

Supplementary Materials: Mean Independent Component Analysis for Multivariate Time Series

Chung Eun Lee and Zeda Li

Paul H. Chook Department of Information Systems and Statistics, Baruch College

In the supplementary material, we include the proofs for Theorem 1 and Theorem 2 and state additional Theorem S1 with its proof, and report additional simulations, real data application results, and figures.

1 Proofs of Theorem 1 and Theorem 2

We shall first state the lemmas used in the proof for Theorem 1.

Lemma 1. *Let h_0 be fixed. For any $\mathbf{U}, \mathbf{V} \in \mathcal{H}_{\mathcal{D}}$ when p is fixed or p diverges, it holds that*

$$|\mathcal{S}_{h_0}(\mathbf{U}) - \mathcal{S}_{h_0}(\mathbf{V})| \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}),$$

where c is a positive constant, and

$$\left| \widehat{\mathcal{S}}_{h_0}(\mathbf{U}) - \widehat{\mathcal{S}}_{h_0}(\mathbf{V}) \right| \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}),$$

where $c = O_p(1)$.

Proof of Lemma 1. Let $\mathbf{U} = (\mathbf{U}_{\cdot 1}, \dots, \mathbf{U}_{\cdot p})$ and $\mathbf{V} = (\mathbf{V}_{\cdot 1}, \dots, \mathbf{V}_{\cdot p})$, where $\mathbf{U}_{\cdot i} = (\mathbf{U}_{1i}, \dots, \mathbf{U}_{pi})^\top$, $\mathbf{V}_{\cdot j} = (\mathbf{V}_{1j}, \dots, \mathbf{V}_{pj})^\top$. We assume that $\mathbf{U}_{\cdot i}^\top \mathbf{V}_{\cdot i} \in [0, 1]$ and $\mathcal{D}(\mathbf{U}, \mathbf{V}) = \sqrt{\sum_{i=1}^p \{1 - \text{tr}(\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top)\}}$ by arranging the orders of $\{\mathbf{U}_{\cdot i}\}_{i=1}^p$ and $\{\mathbf{V}_{\cdot i}\}_{i=1}^p$. Recall

that

$$\begin{aligned}
(tr \{ \mathbf{U} \mathbf{U}^\top \}) \mathcal{S}_{h_0}(\mathbf{U}) &= \sum_{h=0}^{h_0} \sum_{i=1}^p \sum_{j \neq i} -\mathbb{E} [\mathbf{Y}_t \mathbf{U}_{\cdot j} \mathbf{U}_{\cdot j}^\top \mathbf{Y}'_t | \mathbf{U}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h})] \\
&= \sum_{h=0}^{h_0} \sum_{i=1}^p -\mathbb{E} [\mathbf{Y}_t (\mathbf{I} - \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top) \mathbf{Y}'_t \| \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \|].
\end{aligned}$$

By the Cauchy Schwarz inequality, we have

$$\begin{aligned}
& |\mathcal{S}_{h_0}(\mathbf{U}) - \mathcal{S}_{h_0}(\mathbf{V})| \\
& \leq \frac{1}{p} \sum_{h=0}^{h_0} \sum_{i=1}^p |-\mathbb{E} [\mathbf{Y}_t^\top (\mathbf{I} - \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top) \mathbf{Y}'_t \{ \| \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \| - \| \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \| \}]| \\
& + \frac{1}{p} \sum_{h=0}^{h_0} \sum_{i=1}^p |-\mathbb{E} [\mathbf{Y}_t^\top (\mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top - \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top) \mathbf{Y}'_t \| \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \|]| \\
& \leq \frac{1}{p} \sum_{h=0}^{h_0} \sum_{i=1}^p \{ tr (\mathbf{I} - \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top) \}^{1/2} \left(\mathbb{E} \| (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top - \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \|^2 \right)^{1/2} \\
& + \frac{1}{p} \sum_{h=0}^{h_0} \sum_{i=1}^p \left(\mathbb{E} \left[\{ \mathbf{Y}_t^\top (\mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top - \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top) \mathbf{Y}'_t \}^2 \right] \right)^{1/2} \left(\mathbb{E} \| \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \|^2 \right)^{1/2} \\
& \leq \frac{1}{p} \sum_{h=0}^{h_0} \sum_{i=1}^p c \sqrt{p} \cdot \| \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top - \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top \|_F \leq \frac{c}{p} \sqrt{p} \cdot \sum_{h=0}^{h_0} \sum_{i=1}^p \sqrt{2} \{ 1 - tr (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) \}^{1/2} \quad (\text{S1}) \\
& \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}). \quad (\text{S2})
\end{aligned}$$

Note that (S1) is obtained by

$$\begin{aligned}
\mathbb{E} \| \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \|^2 &\leq c \cdot tr [\mathbb{E} \{ \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top \mathbf{Y}_{t-h} \mathbf{Y}_{t-h}^\top \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top \}] = c \cdot \| \mathbf{V}_{\cdot i} \|^4, \\
\mathbb{E} \| (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top - \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \|^2 \\
&\leq c \cdot tr [\mathbb{E} \{ (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top - \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) \mathbf{Y}_{t-h} \mathbf{Y}_{t-h}^\top (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top - \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) \}] = c \cdot \| \mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top - \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top \|_F^2,
\end{aligned}$$

and (S2) is implied by

$$\sum_{i=1}^p \{ 1 - tr (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) \}^{1/2} \leq \sqrt{p} \left(\sum_{i=1}^p \{ 1 - tr (\mathbf{U}_{\cdot i} \mathbf{U}_{\cdot i}^\top \mathbf{V}_{\cdot i} \mathbf{V}_{\cdot i}^\top) \} \right)^{1/2} = \sqrt{p} \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}).$$

By using the similar arguments, we can show that $\left| \widehat{\mathcal{S}}_{h_0}(\mathbf{U}) - \widehat{\mathcal{S}}_{h_0}(\mathbf{V}) \right| \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V})$. This completes the proof of Lemma 1. \diamond

When p diverges, we shall often use Lemma 3, whose proof relies on Lemma 2. Lemma 2 is proved in Appendix of Dette & Spreckelsen (2004), hence its proof is omitted. The following lemma indicates that similar result of Lemma 1 in Yoshihara (1976) holds for the varying kernel under a certain finite moment condition.

Lemma 2 (Lemma A.1. in Dette & Spreckelsen (2004)). *Let $\varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_r})$, $t_1 < t_2 < \dots < t_r$ be a symmetric kernel which depends on n . If $(\mathbf{Z}_t)_{t=1}^n$ is an absolutely regular process with mixing coefficients $\beta(j)$ and if for some $\delta > 0$, $1 \leq j < r$ and for all p , there exists a finite constant \widetilde{C} such that*

$$M_n = \max \left\{ \mathbb{E} \left| \varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_r}) \right|^{1+\delta}, \mathbb{E} \left| \varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_j}, \mathbf{Z}'_{t_{j+1}}, \dots, \mathbf{Z}'_{t_r}) \right|^{1+\delta} \right\} < \widetilde{C} < \infty,$$

then

$$\left| \mathbb{E} \varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_r}) - \mathbb{E} \varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_j}, \mathbf{Z}'_{t_{j+1}}, \dots, \mathbf{Z}'_{t_r}) \right| \leq 4M_n^{\frac{1}{1+\delta}} \beta^{\frac{\delta}{1+\delta}}(t_{j+1} - t_j),$$

where $(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_j})$ are independent of $(\mathbf{Z}'_{t_{j+1}}, \dots, \mathbf{Z}'_{t_r})$.

Lemma 3. *Let $\varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_r})$, $t_1 < t_2 < \dots < t_r$, $r = 2, 3, 4$ be a symmetric kernel which depends on n and $(\mathbf{Z}_t)_{t=1}^n$ is an absolutely regular process with mixing coefficients $\beta(j)$. Assume that for any $z_{t_1}, \dots, z_{t_{r-1}}$, $\mathbb{E} \varphi_n(z_{t_1}, \dots, z_{t_{r-1}}, \mathbf{Z}_{t_r}) = 0$ almost surely. If for some $\delta > 0$, and for all p , there exists a finite constant \widetilde{C} such that $\sum_{k=1}^{\infty} k \beta^{\frac{\delta}{1+\delta}}(k) < \infty$ and*

$$\mathbb{E} \left| \varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_r}) \right|^{2(1+\delta)} < \widetilde{C} < \infty, \quad \mathbb{E} \left| \varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \dots, \mathbf{Z}'_{t_r}) \right|^{2(1+\delta)} < \widetilde{C} < \infty,$$

then

$$\mathbb{E} \left[\left\{ \frac{1}{\binom{n}{r}} \sum_{t_1 < \dots < t_r} \varphi_n(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_r}) \right\}^2 \right] = O(n^{-1}).$$

Proof of Lemma 3. We shall consider the case when $r = 3$ and show the proof since the proofs in the other cases $r = 2, 4$ are similar. By Cauchy Schwarz inequality and $\mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})|^{2(1+\delta)} < \infty$, for any (t_1, \dots, t_6) , we have

$$\begin{aligned} & \mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})\varphi_n(\mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})|^{1+\delta} \\ & \leq \left(\mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})|^{2(1+\delta)} \right)^{1/2} \left(\mathbb{E} |\varphi_n(\mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})|^{2(1+\delta)} \right)^{1/2} \\ & < \infty. \end{aligned}$$

By this fact, we have $\mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})\varphi_n(\mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})| < \infty$, where there are at least two identical indices among (t_1, \dots, t_6) , e.g., $(t_1, \dots, t_6) = (1, 2, 3, 4, 5, 1)$, $(t_1, \dots, t_6) = (1, 2, 3, 2, 5, 1)$. This further implies that

$$\begin{aligned} & \mathbb{E} \left\{ \frac{1}{\binom{n}{3}} \sum_{t_1 < t_2 < t_3} \varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}) \right\}^2 \\ & = O(n^{-6}) \sum_{t_1 < t_2 < t_3, t_4 < t_5 < t_6} \mathbb{E} \varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})\varphi_n(\mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6}) \\ & \leq O(n^{-6}) \sum_{t_1 < t_2 < t_3 < t_4 < t_5 < t_6} |\mathbb{E} h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})| + O(n^{-1}), \end{aligned}$$

where $h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6}) = \frac{1}{6!} \sum_{(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_6})} \varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})\varphi_n(\mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})$, $\sum_{(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_6})}$ denotes the summation over all permutations of the 6-tuple of indices of $(\mathbf{Z}_{t_1}, \dots, \mathbf{Z}_{t_6})$. Notice that for any $t_1 < t_2 < t_3 < t_4 < t_5 < t_6$, we have $\mathbb{E} h_n(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3}, \mathbf{Z}'_{t_4}, \mathbf{Z}'_{t_5}, \mathbf{Z}'_{t_6}) = \mathbb{E} h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}'_{t_6}) = 0$ since φ_n is degenerate. Also, for any $t_1 < t_2 < t_3 < t_4 < t_5 < t_6$, we have

$$\begin{aligned} & \mathbb{E} |h_n(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3}, \mathbf{Z}'_{t_4}, \mathbf{Z}'_{t_5}, \mathbf{Z}'_{t_6})|^{1+\delta} \\ & \leq \left(\mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3})|^{2(1+\delta)} \right)^{1/2} \left(\mathbb{E} |\varphi_n(\mathbf{Z}'_{t_4}, \mathbf{Z}'_{t_5}, \mathbf{Z}'_{t_6})|^{2(1+\delta)} \right)^{1/2} < \infty \end{aligned}$$

under the assumptions that $\mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})|^{2(1+\delta)} < \tilde{C} < \infty$ and $\mathbb{E} |\varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3})|^{2(1+\delta)} < \tilde{C} < \infty$. Similarly, we obtain $\mathbb{E} |h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}'_{t_6})|^{1+\delta} < \infty$. Thus, we apply

