta_{n,p}\hat{\sigma}_{nt,uv0})\Big\},
$$

where $\delta_{n,p}$ and λ_n are tuning parameters. The tuning parameters are chosen such that the second part of $\mathcal{M}^*_n$ equals zero with probability tending to one under H_0, and it converges to a large number under sparse alternatives.

We now discuss the choices of tuning parameters for the above power enhanced test statistic. Let $R = (\rho_{ij})$ be the correlation matrix corresponding to the common covariance Σ_1 under H_0. Define $N_j(\alpha) = \text{card}\{i : |\rho_{ij}| > (\log p)^{-1-\alpha}\}$ and $\Lambda(r) = \{i : |\rho_{ij}| > r \text{ for some } j \ne i\}$. We assume the following condition used in Cai et al. (2013).

Condition S.1. Suppose that there exists a α and a set $\pi \subset \{1, \ldots, p\}$ whose size is $o(p)$ such that $\max_{1\le j\le p, j\notin\pi} N_j(\alpha) = o(p^\gamma)$ for all $\gamma > 0$. In addition, there exists a $r < 1$ and a sequence of numbers $\Lambda_{p,r} = o(p)$ so that card$\{\Lambda(r)\} \le \Lambda_{p,r}$.

Define $l_{s_1 s_2} = \max_{1\le u\le v\le p}(\hat\sigma_{s_1,uv} - \hat\sigma_{s_2,uv})^2/\sigma_{ns_1s_2,uv0}$ where $\sigma^2_{ns_1s_2,uv0} = \text{var}\{(\hat\sigma_{s_1,uv} - \hat\sigma_{s_2,uv})^2\}$ under H_0. Similar to the proof of Theorem 1 in Cai et al. (2013), under Condition S.1 and H_0, we can show that

$$\text{pr}\{l_{s_1s_2} - 4\log(p) + \log\log(p) \le t\} \to \exp\{-\exp(-t/2)/\surd(8\pi)\}. \tag{S.8}$$

Define $L_{uv} = \hat D_{nt,uv}/\hat\sigma_{nt,uv0}$ and $L_n = \max_{1\le u\le v\le p} L_{uv}$. Denote the second term in M_n^* as $M_{n1}^* = \lambda_n \sum_{u\le v} I(\hat D_{nt,uv} > \delta_{n,p}\hat\sigma_{nt,uv0})$. Because $\sum_{s_1=1}^{t}\sum_{s_2=t+1}^{T} \sigma_{ns_1s_2,uv0}/\sigma_{nt,uv0} \le K$, uniformly for all u, v for a constant $K > 0$, and uniform consistency of $\hat\sigma_{nt,uv0}$ to $\sigma_{nt,uv0}$, we have, under H_0,

$$\begin{aligned}
\text{pr}(M_{n1}^* = 0) &\ge \text{pr}(L_n \le \delta_{n,p}) = \text{pr}(\max_{1\le u\le v\le p} \hat D_{nt,uv}/\hat\sigma_{nt,uv0} \le \delta_{n,p}) \\
&= \text{pr}(\max_{1\le u\le v\le p} \sum_{s_1=1}^{t}\sum_{s_2=t+1}^{T} \frac{(\hat\sigma_{s_1,uv} - \hat\sigma_{s_2,uv})^2}{\sigma_{ns_1s_2,uv0}} \frac{\sigma_{ns_1s_2,uv0}}{\sigma_{nt,uv0}} \le \delta_{n,p}) \\
&\ge \text{pr}(\max_{1\le u\le v\le p} \max_{\substack{1\le s_1\le t,\\ t+1\le s_2\le T}} \frac{(\hat\sigma_{s_1,uv} - \hat\sigma_{s_2,uv})^2}{\sigma_{ns_1s_2,uv0}} \sum_{s_1=1}^{t}\sum_{s_2=t+1}^{T} \frac{\sigma_{ns_1s_2,uv0}}{\sigma_{nt,uv0}} \le \delta_{n,p}) \\
&\ge \text{pr}(\max_{1\le u\le v\le p} \max_{\substack{1\le s_1\le t,\\ t+1\le s_2\le T}} (\hat\sigma_{s_1,uv} - \hat\sigma_{s_2,uv})^2/\sigma_{ns_1s_2,uv0} \le \delta_{n,p}/K) \\
&\ge 1 - \sum_{s_1=1}^{t}\sum_{s_2=t+1}^{T} \text{pr}(l_{s_1s_2} > \delta_{n,p}/K).
\end{aligned}$$