Lemma A.1 in [Dette & Spreckelsen \(2004\)](#) with $j = 1, 5$ and obtain

$$\begin{aligned}
& O(n^{-6}) \sum_{t_1 < t_2 < t_3 < t_4 < t_5 < t_6} |\mathbb{E} h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})| \\
& \leq O(n^{-6}) C_{n,1}^{\frac{1}{1+\delta}} \sum_{t_1 < t_2 < t_3 < t_4 < t_5 < t_6} \min \left(\beta^{\frac{\delta}{1+\delta}}(t_6 - t_5), \beta^{\frac{\delta}{1+\delta}}(t_2 - t_1) \right) \\
& \leq O(n^{-6}) C_{n,1}^{\frac{1}{1+\delta}} \sum_{\substack{t_1 < t_2 < t_3 < t_4 < t_5 < t_6 \\ t_2 - t_1 \leq t_6 - t_5}} \beta^{\frac{\delta}{1+\delta}}(t_6 - t_5) \\
& + O(n^{-6}) C_{n,1}^{\frac{1}{1+\delta}} \sum_{\substack{t_1 < t_2 < t_3 < t_4 < t_5 < t_6 \\ t_6 - t_5 < t_2 - t_1}} \beta^{\frac{\delta}{1+\delta}}(t_2 - t_1), \tag{S3}
\end{aligned}$$

where

$$\begin{aligned}
C_{n,1} = \max & \left\{ \max_{t_1 < t_2 < t_3 < t_4 < t_5 < t_6} \mathbb{E} |h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}_{t_6})|^{1+\delta}, \right. \\
& \max_{t_1 < t_2 < t_3 < t_4 < t_5 < t_6} \mathbb{E} |h_n(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3}, \mathbf{Z}'_{t_4}, \mathbf{Z}'_{t_5}, \mathbf{Z}'_{t_6})|^{1+\delta}, \\
& \left. \max_{t_1 < t_2 < t_3 < t_4 < t_5 < t_6} \mathbb{E} |h_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}, \mathbf{Z}_{t_4}, \mathbf{Z}_{t_5}, \mathbf{Z}'_{t_6})|^{1+\delta} \right\} < \infty.
\end{aligned}$$

Next, we can show that

$$\sum_{\substack{t_1 < t_2 < t_3 < t_4 < t_5 < t_6 \\ t_2 - t_1 \leq t_6 - t_5}} \beta^{\frac{\delta}{1+\delta}}(t_6 - t_5) \leq n^4 \sum_{k=1}^n k \beta^{\frac{\delta}{1+\delta}}(k). \tag{S4}$$

$$\sum_{\substack{t_1 < t_2 < t_3 < t_4 < t_5 < t_6 \\ t_6 - t_5 < t_2 - t_1}} \beta^{\frac{\delta}{1+\delta}}(t_2 - t_1) \leq n^4 \sum_{k=1}^n k \beta^{\frac{\delta}{1+\delta}}(k), \tag{S5}$$

Since the proofs are similar, we shall focus on showing (S4). Notice that we have $t_2 - t_1 \leq t_6 - t_5$. Hence,

$$\begin{aligned}
& \sum_{\substack{t_1 < t_2 < t_3 < t_4 < t_5 < t_6 \\ t_2 - t_1 \leq t_6 - t_5}} \beta^{\frac{\delta}{1+\delta}}(t_6 - t_5) \leq n^3 \sum_{t_5=1}^{n-1} \sum_{t_6=t_5+1}^n (t_6 - t_5) \beta^{\frac{\delta}{1+\delta}}(t_6 - t_5) \\
& \leq n^4 \sum_{k=1}^n k \beta^{\frac{\delta}{1+\delta}}(k).
\end{aligned}$$

Finally, by (S3) and $\sum_{k=1}^{\infty} k\beta^{\frac{\delta}{1+\delta}}(k) < \infty$, we obtain

$$\mathbb{E} \left\{ \frac{1}{\binom{n}{3}} \sum_{t_1 < t_2 < t_3} \varphi_n(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}) \right\}^2 \leq O(n^{-2}) \cdot C_{n,1}^{\frac{1}{1+\delta}} \sum_{k=1}^n k\beta^{\frac{\delta}{1+\delta}}(k) + O(n^{-1}) = O(n^{-1}).$$

This completes the proof of Lemma 3. \diamond

Proof of Theorem 1. We shall first show the proof of the first assertion in Theorem 1 which is the case when p is fixed. Recall that

$$\begin{aligned} \mathcal{S}_{h_0}(\tilde{\mathbf{A}}) &= \left(\frac{1}{\text{tr} \{ \tilde{\mathbf{A}} \tilde{\mathbf{A}}^\top \}} \right) \sum_{h=0}^{h_0} \sum_{i=1}^p \sum_{j \neq i} \text{MDD}^2(\tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}) \\ &= \left(\frac{1}{p} \right) \sum_{h=0}^{h_0} \sum_{i=1}^p \text{MDD}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}), \end{aligned}$$

where $\tilde{\mathbf{A}}_{.-i} = (\tilde{\mathbf{A}}_{.1}, \dots, \tilde{\mathbf{A}}_{.i-1}, \tilde{\mathbf{A}}_{.i+1}, \dots, \tilde{\mathbf{A}}_{.p}) \in \mathbb{R}^{p \times (p-1)}$. First, we show that for any $\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}$,

$$\left| \widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}) - \text{MDD}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}) \right| = O_p(n^{-1/2}). \quad (\text{S6})$$

which implies

$$\begin{aligned} & \left| \widehat{\mathcal{S}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\tilde{\mathbf{A}}) \right| \\ & \leq \left(\frac{1}{p} \right) \sum_{h=0}^{h_0} \sum_{i=1}^p \left| \widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}) - \text{MDD}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}) \right| \\ & = O_p(n^{-1/2}). \end{aligned} \quad (\text{S7})$$

Notice that $\widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h})$ can be rewritten as

$$\widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{t-h}) = \frac{(n-h-1)}{(n-h)} \{ \mathcal{U}_{n,1} + \mathcal{U}_{n,2} + \mathcal{U}_{n,3} + \mathcal{U}_{n,4} \}, \quad (\text{S8})$$

where

$$\mathcal{U}_{n,1} = -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_{t_1})^\top (\tilde{\mathbf{A}}_{.-i}^\top \mathbf{Y}_{t_2}) \left| \tilde{\mathbf{A}}_{.i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|,$$

$$\begin{aligned}\mathcal{U}_{n,2} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1})^\top (-\tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}}) \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|, \\ \mathcal{U}_{n,3} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (-\tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}})^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2}) \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|, \\ \mathcal{U}_{n,4} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (\tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}})^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}}) \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|,\end{aligned}$$

and $\bar{\mathbf{Y}} = \frac{1}{(n-h)} \sum_{t=h+1}^n \mathbf{Y}_t$. Here, $\mathcal{U}_{n,1}$ is a U-statistic of order 2 with the kernel,

$$g_1(\mathbf{Z}_t, \mathbf{Z}'_t) = -\frac{1}{2} \left\{ (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t)^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}'_t) + (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}'_t)^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t) \right\} \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right|,$$

where $\mathbf{Z}_t = (\mathbf{Y}_t, \mathbf{Y}_{t-h})$. Under the second condition in Assumption 1 and by the Cauchy Schwarz inequality, for any t_1 and t_2 , we have

$$\begin{aligned}& \mathbb{E} |g_1(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})|^{2+\delta} \\ & \leq c \cdot \left(\mathbb{E} \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\|^{6+3\delta} \right)^{1/3} \left(\mathbb{E} \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right\|^{6+3\delta} \right)^{1/3} \left(\mathbb{E} \left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t_1-h} \right|^{6+3\delta} + \mathbb{E} \left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t_2-h} \right|^{6+3\delta} \right)^{1/3} \\ & \leq c \cdot \left(\left\| \tilde{\mathbf{A}}_{\cdot-i} \right\|_2^{6+3\delta} \mathbb{E} \left\| \mathbf{Y}_{t_1} \right\|^{6+3\delta} \right)^{1/3} \left(\left\| \tilde{\mathbf{A}}_{\cdot-i} \right\|_2^{6+3\delta} \mathbb{E} \left\| \mathbf{Y}_{t_2} \right\|^{6+3\delta} \right)^{1/3} \left(\left\| \tilde{\mathbf{A}}_{\cdot i} \right\|^{6+3\delta} \mathbb{E} \left\| \mathbf{Y}_{t_1-h} \right\|^{6+3\delta} \right)^{1/3} \\ & \leq c \cdot \mathbb{E} \left\| \mathbf{Y}_{t_1} \right\|^{6+3\delta} < \infty,\end{aligned} \tag{S9}$$

where $\|\mathbf{B}\|_2$ is a spectral norm of a matrix \mathbf{B} and $\|\tilde{\mathbf{A}}_{\cdot-i}\|_2 = 1$ since $\tilde{\mathbf{A}}_{\cdot-i}$ is semi-orthogonal, i.e., $\tilde{\mathbf{A}}_{\cdot-i}^\top \tilde{\mathbf{A}}_{\cdot-i} = \mathbf{I} \in \mathbb{R}^{p-1}$. This implies that $\mathbb{E} |g_1(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_1+k})|^{2+\delta} < \infty$ for $k = 1, 2, \dots$. Similarly, we have $\mathbb{E} |g_1(\mathbf{Z}_t, \mathbf{Z}'_t)|^{2+\delta} < \infty$. Hence, we apply Theorem 1 in [Yoshihara \(1976\)](#) and obtain

$$\left| \mathcal{U}_{n,1} - \text{MDD}^2(\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h}) \right|^2 = O_p(n^{-1}). \tag{S10}$$

Furthermore, since we have $\mathbb{E} \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\| \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right\| \left\| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| < \infty$ by (S9) and $\mathbb{E} \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\| \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\| \left\| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| < \infty$ after using similar arguments in (S9),

we obtain

$$\begin{aligned}
\mathcal{U}_{n,2} &= \frac{1}{(n-h)^2(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \sum_{t_3} (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1})^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_3}) \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right| \\
&= \frac{1}{(n-h)^2(n-h-1)} \sum_{t_1 \neq t_2 \neq t_3} (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1})^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_3}) \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right| \\
&+ \frac{1}{(n-h)^2(n-h-1)} \sum_{t_1 \neq t_2} (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1})^\top \left\{ (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1}) + (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2}) \right\} \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right| \\
&= \frac{(n-h-2)}{(n-h)} \bar{\mathcal{U}}_{n,2} + O_p(n^{-1}), \tag{S11}
\end{aligned}$$

where $\bar{\mathcal{U}}_{n,2}$ is a U-statistic of order 3 with the kernel,

$$g_2(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t) = \frac{1}{3!} \sum_{(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)} (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t)^\top (\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}'_t) \left| \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}''_{t-h}) \right|,$$

where $\sum_{(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)}$ denotes the summation over all permutations of the 3-tuple of $(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)$.

Under the second condition in Assumption 1 and after applying similar arguments in (S9), we have

$$\begin{aligned}
&\mathbb{E} |g_2(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})|^{2+\delta} \\
&\leq c \cdot \left(\mathbb{E} \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\|^{6+3\delta} \right)^{1/3} \left(\mathbb{E} \left\| \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right\|^{6+3\delta} \right)^{1/3} \left(\mathbb{E} \left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t_1-h} \right|^{6+3\delta} + \mathbb{E} \left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t_3-h} \right|^{6+3\delta} \right)^{1/3} \\
&< \infty
\end{aligned}$$

for any t_1, t_2, t_3 which implies that $\mathbb{E} |g_2(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_1+k_1}, \mathbf{Z}_{t_1+k_1+k_2})|^{2+\delta} < \infty$ for $k_1, k_2 = 1, 2, \dots$. By using similar arguments, we further have $\mathbb{E} |g_2(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)|^{2+\delta} < \infty$. Thus, we apply Theorem 1 in [Yoshihara \(1976\)](#) to $\bar{\mathcal{U}}_{n,2}$ and obtain $\bar{\mathcal{U}}_{n,2} = O_p(n^{-1/2})$ which implies that $\mathcal{U}_{n,2} = O_p(n^{-1/2})$. Similarly, we obtain $\mathcal{U}_{n,j} = O_p(n^{-1/2})$, $j = 3, 4$ which implies (S6) and (S7). As, we have (S7) for any $\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}$, this further implies that

$$\sup_{\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}} \left| \widehat{\mathcal{S}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\tilde{\mathbf{A}}) \right| \rightarrow^p 0.$$

By the result above, Lemma 1, and the argmax mapping theorem (Theorem 3.2.2 and Corollary 3.2.3) in [Vaart & Wellner \(1997\)](#), we obtain

$$\mathcal{D}(\widehat{\mathbf{A}}, \mathbf{A}) \rightarrow^p 0 \text{ as } n \rightarrow \infty.$$

This completes the proof for the first assertion in Theorem 1.

Next, we shall show the second assertion in Theorem 1. The main idea of the proof for diverging p is using Corollary 2.3, Proposition 2.4 in [White & Wooldridge \(1991\)](#), who have shown the consistency of the sieve extremum estimate with dependent data where the sieve extremum estimate is obtained by optimizing an objective function. In their setting, the dimension of sieve grows with the sample size n ; see [Chen \(2007\)](#) and [White & Wooldridge \(1991\)](#) for more details. Also, we shall follow the proof in [Dette & Spreckelsen \(2004\)](#), who have shown the convergence rate of U-statistic with a varying kernel and dependent data. In particular, we shall use Lemma A.1. in [Dette & Spreckelsen \(2004\)](#) and Lemma 3 to find the order in probability for $\left| \widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) \right|$ for any $\widetilde{\mathbf{A}}$.

Recall our objective function,

$$\begin{aligned} \mathcal{S}_{h_0}(\widetilde{\mathbf{A}}) &= \left(\frac{1}{p} \right) \sum_{h=0}^{h_0} \sum_{i=1}^p \text{MDD}^2 \left(\widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) \\ &= \left(\frac{p\sqrt{p}}{p} \right) \sum_{h=0}^{h_0} \sum_{i=1}^p \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right). \end{aligned}$$

We further note that our objective function depends on n since p grows with n . For the ease of notation, we shall suppress the index n for our objective function.

We shall show that

$$\left| \widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) \right| = O_p(n^{-1/2}) \quad (\text{S12})$$

By (S8), we have

$$\widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \widetilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) = \frac{(n-h-1)}{(n-h)} \left\{ \widetilde{\mathcal{U}}_{n,1} + \widetilde{\mathcal{U}}_{n,2} + \widetilde{\mathcal{U}}_{n,3} + \widetilde{\mathcal{U}}_{n,4} \right\}, \quad (\text{S13})$$

where

$$\begin{aligned}\tilde{\mathcal{U}}_{n,1} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right) \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|, \\ \tilde{\mathcal{U}}_{n,2} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right)^\top \left(-\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}} \right) \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|, \\ \tilde{\mathcal{U}}_{n,3} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(-\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}} \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right) \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|, \\ \tilde{\mathcal{U}}_{n,4} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}} \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \bar{\mathbf{Y}} \right) \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right|.\end{aligned}$$

Here, $\tilde{\mathcal{U}}_{n,1}$ is a U-statistic of order 2 with a varying kernel which depends on n since p grows with n .

$$\begin{aligned}g_{n,1}(\mathbf{Z}_t, \mathbf{Z}'_t) &= -\frac{1}{2} \left\{ \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}'_t \right) + \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}'_t \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \right) \right\} \\ &\quad \times \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right|.\end{aligned}$$

Under the fourth condition in Assumption 1, we shall first show that

$$\mathbb{E}|g_{n,1}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})|^{2+2\delta} < c \cdot C < \infty, \quad \mathbb{E}|g_{n,1}(\mathbf{Z}_t, \mathbf{Z}'_t)|^{2+2\delta} < c \cdot C < \infty \text{ for any } t_1, t_2, p. \text{(S14)}$$

By the Cauchy Schwarz inequality, we have

$$\begin{aligned}
& \mathbb{E}|g_{n,1}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})|^{2+2\delta} \\
& \leq c \cdot \mathbb{E} \left\{ \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\|^{2+2\delta} \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right\|^{2+2\delta} \left(\frac{1}{\sqrt{p}} \right)^{2+2\delta} \left(|\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1-h}|^{2+2\delta} + |\tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2-h}|^{2+2\delta} \right) \right\} \\
& \leq c \cdot \left(\mathbb{E} \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \right\|^{6+6\delta} \right)^{1/3} \left(\mathbb{E} \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \right\|^{6+6\delta} \right)^{1/3} \left(\mathbb{E} \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \right|^{6+6\delta} \right)^{1/3} \\
& \leq c \cdot \left(\|\tilde{\mathbf{A}}_{\cdot-i}\|_2^{6+6\delta} \mathbb{E} \left\| \frac{1}{\sqrt{p}} \mathbf{Y}_t \right\|^{6+6\delta} \right)^{1/3} \left(\|\tilde{\mathbf{A}}_{\cdot-i}\|_2^{6+6\delta} \mathbb{E} \left\| \frac{1}{\sqrt{p}} \mathbf{Y}_t \right\|^{6+6\delta} \right)^{1/3} \\
& \times \left(\|\tilde{\mathbf{A}}_{\cdot-i}\|^{6+6\delta} \mathbb{E} \left\| \frac{1}{\sqrt{p}} \mathbf{Y}_t \right\|^{6+6\delta} \right)^{1/3} \\
& \leq c \cdot \left(\mathbb{E} \left\| \frac{1}{\sqrt{p}} \mathbf{Y}_t \right\|^{6+6\delta} \right) < c \cdot C < \infty, \tag{S15}
\end{aligned}$$

where $\tilde{\mathbf{A}}_{\cdot-i}$ is semi-orthogonal. Similarly, we obtain $\mathbb{E}|g_{n,1}(\mathbf{Z}_t, \mathbf{Z}'_t)|^{2+2\delta} < c \cdot C < \infty$.

Next, similar to the proof in [Dette & Spreckelsen \(2004\)](#), we shall apply Hoeffding decomposition to $\tilde{\mathcal{U}}_{n,1}$, i.e.,

$$\tilde{\mathcal{U}}_{n,1} = \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t-h} \right) + H_n^{(1)} + H_n^{(2)},$$

where $H_n^{(1)} = \frac{2}{(n-h)} \sum_t g_{n,1}^{(1)}(\mathbf{Z}_t)$, $H_n^{(2)} = \frac{1}{(n-h)(n-h-1)} \sum_{t_1 \neq t_2} g_{n,1}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})$, $g_{n,1}^{(1)}(z) = \mathbb{E} \{g_{n,1}(z, \mathbf{Z}')\} - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t-h} \right)$, $g_{n,1}^{(2)}(z, z') = g_{n,1}(z, z') - g_{n,1}^{(1)}(z) - g_{n,1}^{(1)}(z') - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t-h} \right)$.

By [\(S14\)](#) and $\text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t-h} \right) < \infty$ under the fourth condition in Assumption 1, for any $t_1 \neq t_2$, we have

$$\mathbb{E} \left\{ g_{n,1}^{(1)}(\mathbf{Z}_t)^2 \right\} < \infty, \mathbb{E} \left| g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}_{t_2}) \right|^{1+\delta} < \infty, \mathbb{E} \left| g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}'_{t_2}) \right|^{1+\delta} < \infty \tag{S16}$$

By Lemma A.1. in [Dette & Spreckelsen \(2004\)](#), $\mathbb{E} g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}'_{t_2}) = 0$, and $\sum_{t=1}^n \beta^{\delta/(1+\delta)}(t) < \infty$

∞ , we obtain

$$\begin{aligned}
\mathbb{E} \{ (H_n^{(1)})^2 \} &= O(n^{-2}) \sum_{t_1, t_2} \mathbb{E} \left\{ g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}_{t_2}) \right\} \\
&= O(n^{-1}) \mathbb{E} \left\{ g_{n,1}^{(1)}(\mathbf{Z}_t)^2 \right\} + O(n^{-2}) \sum_{t_1 \neq t_2} \mathbb{E} \left\{ g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}_{t_2}) \right\} \\
&\leq O(n^{-1}) \mathbb{E} \left\{ g_{n,1}^{(1)}(\mathbf{Z}_t)^2 \right\} + O(n^{-2}) \sum_{t=1}^n (n-t) \mathbb{E} \left\{ g_{n,1}^{(1)}(\mathbf{Z}_1) g_{n,1}^{(1)}(\mathbf{Z}_{t+1}) \right\} \\
&\leq O(n^{-1}) + O(n^{-1}) \sum_{t=1}^n \left| \mathbb{E} g_{n,1}^{(1)}(\mathbf{Z}_1) g_{n,1}^{(1)}(\mathbf{Z}_{t+1}) \right| \\
&\leq O(n^{-1}) + O(n^{-1}) \tilde{C}_{n,1}^{\frac{1}{1+\delta}} \sum_{t=1}^n \beta^{\delta/(1+\delta)}(t) = O(n^{-1}), \tag{S17}
\end{aligned}$$

where $\tilde{C}_{n,1} = \max \left\{ \max_{t_1 < t_2} \mathbb{E} \left| g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}_{t_2}) \right|^{1+\delta}, \max_{t_1 < t_2} \mathbb{E} \left| g_{n,1}^{(1)}(\mathbf{Z}_{t_1}) g_{n,1}^{(1)}(\mathbf{Z}'_{t_2}) \right|^{1+\delta} \right\} < \infty$ by (S16).

After using similar arguments in (S15) and by (S14) and $\text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t-h} \right) < \infty$, for any $t_1 < t_2$ and p , we also have

$$\mathbb{E} \left| g_{n,1}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}) \right|^{2(1+\delta)} < \infty, \quad \mathbb{E} \left| g_{n,1}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}) \right|^{2(1+\delta)} < \infty. \tag{S18}$$

Since we have (S18), we apply Lemma 3 and further obtain

$$\mathbb{E} \{ (H_n^{(2)})^2 \} = O(n^{-1}). \tag{S19}$$

Thus, by (S17) and (S19), we obtain

$$\left| \tilde{\mathcal{U}}_{n,1} - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t-h} \right) \right|^2 = O_p(n^{-1}). \tag{S20}$$

Notice that by the similar arguments used in (S15), we have

$$\begin{aligned}
&\mathbb{E} \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\| \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_2} \right\| \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| < \infty \text{ and} \\
&\mathbb{E} \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\| \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_{t_1} \right\| \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| < \infty \text{ which imply that}
\end{aligned}$$

$$\tilde{\mathcal{U}}_{n,2} = \frac{(n-h-2)}{(n-h)} \bar{\mathcal{U}}_{n,2} + O_p(n^{-1}), \tag{S21}$$

where $\bar{U}_{n,2}$ is a U-statistic of order 3 with a varying kernel,

$$g_{n,2}(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t) = \frac{1}{3!} \sum_{(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}'_t \right) \left| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}''_{t-h}) \right| \quad (\text{S22})$$

By applying Hoeffding decomposition, we have

$$\bar{U}_{n,2} = \bar{H}_n^{(1)} + \bar{H}_n^{(2)} + \bar{H}_n^{(3)}, \quad (\text{S23})$$

where $\bar{H}_n^{(1)} = \frac{3}{(n-h)} \sum_t g_{n,2}^{(1)}(\mathbf{Z}_t)$, $\bar{H}_n^{(2)} = \frac{3}{(n-h)(n-h-1)} \sum_{t_1 \neq t_2} g_{n,2}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})$,
 $\bar{H}_n^{(3)} = \frac{1}{(n-h)(n-h-1)(n-h-2)} \sum_{t_1 \neq t_2 \neq t_3} g_{n,2}^{(3)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})$, $g_{n,2}^{(1)}(z) = \mathbb{E} \{g_{n,2}(z, \mathbf{Z}', \mathbf{Z}'')\}$, $g_{n,2}^{(2)}(z, z') = \mathbb{E} \{g_{n,2}(z, z', \mathbf{Z}'')\} - g_{n,2}^{(1)}(z) - g_{n,2}^{(1)}(z')$, $g_{n,2}^{(3)}(z, z', z'') = g_{n,2}(z, z', z'') - g_{n,2}^{(2)}(z, z') - g_{n,2}^{(2)}(z, z'') - g_{n,2}^{(2)}(z', z'') - g_{n,2}^{(1)}(z) - g_{n,2}^{(1)}(z') - g_{n,2}^{(1)}(z'')$.