Applying the result in (S.8), if $\delta_{n,p}/K - 4\log(p) + \log\log(p) \to \infty$, then $\text{pr}(M_{n1}^* = 0) \to 1$. We suggest choose $\delta_{n,p}$ at the order of $\log(n)\log(p)$ and λ_n to be a constant based on our numerical experiments. In summary, the tuning parameters $\delta_{n,p}$ and λ_n ensure that, under the null hypothesis, M_{n1}^* converges to zero with probability one.

We conducted a numerical simulation to illustrate the performance of the power enhanced test statistic under sparse alternatives. The data were generated according to setting (I) described in § 5 of the main paper except for a sparse alternative design. Specifically, let $k_1 = [T/2]$ be the largest integer no greater than $T/2$. For $t \in \{1, \ldots, k_1\}$, we set $A_{t,h} = A^{(1)}$ for $h \in \{0, \ldots, L\}$. For $t \in \{k_1 + 1, \ldots, T\}$, we set $A_{t,h} = A^{(2)}$, where $A_h^{(1)} = \{0{\cdot}6^{|i-j|} I(|i-j| < p/5)\}$. Under the null hypothesis, $A_h^{(2)}$ was set equal to $A_h^{(1)}$. Under the sparse alternative hypothesis, $A_h^{(2)}$ was the same as $A_h^{(1)}$ except the components within $\{|i-j| < 2, i < p/25\}$ were set to $1{\cdot}4$.

Table S3 reports the empirical size and power of the test based on $\mathcal{M}_n$ and $\mathcal{M}_n^*$. In the simulation, the tuning parameter $\delta_{n,p}$ was set to $0{\cdot}5\log(n)\log(p)$, and λ_n was set to $0{\cdot}15$. We observe that both tests can control the type I error, and the power enhanced test does not inflate the type I error. More importantly, the power enhanced test statistic has greater power under the sparse alternative setting.

Table S3. *Empirical size and power, percentages of simulation replications that reject the null hypothesis for the test statistic $\mathcal{M}_n$ and the power enhanced test statistic $\mathcal{M}_n^*$*

		$\mathcal{M}_n$		$\mathcal{M}_n^*$	
n	p	Null	Alternative	Null	Alternative
40	500	5·2	35·4	6·0	67·6
40	750	3·2	34·6	3·6	62·8
40	1,000	4·6	36·4	4·6	62·4
50	500	5·2	47·8	5·6	91·6
50	750	6·4	47·2	6·6	94·2
50	1,000	3·4	52·6	3·4	97·0
60	500	3·8	56·8	4·4	98·8
60	750	4·8	65·8	5·6	99·2
60	1,000	4·2	66·4	4·2	99·8
80	500	4·0	81·2	4·0	100
80	750	3·2	86·4	3·8	100
80	1,000	4·6	84·8	4·6	100

5. ACCURACY OF CORRELATION MATRIX ESTIMATOR OF V_{nD}

This section aims to evaluate the numerical performance of the correlation matrix estimator, $\hat{V}_{nD}$, proposed immediately following Theorem 3. To measure the difference between $\hat{V}_{nD}$ and V_{nD}, we used the average component-wise quadratic distance, namely, $(T-1)^{-2}\|\hat{V}_{nD} - V_{nD}\|_F^2$. Figure 1 illustrates the average of $(T-1)^{-2}\|\hat{V}_{nD} - V_{nD}\|_F^2$ based on 500 simulation replications conducted in setting (I) under the null hypothesis with $T = 5$. We observe that the correlation matrix estimator, $\hat{V}_{nD}$, is reliable when $n = 40$. The performance further improves as the sample size increases.