Similarly, by Cauchy Schwarz inequality, for any $t_1 < t_2 < t_3$ and p , we have

$$\mathbb{E} |g_{n,2}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})|^{2+2\delta} < c \cdot C < \infty, \quad \mathbb{E} |g_{n,2}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3})|^{2+2\delta} < c \cdot C < \infty \quad (\text{S24})$$

under the fourth condition in Assumption 1. This further implies that similar moment conditions in (S18) hold for $g_{n,2}^{(2)}$ and similar moment conditions in (S16) hold for $g_{n,2}^{(1)}$. Therefore, by the similar arguments in (S17) and Lemma 3, we obtain

$$\mathbb{E} \left\{ (\bar{H}_n^{(1)})^2 \right\} = O(n^{-1}), \quad \mathbb{E} \left\{ (\bar{H}_n^{(2)})^2 \right\} = O(n^{-1}). \quad (\text{S25})$$

By (S24), we further have

$$\mathbb{E} \left| g_{n,2}^{(3)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}) \right|^{2(1+\delta)} < \infty, \quad \mathbb{E} \left| g_{n,2}^{(3)}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3}) \right|^{2(1+\delta)} < \infty.$$

Therefore, we apply Lemma 3 to $g_{n,2}^{(3)}$ and obtain

$$\mathbb{E} \left\{ (\bar{H}_n^{(3)})^2 \right\} = O(n^{-1}). \quad (\text{S26})$$

Thus, we have $\bar{U}_{n,2} = O_p(n^{-1/2})$ which implies that $\tilde{U}_{n,2} = O_p(n^{-1/2})$.

By the similar arguments, we have $\tilde{\mathcal{U}}_{n,j} = O_p(n^{-1/2})$, $j = 3, 4$. Hence, we obtain (S12) for any $\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}$, i.e.,

$$\left| \widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) \right| = O_p(n^{-1/2}).$$

Also, for any $\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}$, we have

$$\begin{aligned} & \left| \widehat{\mathcal{S}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\tilde{\mathbf{A}}) \right| \\ & \leq \frac{p\sqrt{p}}{p} \sum_{h=0}^{h_0} \sum_{i=1}^p \left| \widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{t-h} \right) \right| \\ & = O_p(p\sqrt{p} \cdot n^{-1/2}). \end{aligned} \tag{S27}$$

This further implies that

$$\sup_{\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}} \left| \widehat{\mathcal{S}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\tilde{\mathbf{A}}) \right| \xrightarrow{p} 0 \tag{S28}$$

under the assumption that $p\sqrt{p} \cdot n^{-1/2} \rightarrow 0$ as $n \rightarrow \infty$.

Lastly, since we have Lemma 1 and third condition in Assumption 1, our estimate $\widehat{\mathbf{A}}$ that minimizes our objective function $\widehat{\mathcal{S}}_{h_0}$ is well defined and measurable by Theorem 2.2 in [White & Wooldridge \(1991\)](#). Hence, by Lemma 1, (S28), and the third condition in Assumption 1, we apply Proposition 2.4, Corollary 2.3 in [White & Wooldridge \(1991\)](#) and obtain

$$\mathcal{D}(\widehat{\mathbf{A}}, \mathbf{A}) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty.$$

This completes the proof of Theorem 1. ◇

We shall introduce Lemma 4 which is used to show Theorem 2.

Lemma 4. *Let h_0 and m be fixed. For any $\mathbf{U}, \mathbf{V} \in \mathcal{H}_{\mathcal{D}}$, it holds that*

$$|\mathcal{G}_{h_0}(\mathbf{U}) - \mathcal{G}_{h_0}(\mathbf{V})| \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}),$$

where c is a positive constant, and

$$\left| \widehat{\mathcal{G}}_{h_0}(\mathbf{U}) - \widehat{\mathcal{G}}_{h_0}(\mathbf{V}) \right| \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}),$$

where $c = O_p(1)$.

Proof of Lemma 4. Let $(\mathbf{U}_1, \dots, \mathbf{U}_m)$ and $(\mathbf{V}_1, \dots, \mathbf{V}_m)$ form orthogonal matrices, where $\mathbf{U}_i = \{(\mathbf{U}_i)_{\cdot 1}, \dots, (\mathbf{U}_i)_{\cdot p_i}\} \in \mathbb{R}^{p \times p_i}$, $\mathbf{V}_i = \{(\mathbf{V}_i)_{\cdot 1}, \dots, (\mathbf{V}_i)_{\cdot p_i}\} \in \mathbb{R}^{p \times p_i}$ for any i . By arranging the orders, we assume that $\mathcal{D}^2(\mathbf{U}_i, \mathbf{V}_i) = \sum_{i=1}^m \{p_i - \text{tr}(\mathbf{U}_i \mathbf{U}_i^\top \mathbf{V}_i \mathbf{V}_i^\top)\}$. Recall that

$$\begin{aligned} \mathcal{G}_{h_0}(\mathbf{U}) &= \left(\frac{1}{\text{tr}\{\mathbf{U}\mathbf{U}^\top\}} \right)^{1/2} \sum_{h=0}^{h_0} \sum_{i=1}^m \sum_{j \neq i} -\mathbb{E} \left[\mathbf{Y}_t^\top \mathbf{U}_i \mathbf{U}_i^\top \mathbf{Y}'_t \left\| \mathbf{U}_j^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right\| \right] \\ &= \left(\frac{1}{\text{tr}\{\mathbf{U}\mathbf{U}^\top\}} \right)^{1/2} \sum_{h=0}^{h_0} \sum_{i=1}^m \sum_{j \neq i} -\mathbb{E} \left[\mathbf{Y}_t^\top \mathbf{U}_i \mathbf{U}_i^\top \mathbf{Y}'_t \left\| \mathbf{U}_j \mathbf{U}_j^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right\| \right] \\ &= \left(\frac{1}{\text{tr}\{\mathbf{U}\mathbf{U}^\top\}} \right)^{1/2} \sum_{h=0}^{h_0} \sum_{i=1}^m -\mathbb{E} \left[\mathbf{Y}_t^\top (\mathbf{I} - \mathbf{U}_i \mathbf{U}_i^\top) \mathbf{Y}'_t \left\| \mathbf{U}_i \mathbf{U}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right\| \right] \end{aligned}$$

due to the fact that $\mathbf{U}_i^\top \mathbf{U}_i = \mathbf{I}$. By Cauchy Schwarz inequality and the fact that $\|\mathbf{U}_i \mathbf{U}_i^\top -$

$\mathbf{V}_i \mathbf{V}_i^\top \|_F^2 = 2\{p_i - \text{tr}(\mathbf{U}_i \mathbf{U}_i^\top \mathbf{V}_i \mathbf{V}_i^\top)\}$, $\mathbb{E} \|\mathbf{V}_i \mathbf{V}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h})\|^2 \leq c \cdot p_i$, we have

$$\begin{aligned}
& |\mathcal{G}_{h_0}(\mathbf{U}) - \mathcal{G}_{h_0}(\mathbf{V})| \\
& \leq \frac{1}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m |-\mathbb{E} [\mathbf{Y}_t^\top (\mathbf{I} - \mathbf{U}_i \mathbf{U}_i^\top) \mathbf{Y}'_t \{ \|\mathbf{U}_i \mathbf{U}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h})\| - \|\mathbf{V}_i \mathbf{V}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h})\| \}]| \\
& + \frac{1}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m |-\mathbb{E} [\mathbf{Y}_t^\top (\mathbf{V}_i \mathbf{V}_i^\top - \mathbf{U}_i \mathbf{U}_i^\top) \mathbf{Y}'_t \|\mathbf{V}_i \mathbf{V}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h})\|]| \\
& \leq \frac{1}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m \{ \text{tr}(\mathbf{I} - \mathbf{U}_i \mathbf{U}_i^\top) \}^{1/2} \left(\mathbb{E} \left[\left\| \{ \mathbf{U}_i \mathbf{U}_i^\top - \mathbf{V}_i \mathbf{V}_i^\top \}^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right\|^2 \right] \right)^{1/2} \\
& + \frac{1}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m \left(\mathbb{E} \left[\{ \mathbf{Y}_t^\top (\mathbf{U}_i \mathbf{U}_i^\top - \mathbf{V}_i \mathbf{V}_i^\top) \mathbf{Y}'_t \}^2 \right] \right)^{1/2} \left(\mathbb{E} \|\mathbf{V}_i \mathbf{V}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h})\|^2 \right)^{1/2} \\
& \leq c \cdot \frac{1}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m \sqrt{p} \|\mathbf{U}_i \mathbf{U}_i^\top - \mathbf{V}_i \mathbf{V}_i^\top\|_F \\
& \leq c \cdot \frac{\sqrt{p}}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sqrt{m} \left(\sum_{i=1}^m \{ p_i - \text{tr}(\mathbf{U}_i \mathbf{U}_i^\top \mathbf{V}_i \mathbf{V}_i^\top) \} \right)^{1/2} \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V}).
\end{aligned}$$

By using the similar arguments, we can show $|\widehat{\mathcal{G}}_{h_0}(\mathbf{U}) - \widehat{\mathcal{G}}_{h_0}(\mathbf{V})| \leq c \cdot \mathcal{D}(\mathbf{U}, \mathbf{V})$. This completes the proof of Lemma 4. \diamond

Proof of Theorem 2. We shall show the first assertion in Theorem 2 which is the case when p is fixed. Recall that the objective function is

$$\mathcal{G}_{h_0}(\widetilde{\mathbf{A}}) = \left(\frac{1}{p \cdot m} \right)^{1/2} \sum_{h=0}^{h_0} \sum_{i=1}^m \sum_{j \neq i}^m \text{MDD}^2 \left(\widetilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \widetilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h} \right),$$

where $\mathbf{A} = (\widetilde{\mathbf{A}}_1, \dots, \widetilde{\mathbf{A}}_m) \in \mathcal{H}_{\mathcal{D}}$. Notice that the similar decomposition in (S8) remains valid for $\widehat{\text{MDD}}^2 \left(\widetilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \widetilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h} \right)$, i.e.,

$$\widehat{\text{MDD}}^2 \left(\widetilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \widetilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h} \right) = \frac{(n-h-1)}{(n-h)} \{ U_{n,1} + U_{n,2} + U_{n,3} + U_{n,4} \},$$

where

$$\begin{aligned}
U_{n,1} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1})^\top (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_2}) \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|, \\
U_{n,2} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1})^\top (-\tilde{\mathbf{A}}_i^\top \bar{\mathbf{Y}}) \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|, \\
U_{n,3} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (-\tilde{\mathbf{A}}_i^\top \bar{\mathbf{Y}})^\top (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_2}) \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|, \\
U_{n,4} &= -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} (\tilde{\mathbf{A}}_i^\top \bar{\mathbf{Y}})^\top (\tilde{\mathbf{A}}_i^\top \bar{\mathbf{Y}}) \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|,
\end{aligned}$$

and $U_{n,1}$ is a U-statistic of order 2 with the kernel,

$$\tilde{g}_1(\mathbf{Z}_t, \mathbf{Z}'_t) = -\frac{1}{2} \left\{ (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t)^\top (\tilde{\mathbf{A}}_i^\top \mathbf{Y}'_t) + (\tilde{\mathbf{A}}_i^\top \mathbf{Y}'_t)^\top (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t) \right\} \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right\|.$$

Under the second condition in Assumption 1 and by the same arguments in (S9), we have $\mathbb{E}|\tilde{g}_1(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_1+k})|^{2+\delta} < \infty$ for $k = 1, 2, \dots$ and $\mathbb{E}|\tilde{g}_1(\mathbf{Z}_t, \mathbf{Z}'_t)|^{2+\delta} < \infty$. Therefore, after applying Theorem 1 in Yoshihara (1976), we obtain

$$\left| U_{n,1} - \text{MDD}^2(\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h}) \right|^2 = O_p(n^{-1}).$$

Similar to (S11) and by $\mathbb{E}\|\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1}\| \cdot \|\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1}\| \cdot \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| < \infty$, $\mathbb{E}\|\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1}\| \cdot \|\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_2}\| \cdot \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| < \infty$, we have

$$\begin{aligned}
U_{n,2} &= \frac{1}{(n-h)^2(n-h-1)} \sum_{t_1 \neq t_2 \neq t_3} (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1})^\top (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_3}) \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| \\
&+ \frac{1}{(n-h)^2(n-h-1)} \sum_{t_1 \neq t_2} (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1})^\top \left\{ (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_1}) + (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t_2}) \right\} \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\| \\
&= \frac{(n-h-2)}{(n-h)} \bar{U}_{n,2} + O_p(n^{-1}),
\end{aligned}$$

where $\bar{U}_{n,2}$ is a U-statistic of order 3 with the kernel,

$$\tilde{g}_2(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t) = \frac{1}{3!} \sum_{(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)} (\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t)^\top (\tilde{\mathbf{A}}_i^\top \mathbf{Y}'_t) \left\| \tilde{\mathbf{A}}_j^\top (\mathbf{Y}_{t-h} - \mathbf{Y}''_{t-h}) \right\|.$$

Note that we further have $\mathbb{E}|\tilde{g}_2(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_1+k_1}, \mathbf{Z}_{t_1+k_1+k_2})|^{2+\delta} < \infty$ for $k_1, k_2 = 1, 2, \dots$ and $\mathbb{E}|\tilde{g}_2(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)|^{2+\delta} < \infty$ under the second condition in Assumption 1. Hence by Theorem 1 in [Yoshihara \(1976\)](#), we obtain $\bar{U}_{n,2} = O_p(n^{-1/2})$ which implies that $U_{n,2} = O_p(n^{-1/2})$. Similarly, we have $U_{n,j} = O_p(n^{-1/2})$, $j = 3, 4$ which implies that

$$\left| \widehat{\text{MDD}}^2 \left(\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h} \right) \right| = O_p(n^{-1/2})$$

and

$$\left| \widehat{\mathcal{G}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{G}_{h_0}(\tilde{\mathbf{A}}) \right| = O_p(n^{-1/2}) \text{ for any } \tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}. \quad (\text{S29})$$

This further implies that

$$\sup_{\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}} \left| \widehat{\mathcal{G}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{G}_{h_0}(\tilde{\mathbf{A}}) \right| \rightarrow^p 0.$$

By the result above, [Lemma 4](#), and the argmax mapping theorem ([Theorem 3.2.2](#) and [Corollary 3.2.3](#)) in [van der Vaart and Wellner \(1996\)](#), we have

$$\mathcal{D}(\widehat{\mathbf{A}}, \mathbf{A}) \rightarrow^p 0 \text{ as } n \rightarrow \infty.$$

Next, we shall show the second assertion in [Theorem 2](#) which is the case when p diverges. Note that the objective function is

$$\begin{aligned} \mathcal{G}_{h_0}(\tilde{\mathbf{A}}) &= \left(\frac{1}{\text{tr} \left\{ \tilde{\mathbf{A}} \tilde{\mathbf{A}}^\top \right\} m} \right)^{1/2} \sum_{h=0}^{h_0} \sum_{i=1}^m \sum_{j \neq i} \text{MDD}^2 \left(\tilde{\mathbf{A}}_i^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_j^\top \mathbf{Y}_{t-h} \right) \\ &= \frac{p\sqrt{p}}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right), \end{aligned}$$

where $\tilde{\mathbf{A}}_{-i} = (\tilde{\mathbf{A}}_1, \dots, \tilde{\mathbf{A}}_{i-1}, \tilde{\mathbf{A}}_{i+1}, \dots, \tilde{\mathbf{A}}_m) \in \mathbb{R}^{p \times (p-p_i)}$, and similar decomposition in [\(S13\)](#) remains valid for $\widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right)$, i.e.,

$$\widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right) = \frac{(n-h-1)}{(n-h)} \left\{ \tilde{U}_{n,1} + \tilde{U}_{n,2} + \tilde{U}_{n,3} + \tilde{U}_{n,4} \right\},$$

where

$$\tilde{U}_{n,1} = -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_{t_1} \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_{t_2} \right) \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|,$$

$$\tilde{U}_{n,2} = -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_{t_1} \right)^\top \left(-\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \bar{\mathbf{Y}} \right) \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|,$$

$$\tilde{U}_{n,3} = -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(-\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \bar{\mathbf{Y}} \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_{t_2} \right) \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|,$$

$$\tilde{U}_{n,4} = -\frac{1}{(n-h)(n-h-1)} \sum_{t_1} \sum_{t_2 \neq t_1} \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \bar{\mathbf{Y}} \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \bar{\mathbf{Y}} \right) \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top (\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h}) \right\|,$$

and $\tilde{U}_{n,1}$ is a U-statistic of order 2 with a varying kernel which depends on n .

$$\begin{aligned} \tilde{g}_{n,1}(\mathbf{Z}_t, \mathbf{Z}'_t) &= -\frac{1}{2} \left\{ \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}'_t \right) + \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}'_t \right)^\top \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \right) \right\} \\ &\quad \times \left\| \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top (\mathbf{Y}_{t-h} - \mathbf{Y}'_{t-h}) \right\|. \end{aligned}$$

By the similar arguments in (S15), for any $t_1 < t_2$ and p , we have

$$\mathbb{E} |\tilde{g}_{n,1}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})|^{2+2\delta} < c \cdot C < \infty, \quad \mathbb{E} |\tilde{g}_{n,1}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2})|^{2+2\delta} < c \cdot C < \infty \quad (\text{S30})$$

under the fourth condition in Assumption 1. By the Hoeffding decomposition, we further have

$$\tilde{U}_{n,1} = \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right) + \tilde{H}_n^{(1)} + \tilde{H}_n^{(2)},$$

where $\tilde{H}_n^{(1)} = \frac{2}{n-h} \sum_t \tilde{g}_{n,1}^{(1)}(\mathbf{Z}_t)$, $\tilde{H}_n^{(2)} = \frac{1}{(n-h)(n-h-1)} \sum_{t_1 \neq t_2} \tilde{g}_{n,1}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})$, $\tilde{g}_{n,1}^{(1)}(z) = \mathbb{E} \{ \tilde{g}_{n,1}(z, \mathbf{Z}') \} - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right)$, $\tilde{g}_{n,1}^{(2)}(z, z') = \tilde{g}_{n,1}(z, z') - \tilde{g}_{n,1}^{(1)}(z) - \tilde{g}_{n,1}^{(1)}(z')$
 $- \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right)$. Note that for any $t_1 \neq t_2$, we also have $\mathbb{E} \left\{ \tilde{g}_{n,1}^{(1)}(\mathbf{Z}_t)^2 \right\} < \infty$, $\mathbb{E} \left| \tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t_1}) \tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t_2}) \right|^{1+\delta} < \infty$, and $\mathbb{E} \left| \tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t_1}) \tilde{g}_{n,1}^{(1)}(\mathbf{Z}'_{t_2}) \right|^{1+\delta} < \infty$. By Lemma A.1. in

Dette & Spreckelsen (2004) and $\mathbb{E}\tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t_1})\tilde{g}_{n,1}^{(1)}(\mathbf{Z}'_{t_2}) = 0$, we obtain

$$\begin{aligned}
\mathbb{E}\left\{(\tilde{H}_n^{(1)})^2\right\} &= O(n^{-2})\sum_{t_1,t_2}\mathbb{E}\left\{\tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t_1})\tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t_2})\right\} \\
&\leq O(n^{-1})\mathbb{E}\left\{\tilde{g}_{n,1}^{(1)}(\mathbf{Z}_t)^2\right\} + O(n^{-2})\sum_{t=1}^n(n-t)\mathbb{E}\left\{\tilde{g}_{n,1}^{(1)}(\mathbf{Z}_1)\tilde{g}_{n,1}^{(1)}(\mathbf{Z}_{t+1})\right\} \\
&\leq O(n^{-1}) + O(n^{-1})\sum_{t=1}^n\beta^{\delta/(1+\delta)}(t) = O(n^{-1}). \tag{S31}
\end{aligned}$$

By (S30) and $\text{MDD}^2\left(\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_t \mid \frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_{t-h}\right) < \infty$, for any $t_1 < t_2$ and p , we also have

$$\mathbb{E}\left|\tilde{g}_{n,1}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})\right|^{2(1+\delta)} < \infty, \quad \mathbb{E}\left|\tilde{g}_{n,1}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2})\right|^{2(1+\delta)} < \infty.$$

Thus, we apply Lemma 3 to $\tilde{g}_{n,1}^{(2)}$ and obtain

$$\mathbb{E}\left\{(\tilde{H}_n^{(2)})^2\right\} = O(n^{-1}). \tag{S32}$$

Hence we have

$$\left|\tilde{U}_{n,1} - \text{MDD}^2\left(\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_t \mid \frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_{t-h}\right)\right|^2 = O_p(n^{-1}).$$

By the similar arguments used in (S15), we have $\mathbb{E}\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_{t_1}\right\|\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_{t_2}\right\|\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top(\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h})\right\| < \infty$ and $\mathbb{E}\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_{t_1}\right\|\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_{t_1}\right\|\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top(\mathbf{Y}_{t_1-h} - \mathbf{Y}_{t_2-h})\right\| < \infty$ which imply that

$$\tilde{U}_{n,2} = \frac{(n-h-2)}{(n-h)}\bar{U}_{n,2} + O_p(n^{-1}),$$

where $\bar{U}_{n,2}$ is a U-statistic of order 3 with a varying kernel,

$$\tilde{g}_{n,2}(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t) = \frac{1}{3!}\sum_{(\mathbf{Z}_t, \mathbf{Z}'_t, \mathbf{Z}''_t)}\left(\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}_t\right)^\top\left(\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top\mathbf{Y}'_t\right)\left\|\frac{1}{\sqrt{p}}\tilde{\mathbf{A}}_{-i}^\top(\mathbf{Y}_{t-h} - \mathbf{Y}''_{t-h})\right\|.$$

By Hoeffding decomposition, we obtain

$$\bar{U}_{n,2} = \check{H}_n^{(1)} + \check{H}_n^{(2)} + \check{H}_n^{(3)},$$

where $\check{H}_n^{(1)} = \frac{3}{(n-h)} \sum_t \tilde{g}_{n,2}^{(1)}(\mathbf{Z}_t)$, $\check{H}_n^{(2)} = \frac{3}{(n-h)(n-h-1)} \sum_{t_1 \neq t_2} \tilde{g}_{n,2}^{(2)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2})$,
 $\check{H}_n^{(3)} = \frac{1}{(n-h)(n-h-1)(n-h-2)} \sum_{t_1 \neq t_2 \neq t_3} \tilde{g}_{n,2}^{(3)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})$, $\tilde{g}_{n,2}^{(1)}(z) = \mathbb{E} \{ \tilde{g}_{n,2}(z, \mathbf{Z}', \mathbf{Z}'') \}$, $\tilde{g}_{n,2}^{(2)}(z, z') =$
 $\mathbb{E} \{ \tilde{g}_{n,2}(z, z', \mathbf{Z}'') \} - \tilde{g}_{n,2}^{(1)}(z) - \tilde{g}_{n,2}^{(1)}(z')$, $\tilde{g}_{n,2}^{(3)}(z, z', z'') = \tilde{g}_{n,2}(z, z', z'') - \tilde{g}_{n,2}^{(2)}(z, z') - \tilde{g}_{n,2}^{(2)}(z, z'') -$
 $\tilde{g}_{n,2}^{(2)}(z', z'') - \tilde{g}_{n,2}^{(1)}(z) - \tilde{g}_{n,2}^{(1)}(z') - \tilde{g}_{n,2}^{(1)}(z'')$.