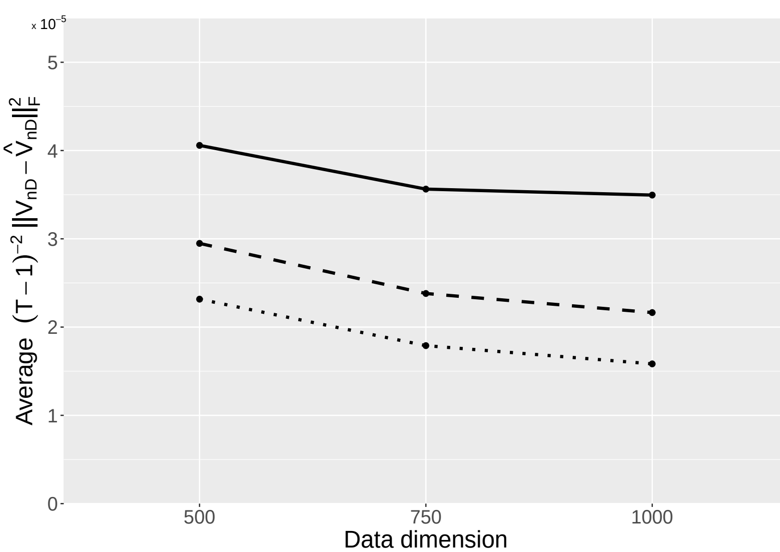

Fig. 1. The average component-wise quadratic distance between $\hat{V}_{nD}$ and V_{nD}. The top solid line is for $n = 40$; the middle dashed line is for $n = 50$; the bottom dotted line is for $n = 60$. The scale of the y-axis is 10^{-5}.

6. COMPARISON WITH A PAIR-WISE BASED METHOD

In this section, we compare our proposed method with a pair-wise based method that is similar to the method proposed by Zalesky et al. (2014). In the pair-wise based method, we first obtain a p-value for testing the homogeneity of each component of the covariance matrix for every pair of coordinates (u, v) with $u \leq v$ and $u, v \in \{1, \ldots, p\}$, and then apply the Bonferroni correction to all the p-values to control the family-wise error rate.

In the first step, for each pair (u, v) with $u \leq v$ and $u, v \in \{1, \ldots, p\}$, we test the following hypothesis

$$H_{0,uv} : \sigma_{1,uv} = \cdots = \sigma_{T,uv},$$

versus

$$H_{1,uv} : \sigma_{1,uv} = \cdots = \sigma_{k_1,uv} \neq \sigma_{k_1+1,uv} = \cdots = \sigma_{k_q,uv} \neq \sigma_{k_q+1,uv} = \cdots = \sigma_{T,uv}.$$

To test $H_{0,uv}$, we apply the statistic $\hat{D}_{nt,uv}$ defined in § 4, and define $\hat{D}_{n,uv} = n \sum_{t=1}^{T-1} \hat{D}_{nt,uv}$. Under $H_{0,uv}$, the asymptotic distribution of $\hat{D}_{n,uv}$ is $\sum_{l=1}^{\infty} \lambda_l \chi_l^2$, where χ_l^2 are independent chi-square distributions with degree of freedom 1, and λ_l's are the eigenvalues of the kernel of $\hat{D}_{n,uv}$. In practice, one can approximate the weighted chi-square distribution using a scaled chi-square distribution. Thus, we approximate the distribution of $\hat{D}_{n,uv}$ by $b\chi_\nu^2$, where $b = \sigma_{uv}^2/(2\mu_{uv})$ and $\nu = 2\mu_{uv}^2/\sigma_{uv}^2$. Here μ_{uv} and σ_{uv}^2 are the mean and variance of $\hat{D}_{n,uv}$ under $H_{0,uv}$, respectively. The variance of $\hat{D}_{n,uv}$ under $H_{0,uv}$ is