Similarly, by Cauchy Schwarz inequality, for any $t_1 < t_2 < t_3$ and p , we have

$$\mathbb{E} |\tilde{g}_{n,2}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3})|^{2+2\delta} < c \cdot C < \infty, \quad \mathbb{E} |\tilde{g}_{n,2}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3})|^{2+2\delta} < c \cdot C < \infty \quad (\text{S33})$$

under the fourth condition in Assumption 1. Hence, by using similar arguments in (S31) and (S32), we obtain

$$\mathbb{E} \{ (\check{H}_n^{(1)})^2 \} = O(n^{-1}), \quad \mathbb{E} \{ (\check{H}_n^{(2)})^2 \} = O(n^{-1}).$$

By (S33), we also have

$$\mathbb{E} \left| \tilde{g}_{n,2}^{(3)}(\mathbf{Z}_{t_1}, \mathbf{Z}_{t_2}, \mathbf{Z}_{t_3}) \right|^{2(1+\delta)} < \infty, \quad \mathbb{E} \left| \tilde{g}_{n,2}^{(3)}(\mathbf{Z}_{t_1}, \mathbf{Z}'_{t_2}, \mathbf{Z}'_{t_3}) \right|^{2(1+\delta)} < \infty.$$

Therefore, we apply Lemma 3 to $\tilde{g}_{n,2}^{(3)}$ and obtain

$$\mathbb{E} \{ (\check{H}_n^{(3)})^2 \} = O(n^{-1}).$$

Thus, we have $\bar{U}_{n,2} = O_p(n^{-1/2})$ which implies that $\tilde{U}_{n,2} = O_p(n^{-1/2})$. Similarly, we can obtain $\tilde{U}_{n,j} = O_p(n^{-1/2})$, $j = 3, 4$ by Lemma 3 and using similar arguments. Therefore, we ultimately have

$$\left| \widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right) \right| = O_p(n^{-1/2}).$$

Hence, for any $\tilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}$, we further have

$$\begin{aligned} & \left| \hat{\mathcal{G}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{G}_{h_0}(\tilde{\mathbf{A}}) \right| \\ & \leq \frac{p\sqrt{p}}{\sqrt{p \cdot m}} \sum_{h=0}^{h_0} \sum_{i=1}^m \left| \widehat{\text{MDD}}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right) - \text{MDD}^2 \left(\frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_{-i}^\top \mathbf{Y}_t \mid \frac{1}{\sqrt{p}} \tilde{\mathbf{A}}_i^\top \mathbf{Y}_{t-h} \right) \right| \\ & \leq O_p(p \cdot \sqrt{m} \cdot n^{-1/2}). \end{aligned} \quad (\text{S34})$$

Under the assumption that $p \cdot \sqrt{m} \cdot n^{-1/2} \rightarrow 0$ as $n \rightarrow \infty$, we obtain

$$\sup_{\tilde{\mathbf{A}} \in \mathcal{H}_{\tilde{\mathcal{D}}}} \left| \widehat{\mathcal{G}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{G}_{h_0}(\tilde{\mathbf{A}}) \right| \rightarrow^p 0. \quad (\text{S35})$$

Finally, since we have Lemma 4 and third condition in Assumption 1, our estimate $\widehat{\mathbf{A}}$ that minimizes our objective function $\widehat{\mathcal{G}}_{h_0}$ is well defined and measurable by Theorem 2.2 in [White & Wooldridge \(1991\)](#). Furthermore, since Lemma 4 and (S35) are satisfied and we have the third condition in Assumption 1, we apply Proposition 2.4, Corollary 2.3 in [White & Wooldridge \(1991\)](#) and obtain

$$\mathcal{D}(\widehat{\mathbf{A}}, \mathbf{A}) \rightarrow^p 0 \text{ as } n \rightarrow \infty.$$

This completes the proof of Theorem 2. ◇

2 Additional Theorem and the Proof

Assumption S1. *When p diverges, for any vector v with $\|v\| = 1$, there exists a constant $C' > 0$ such that $\mathbb{E} \{ \exp(2|v^\top \mathbf{Y}_t|) \} < C'$ for all p . The data $(\mathbf{Y}_t)_{t=1}^n$ is strictly stationary and α -mixing with $\alpha_{k,p} \rightarrow 0$ as $k \rightarrow \infty$, where $\alpha_{k,p} = \sup_i \sup_{A \in \mathcal{F}_{-\infty}^i, B \in \mathcal{F}_{i+k}^\infty} |P(A \cap B) - P(A)P(B)|$, and \mathcal{F}_i^j is the σ -field generated by $\{\mathbf{Y}_t : i \leq t \leq j\}$. For all $k \geq 1$, there exists a constant $C'' > 0$ such that $\sup_{p \geq 1} \alpha_{k,p} \leq \exp(-C'' k^{r_2})$, where $r_2 \in (0, 1]$ is a constant.*

Theorem S1. *Let h_0 be a fixed integer and p grow with n such that $p(\log(p))^{4/\gamma'} / n \rightarrow 0$ as $n \rightarrow \infty$, where $\gamma'^{-1} = 1 + r_2^{-1}$. Under Assumption S1 and the first, third conditions in Assumption 1 in our manuscript, we have $\mathcal{D}(\mathbf{A}, \widehat{\mathbf{A}}) \rightarrow^p 0$ as $n \rightarrow \infty$.*

Proof for Theorem SI. Note that for any $\varepsilon > 0$ and $\tilde{\mathbf{A}}$, we have

$$\begin{aligned} & P\left(\left|\widehat{\mathcal{S}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\tilde{\mathbf{A}})\right| > \varepsilon\right) \\ & \leq \sum_{h=0}^{h_0} \sum_{i=1}^p \sum_{j \neq i} P\left(\left|\widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h}) - \text{MDD}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h})\right| > \frac{\varepsilon}{(p-1)(h_0+1)}\right). \end{aligned} \quad (\text{S36})$$

Next, we shall show that for any $\varepsilon > 0$ and $\tilde{\mathbf{A}}$,

$$\begin{aligned} & P\left(\left|\widehat{\mathcal{S}}_{h_0}(\tilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\tilde{\mathbf{A}})\right| > \varepsilon\right) \\ & \leq c \cdot p^2 n^2 \exp(-c(\varepsilon/p)^\gamma n^{2\gamma}) + c \cdot p^2 n^2 \exp(-c(\varepsilon/p)^{\gamma'/2} n^{2\gamma'}) + c \cdot p^2 n^2 \exp(-c(\varepsilon/p)^{\gamma''/2} n^{2\gamma''}) \\ & + c \cdot p^2 n \exp(-c(\varepsilon/p)^{\gamma'/4} n^{\gamma'}) + c \cdot p^2 n \exp(-c(\varepsilon/p)^{\gamma'/2} n^{\gamma'}) + c \cdot p^2 \exp(-c(\varepsilon/p)^{1/2} n) \\ & + c \cdot p^2 \exp(-c(\varepsilon/p)n) + c \cdot p^2 \exp(-c(\varepsilon/p)n^2) + c \cdot p^2 \exp(-c(\varepsilon/p)^2 n^2) \end{aligned}$$

by following the idea in the proof of [Chang et al. \(2018\)](#).

Recall that $\mathbb{E}(\mathbf{Y}_t) = 0$. Notice that we have

$$\widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h}) = I + II + III,$$

where

$$\begin{aligned} I &= \frac{-1}{n^2} \sum_{k_1 \neq k_2} (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1})^\top (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2}) |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|, \\ II &= \frac{-1}{n^2} \sum_{k_1, k_2} (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1})^\top (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2}) \frac{1}{n^2} \sum_{k_1 \neq k_2} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|, \\ III &= \frac{2}{n^3} \sum_{k_1, k_2, k_3} (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1})^\top (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_3}) |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_3-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|. \end{aligned}$$

We shall first show that for any $s > 0$,

$$P\left(\left|I - \text{MDD}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h})\right| > s\right) \leq c \cdot n^2 \exp(-cs^\gamma n^{2\gamma}) + c \cdot \exp(-cs^2 n^2), \quad (\text{S37})$$

where $\gamma^{-1} = 3 + r_2^{-1}$.

We let

$$I = \frac{-1}{n(n-1)} \sum_{k_1 \neq k_2} (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1})^\top (\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2}) |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|$$

which will not affect the result in (S37). We obtain the following for any $s > 0$.

$$\begin{aligned} & P \left(\left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| - \text{MDD}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h}) \right| > s \right) \\ &= P \left(\left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} \right| > s^\epsilon, \right. \\ & \quad \left. \left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| - \text{MDD}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h}) \right| > s \right) \\ &+ P \left(\left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} \right| \leq s^\epsilon, \right. \\ & \quad \left. \left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| - \text{MDD}^2(\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{t-h}) \right| > s \right) \\ &\leq P \left(\left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} \right| > s^\epsilon \right) \\ &+ P \left(|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| + \mathbb{E} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| > s^{1-\epsilon} \right) \\ &\leq c \cdot \exp(-s^{\epsilon/2}) + c \cdot \exp(-s^{(1-\epsilon)}), \end{aligned}$$

where

$$\begin{aligned} & P \left(\left| -\tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} \right| > s^\epsilon \right) \leq P \left(\left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \right| > s^{\epsilon/2} \right) + P \left(\left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_2} \right| > s^{\epsilon/2} \right) \\ &\leq c \cdot \exp(-s^{\epsilon/2}) \mathbb{E} \left\{ \exp \left(\left| \tilde{\mathbf{A}}_{\cdot i}^\top \mathbf{Y}_{k_1} \right| \right) \right\} = c \cdot \exp(-s^{\epsilon/2}), \end{aligned}$$

and

$$\begin{aligned} & P \left(|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| + \mathbb{E} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| > s^{(1-\epsilon)} \right) \\ &\leq \exp(-s^{(1-\epsilon)}) \mathbb{E} \left\{ \exp(|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}| + \mathbb{E} |\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|) \right\} \\ &\leq c \cdot \exp(-s^{(1-\epsilon)}) \mathbb{E} \left\{ \exp(|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h}|) \cdot \exp(|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|) \right\} \\ &\leq c \cdot \exp(-s^{(1-\epsilon)}) \mathbb{E} \left\{ \exp(2|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_1-h}|) \right\}^{1/2} \mathbb{E} \left\{ \exp(2|\tilde{\mathbf{A}}_{\cdot j}^\top \mathbf{Y}_{k_2-h}|) \right\}^{1/2} \\ &\leq c \cdot \exp(-s^{(1-\epsilon)}), \end{aligned}$$

by Markov inequality and Cauchy Schwarz inequality under Assumption S1. By setting $(1 - \epsilon) = \epsilon/2$, we have $\epsilon = \frac{1}{1+1/2}$. Thus, we obtain

$$\begin{aligned} & P \left(\left| -\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{k_2} \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{k_2-h} \right| - \text{MDD}^2(\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{t-h}) \right| > s \Big) \\ & \leq c \cdot \exp(-s^{1/3}) \end{aligned}$$

We further apply Theorem 1 in [Merlevède et al. \(2011\)](#) and obtain that

$$\begin{aligned} & P \left(\left| I - \text{MDD}^2(\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{t-h}) \right| > s \right) \\ & = P \left(\left| \frac{1}{n(n-1)} \sum_{k_1 \neq k_2} \left\{ -\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{k_1} \tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_{k_2} \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{k_1-h} - \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{k_2-h} \right\} \right. \right. \\ & \quad \left. \left. - \text{MDD}^2(\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{t-h}) \right| > s \right) \\ & \leq c \cdot n^2 \exp(-cs^\gamma n^{2\gamma}) + c \cdot \exp(-cs^2 n^2), \end{aligned}$$

where $\gamma^{-1} = (1/3)^{-1} + r_2^{-1}$.

By using the similar arguments, we have

$$P(|III| > s) \leq c \cdot n \exp(-cs^{\gamma'/4} n^{\gamma'}) + c \cdot n^2 \exp(-cs^{\gamma'/2} n^{2\gamma'}) + c \cdot \exp(-cs^{1/2} n) + c \cdot \exp(-csn^2),$$

$$P(|III| > s) \leq c \cdot n \exp(-cs^{\gamma''/2} n^{\gamma''}) + c \cdot n^2 \exp(-cs^{\gamma''/2} n^{2\gamma''}) + c \cdot \exp(-csn) + c \cdot \exp(-csn^2),$$

where $(\gamma')^{-1} = 1^{-1} + r_2^{-1}$, $\gamma''^{-1} = (1/2)^{-1} + r_2^{-1}$. Hence, for any $s > 0$, we finally have

$$\begin{aligned} & P \left(\left| \widehat{\text{MDD}}^2(\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{t-h}) - \text{MDD}^2(\tilde{\mathbf{A}}_{.i}^\top \mathbf{Y}_t \mid \tilde{\mathbf{A}}_{.j}^\top \mathbf{Y}_{t-h}) \right| > s \right) \\ & \leq c \cdot n^2 \exp(-cs^\gamma n^{2\gamma}) + c \cdot n^2 \exp(-cs^{\gamma'/2} n^{2\gamma'}) + c \cdot n^2 \exp(-cs^{\gamma''/2} n^{2\gamma''}) + c \cdot n \exp(-cs^{\gamma'/4} n^{\gamma'}) \\ & \quad + c \cdot n \exp(-cs^{\gamma'/2} n^{\gamma'}) + c \cdot \exp(-cs^{1/2} n) + c \cdot \exp(-csn) + c \cdot \exp(-csn^2) + c \cdot \exp(-cs^2 n^2). \end{aligned}$$

By (S36), we finally obtain

$$\begin{aligned}
& P \left(\left| \widehat{\mathcal{S}}_{h_0}(\widetilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\widetilde{\mathbf{A}}) \right| > \varepsilon \right) \\
& \leq c \cdot p^2 n^2 \exp(-c(\varepsilon/p)^\gamma n^{2\gamma}) + c \cdot p^2 n^2 \exp(-c(\varepsilon/p)^{\gamma'/2} n^{2\gamma'}) + c \cdot p^2 n^2 \exp(-c(\varepsilon/p)^{\gamma''/2} n^{2\gamma''}) \\
& + c \cdot p^2 n \exp(-c(\varepsilon/p)^{\gamma'/4} n^{\gamma'}) + c \cdot p^2 n \exp(-c(\varepsilon/p)^{\gamma'/2} n^{\gamma'}) + c \cdot p^2 \exp(-c(\varepsilon/p)^{1/2} n) \\
& + c \cdot p^2 \exp(-c(\varepsilon/p)n) + c \cdot p^2 \exp(-c(\varepsilon/p)n^2) + c \cdot p^2 \exp(-c(\varepsilon/p)^2 n^2)
\end{aligned}$$

Thus, for any $\widetilde{\mathbf{A}}$, we have

$$\left| \widehat{\mathcal{S}}_{h_0}(\widetilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\widetilde{\mathbf{A}}) \right| = O_p \left(\frac{p}{n} (\log(p))^{4/\gamma'} \right)$$

which further implies

$$\sup_{\widetilde{\mathbf{A}} \in \mathcal{H}_{\mathcal{D}}} \left| \widehat{\mathcal{S}}_{h_0}(\widetilde{\mathbf{A}}) - \mathcal{S}_{h_0}(\widetilde{\mathbf{A}}) \right| \xrightarrow{p} 0$$

under the assumption that $p(\log(p))^{4/\gamma'}/n \rightarrow 0$ as $n \rightarrow \infty$. By using the same arguments in the proof of Theorem 1, our estimate $\widehat{\mathbf{A}}$ is well defined and measurable, and we obtain

$$\mathcal{D} \left(\widehat{\mathbf{A}}, \mathbf{A} \right) \rightarrow 0.$$

This completes the proof of Theorem S1. ◇

3 Additional Simulations, Real Data Application Results, and Figures

In this section, we report additional simulation results in Section 5 and additional real data application results, figures in Section 6.

Table S1: Reported are the average and the standard deviation of \mathcal{D}^2 -distance with $h_0 = 2$ based on 500 replications. Three methods are compared: our method (MICA), the method in [Matteson & Tsay \(2011, DOC\)](#), the method in [Belouchrani et al. \(1997, SOBI\)](#).

	Dist.	p	Methods	$n = 50$	$n = 100$	$n = 200$
Example I	Normal	5	MICA	0.192 (0.071)	0.165 (0.071)	0.134 (0.068)
			DOC	0.184 (0.070)	0.152 (0.068)	0.116 (0.065)
			SOBI	0.178 (0.060)	0.148 (0.051)	0.114 (0.062)
		10	MICA	0.370 (0.051)	0.343 (0.055)	0.317 (0.055)
			DOC	0.364 (0.052)	0.325 (0.055)	0.279 (0.060)
			SOBI	0.368 (0.050)	0.324 (0.078)	0.279 (0.089)
	t	5	MICA	0.141 (0.070)	0.109 (0.065)	0.078 (0.057)
			DOC	0.208 (0.083)	0.169 (0.078)	0.125 (0.076)
			SOBI	0.220 (0.089)	0.166 (0.083)	0.132 (0.078)
		10	MICA	0.321 (0.060)	0.280 (0.058)	0.237 (0.057)
			DOC	0.401 (0.064)	0.350 (0.071)	0.296 (0.077)
			SOBI	0.404 (0.063)	0.367 (0.073)	0.304 (0.075)
exp	5	MICA	0.161 (0.069)	0.115 (0.070)	0.076 (0.052)	
		DOC	0.232 (0.068)	0.215 (0.069)	0.183 (0.068)	
		SOBI	0.192 (0.076)	0.152 (0.074)	0.112 (0.066)	
	10	MICA	0.357 (0.057)	0.327 (0.057)	0.272 (0.058)	
		DOC	0.383 (0.051)	0.358 (0.053)	0.320 (0.059)	
		SOBI	0.370 (0.050)	0.329 (0.059)	0.282 (0.062)	
Example II	t	5	MICA	0.136 (0.075)	0.087 (0.068)	0.041 (0.036)
			DOC	0.280 (0.091)	0.275 (0.084)	0.270 (0.086)
			SOBI	0.269 (0.085)	0.271 (0.082)	0.262 (0.083)
		10	MICA	0.302 (0.071)	0.197 (0.083)	0.110 (0.064)
			DOC	0.457 (0.067)	0.444 (0.068)	0.435 (0.074)
			SOBI	0.449 (0.052)	0.442 (0.054)	0.441 (0.053)
	exp	5	MICA	0.169 (0.068)	0.079 (0.063)	0.026 (0.027)
			DOC	0.266 (0.083)	0.263 (0.083)	0.257 (0.081)
			SOBI	0.270 (0.084)	0.269 (0.080)	0.261 (0.083)
		10	MICA	0.358 (0.057)	0.270 (0.065)	0.143 (0.059)
			DOC	0.438 (0.055)	0.431 (0.057)	0.424 (0.059)
			SOBI	0.450 (0.050)	0.443 (0.053)	0.445 (0.054)

We compare the simulation results for larger dimension, $p = 20, 50$. For the method [Matteson & Tsay \(2011\)](#), we only report the result for $p = 20$ due to the computational time. For instance, it took more than 24 hours to generate a result for one replicate when $p = 50$. Table S2: Reported are the average and the standard deviation of \mathcal{D}^2 -distance with $h_0 = 1$ based on 500 replications. Three methods are compared: our method (MICA), the method in [Matteson & Tsay \(2011, DOC\)](#), the method in [Belouchrani et al. \(1997, SOBI\)](#).

Dist.	p	Methods	$n = 100$	$n = 200$
Normal	20	MICA	0.556 (0.257)	0.509 (0.250)
		DOC	0.512 (0.223)	0.465 (0.212)
		SOBI	0.529 (0.237)	0.483 (0.208)
	50	MICA	0.775 (0.298)	0.736 (0.268)
		SOBI	0.752 (0.302)	0.724 (0.284)
Example I	20	MICA	0.511 (0.273)	0.482 (0.263)
		DOC	0.551 (0.265)	0.512 (0.239)
		SOBI	0.558 (0.274)	0.513 (0.243)
	50	MICA	0.730 (0.301)	0.709 (0.274)
		SOBI	0.765 (0.328)	0.747 (0.298)
exp	20	MICA	0.501 (0.202)	0.462 (0.197)
		DOC	0.525 (0.224)	0.476 (0.201)
		SOBI	0.531 (0.239)	0.485 (0.210)
	50	MICA	0.729 (0.291)	0.698 (0.277)
		SOBI	0.743 (0.304)	0.722 (0.285)
Example II	20	MICA	0.444 (0.231)	0.344 (0.171)
		DOC	0.615 (0.331)	0.616 (0.334)
		SOBI	0.611 (0.344)	0.610 (0.331)
	50	MICA	0.533 (0.382)	0.485 (0.327)
		SOBI	0.778 (0.493)	0.776 (0.490)
exp	20	MICA	0.455 (0.294)	0.397 (0.256)
		DOC	0.598 (0.287)	0.597 (0.288)
		SOBI	0.610 (0.319)	0.609 (0.329)
	50	MICA	0.563 (0.361)	0.512 (0.302)
		SOBI	0.777 (0.493)	0.778 (0.491)

For the existing method TS-PCA, we only report the results for $p = 6$ when $h_0 = 1$ since the proportion of the correct segmentation π for $p = 12$ is extremely small, e.g., $\pi \leq 0.04$.