$$\sigma_{uv}^2 = \sum_{t=1}^{T-1} \sum_{q=1}^{T-1} \sum_{s_1=1}^{t} \sum_{h_1=1}^{q} \sum_{s_2=t+1}^{T} \sum_{h_2=q+1}^{T} \sum_{\substack{k,l,\\ s,t \in \{1,2\}}} (-1)^{|k-l|+|s-t|} G_{s_k s_l h_t h_s}^{(uv)},$$

where $G_{s_k s_l h_t h_s}^{(uv)}$ is defined in § 4. The mean of $\hat{D}_{n,uv}$ under the null $H_{0,uv}$ is

$$\mu_{uv} = \sum_{t=1}^{T-1} \sum_{s_1=1}^{t} \sum_{s_2=t+1}^{T} \sum_{a,b \in \{1,2\}} (-1)^{|a-b|} \{C_{s_a s_b}^{(uu)} C_{s_a s_b}^{(vv)} + C_{s_a s_b}^{(uv)} C_{s_a s_b}^{(vu)}\}.$$

We then approximate the distribution of $\hat{D}_{n,uv}$ by $\hat{b}\chi_{\hat{\nu}}^2$ where $\hat{b} = \hat{\sigma}_{uv}^2/(2\hat{\mu}_{uv})$ and $\hat{\nu} = 2\hat{\mu}_{uv}^2/\hat{\sigma}_{uv}^2$. The p-value for the (u, v) pair is computed as $p_{uv} = \mathrm{pr}(\hat{b}\chi_{\hat{\nu}}^2 > \hat{D}_{n,uv})$.

In the second step, we apply the Bonferroni correction to control the family-wise error rate. Define $p_{\min} = \min_{u \leq v} p_{uv}$ as the minimum of all the pair-wise p-values. If $p_{\min} < 2\alpha/\{p(p+1)\}$, then we reject the null hypothesis on the homogeneity of covariance matrices at the α level.

To compare the proposed methods with the pair-wise based method, we conducted a simulation study using the simulation setup given in §4. The simulation results are summarized in Table S4. We observe that the pair-wise based method has very conservative size under the null hypothesis when sample size is less than 80, but it improves as sample size increases. Under the alternatives, the power of the pair-wise based method is low for the small sample cases, but it increases as sample size increases to 80. However, in all the cases, our proposed power enhanced method has superior power than the pair-wise based method.

REFERENCES

BICKEL, P. J. & RITOV, Y. (2009). An exponential inequality for U-statistics with applications to testing. *Probab. Eng. Inf. Sci.* **9**, 39–52.

Table S4. *Empirical size and power, percentages rejecting the null hypotheses in the simulations, for the pair-wise based test and the power enhanced test statistic* $\mathcal{M}_n^*$

		Pair-wise based test		$\mathcal{M}_n^*$	
n	p	Null	Alternative	Null	Alternative
40	500	0·2	0·2	6·0	67·6
40	750	0·0	0·4	3·6	62·8
40	1,000	0·0	0·0	4·6	62·4
50	500	0·4	0·4	5·6	91·6
50	750	0·2	0·0	6·6	94·2
50	1,000	0·2	0·2	3·4	97·0
60	500	0·6	12·2	4·4	98·8
60	750	0·2	4·8	5·6	99·2
60	1,000	0·6	1·0	4·2	99·8
80	500	0·4	97·6	4·0	100
80	750	2·4	98·8	3·8	100
80	1,000	2·0	96·8	4·6	100

CAI, T., LIU, W. & XIA, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Am. Statist. Assoc.* **108**, 265–277.

CHEN, S. X., ZHANG, L. & ZHONG, P.-S. (2010). Testing high dimensional covariance matrices. *J. Am. Statist. Assoc.* **105**, 810–819.

FAN, J., LIAO, Y. & YAO, J. (2015). Power enhancement in high dimensional cross-sectional tests. *Econometrica*, **83**, 1497–1541.

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

ZALESKY, A., FORNITO, A., COCCHI, L., GOLLO, L. L. & BREAKSPEAR, M. (2014). Time-resolved resting-state brain networks. *Proc. Natl. Acad. Sci.* **111**, 10341–10346.