Table S3: Reported are the proportions of correct segmentations, π , the average and the standard deviation of \mathcal{D}^2 -distance when estimated groups are correct with $h_0 = 1$ based on 500 replications. Two methods are compared: our method (GMICA), the method in [Chang et al. \(2018, TS-PCA\)](#)

Dist.	p	Method	$n = 200$		$n = 500$		$n = 1000$		
			π	\mathcal{D}^2	π	\mathcal{D}^2	π	$\tilde{\mathcal{D}}^2$	
Example I	Normal	6	GMICA	0.520	0.060 (0.117)	0.674	0.024 (0.075)	0.740	0.009 (0.005)
		TS-PCA	0.136	0.101 (0.052)	0.126	0.093 (0.054)	0.122	0.064 (0.043)	
	12	GMICA	0.164	0.094 (0.169)	0.274	0.053 (0.127)	0.388	0.019 (0.011)	
	t	6	GMICA	0.416	0.068 (0.143)	0.496	0.032 (0.100)	0.730	0.021 (0.089)
		TS-PCA	0.124	0.120 (0.069)	0.184	0.087 (0.057)	0.194	0.063 (0.052)	
	12	GMICA	0.312	0.092 (0.168)	0.420	0.040 (0.148)	0.544	0.038 (0.112)	
Example II	exp	6	GMICA	0.408	0.059 (0.128)	0.588	0.018 (0.056)	0.798	0.006 (0.005)
		TS-PCA	0.134	0.127 (0.057)	0.140	0.093 (0.080)	0.122	0.062 (0.044)	
	12	GMICA	0.320	0.095 (0.142)	0.420	0.039 (0.081)	0.718	0.024 (0.028)	
	t	6	GMICA	0.546	0.040 (0.091)	0.652	0.026 (0.021)	0.750	0.014 (0.044)
TS-PCA		0.062	0.254 (0.069)	0.062	0.244 (0.092)	0.054	0.227 (0.079)		
12	GMICA	0.214	0.125 (0.081)	0.374	0.064 (0.074)	0.542	0.073 (0.053)		
Example II	exp	6	GMICA	0.664	0.045 (0.053)	0.744	0.034 (0.032)	0.810	0.026 (0.025)
		TS-PCA	0.036	0.287 (0.077)	0.028	0.276 (0.106)	0.016	0.321 (0.089)	
	12	GMICA	0.302	0.128 (0.082)	0.492	0.093 (0.053)	0.722	0.054 (0.042)	

We compare the simulation results for larger dimension, $p = 24$. We only report the results for our method since the proportion of the correct segmentation π for TS-PCA is

extremely small, e.g., $\pi \leq 0.04$.

Table S4: Reported are the proportions of correct segmentations, π , the average and the standard deviation of \mathcal{D}^2 -distance when estimated groups are correct with $h_0 = 5$ based on 500 replications.

	Dist.	p	Method	$n = 500$		$n = 1000$	
				π	\mathcal{D}^2	π	\mathcal{D}^2
Example I	Normal	24	GMICA	0.222	0.074 (0.018)	0.501	0.043 (0.014)
	t	24	GMICA	0.292	0.085 (0.010)	0.462	0.056 (0.008)
	exp	24	GMICA	0.320	0.073 (0.041)	0.500	0.045 (0.047)
Example II	t	24	GMICA	0.279	0.066 (0.039)	0.433	0.049 (0.035)
	exp	24	GMICA	0.325	0.079 (0.042)	0.461	0.061 (0.032)

We further compare the finite sample performance between our approach with the objective function \mathcal{S}_{h_0} and a variant objective function by aggregating MDD jointly. In particular, the joint approach shall construct the objective function by adding $\text{MDD}(\mathbf{A}_{\cdot i}^\top \mathbf{Y}_t \mid \mathbf{A}_{\cdot j}^\top \mathbf{Y}_t, \mathbf{A}_{\cdot j}^\top \mathbf{Y}_{t-1}, \dots, \mathbf{A}_{\cdot j}^\top \mathbf{Y}_{t-h_0})$ for each component. It turns out that the joint approach is very comparable for Example I but inferior to the ones generated by the current approach for Example II. The results are summarized in Table S5.

In Section 6, the community mobility data captures movements across 6 categories of locations: retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential areas. We compare prediction performance by considering each category separately.

Table S5: Reported are the average and the standard deviation of \mathcal{D}^2 -distance with $h_0 = 1$ based on 500 replications. Two methods are compared: MICA and joint approach of MICA (joint MICA).

	Dist.	p	Methods	$n = 50$	$n = 100$	$n = 200$	
Example I	Normal	5	MICA	0.193 (0.069)	0.172 (0.071)	0.147 (0.074)	
			joint MICA	0.193 (0.068)	0.173 (0.071)	0.148 (0.075)	
		10	MICA	0.374 (0.050)	0.349 (0.054)	0.323 (0.055)	
			joint MICA	0.380 (0.051)	0.351 (0.055)	0.335 (0.025)	
		t	5	MICA	0.140 (0.068)	0.102 (0.061)	0.075 (0.056)
				joint MICA	0.141 (0.068)	0.106 (0.060)	0.077 (0.056)
	10		MICA	0.317 (0.062)	0.269 (0.059)	0.231 (0.055)	
			joint MICA	0.322 (0.065)	0.276 (0.059)	0.237 (0.055)	
	exp		5	MICA	0.153 (0.071)	0.109 (0.065)	0.067 (0.050)
				joint MICA	0.153 (0.070)	0.110 (0.064)	0.068 (0.050)
		10	MICA	0.350 (0.058)	0.314 (0.058)	0.268 (0.056)	
			joint MICA	0.355 (0.059)	0.318 (0.059)	0.273 (0.058)	
Example II		t	5	MICA	0.137 (0.076)	0.075 (0.058)	0.037 (0.037)
				joint MICA	0.137 (0.077)	0.076 (0.057)	0.039 (0.038)
	10		MICA	0.292 (0.072)	0.188 (0.075)	0.099 (0.057)	
			joint MICA	0.292 (0.072)	0.192 (0.075)	0.105 (0.060)	
	exp		5	MICA	0.161 (0.072)	0.074 (0.052)	0.021 (0.024)
				joint MICA	0.160 (0.072)	0.076 (0.052)	0.026 (0.020)
		10	MICA	0.349 (0.060)	0.261 (0.064)	0.132 (0.052)	
			joint MICA	0.354 (0.061)	0.264 (0.065)	0.139 (0.055)	

Table S6: Average MSPE for community mobility data in retail and recreation category.

Method		$h_0 = 7$	$h_0 = 14$
	Groups	7 Groups (2,1,1,1,1,1,1)	7 Groups (2,1,1,1,1,1,1)
GMICA	MSPE ($q = 1$)	1.016	1.033
	MSPE ($q = 2$)	0.990	0.996
	Groups	8 Groups	8 Groups
MICA	MSPE ($q = 1$)	1.038	1.009
	MSPE ($q = 2$)	1.015	1.005
	Groups	7 Groups (2,1,1,1,1,1,1)	7 Groups (2,1,1,1,1,1,1)
TS-PCA	MSPE ($q = 1$)	1.296	1.180
	MSPE ($q = 2$)	1.191	1.119
	Groups	8 Groups	8 Groups
DOC	MSPE ($q = 1$)	1.264	1.259
	MSPE ($q = 2$)	1.192	1.187
	Groups	8 Groups	8 Groups
SOBI	MSPE ($q = 1$)	1.253	1.231
	MSPE ($q = 2$)	1.183	1.172
	Groups	8 Groups	8 Groups
sUARMA	MSPE ($q = 1$)	1.618	1.618
	MSPE ($q = 2$)	1.152	1.152
	Groups	1 Group	1 Group
sVARMA	MSPE ($q = 1$)	1.178	1.178
	MSPE ($q = 2$)	1.192	1.192
	Groups	1 Group	1 Group
refVAR	MSPE ($q = 1$)	1.612	1.612
	MSPE ($q = 2$)	2.552	2.552

Table S7: Average MSPE for community mobility data in grocery and pharmacy category.

Method		$h_0 = 7$	$h_0 = 14$
	Groups	6 Groups (2,2,1,1,1,1)	6 Groups (2,2,1,1,1,1)
GMICA	MSPE ($q = 1$)	1.026	0.999
	MSPE ($q = 2$)	0.985	0.985
	Groups	8 Groups	8 Groups
MICA	MSPE ($q = 1$)	1.045	1.048
	MSPE ($q = 2$)	1.039	1.048
	Groups	7 Groups (2,1,1,1,1,1,1)	6 Groups (3,1,1,1,1,1)
TS-PCA	MSPE ($q = 1$)	1.119	1.339
	MSPE ($q = 2$)	1.091	1.269
	Groups	8 Groups	8 Groups
DOC	MSPE ($q = 1$)	1.213	1.267
	MSPE ($q = 2$)	1.121	1.244
	Groups	8 Groups	8 Groups
SOBI	MSPE ($q = 1$)	1.202	1.324
	MSPE ($q = 2$)	1.115	1.298
	Groups	8 Groups	8 Groups
sUARMA	MSPE ($q = 1$)	1.957	1.957
	MSPE ($q = 2$)	1.156	1.156
	Groups	1 Group	1 Group
sVARMA	MSPE ($q = 1$)	1.078	1.078
	MSPE ($q = 2$)	1.178	1.178
	Groups	1 Group	1 Group
refVAR	MSPE ($q = 1$)	1.755	1.755
	MSPE ($q = 2$)	2.845	2.845

Table S8: Average MSPE for community mobility data in park category.

Method		$h_0 = 7$	$h_0 = 14$
	Groups	7 Groups (2,1,1,1,1,1)	7 Groups (2,1,1,1,1,1)
GMICA	MSPE ($q = 1$)	0.975	0.968
	MSPE ($q = 2$)	0.901	0.906
	Groups	8 Groups	8 Groups
MICA	MSPE ($q = 1$)	1.018	1.018
	MSPE ($q = 2$)	0.988	0.989
	Groups	5 Groups (4,1,1,1,1)	5 Groups (4,1,1,1,1)
TS-PCA	MSPE ($q = 1$)	1.013	1.018
	MSPE ($q = 2$)	0.996	0.966
	Groups	8 Groups	8 Groups
DOC	MSPE ($q = 1$)	1.025	1.027
	MSPE ($q = 2$)	0.976	0.972
	Groups	8 Groups	8 Groups
SOBI	MSPE ($q = 1$)	1.013	1.011
	MSPE ($q = 2$)	0.984	0.983
	Groups	8 Groups	8 Groups
sUARMA	MSPE ($q = 1$)	1.518	1.518
	MSPE ($q = 2$)	1.056	1.056
	Groups	1 Group	1 Group
sVARMA	MSPE ($q = 1$)	0.982	0.982
	MSPE ($q = 2$)	2.523	2.523
	Groups	1 Group	1 Group
refVAR	MSPE ($q = 1$)	1.412	1.412
	MSPE ($q = 2$)	0.957	0.957

Table S9: Average MSPE for community mobility data in transit category.

Method		$h_0 = 7$	$h_0 = 14$
	Groups	7 Groups (2,1,1,1,1,1,1)	7 Groups (2,1,1,1,1,1,1)
GMICA	MSPE ($q = 1$)	1.071	1.058
	MSPE ($q = 2$)	0.994	0.977
	Groups	8 Groups	8 Groups
MICA	MSPE ($q = 1$)	1.045	1.101
	MSPE ($q = 2$)	0.996	1.000
	Groups	7 Groups (2,1,1,1,1,1,1)	7 Groups (2,1,1,1,1,1,1)
TS-PCA	MSPE ($q = 1$)	1.233	1.227
	MSPE ($q = 2$)	1.109	1.141
	Groups	8 Groups	8 Groups
DOC	MSPE ($q = 1$)	1.179	1.182
	MSPE ($q = 2$)	1.118	1.128
	Groups	8 Groups	8 Groups
SOBI	MSPE ($q = 1$)	1.216	1.320
	MSPE ($q = 2$)	1.092	1.081
	Groups	8 Groups	8 Groups
sUARMA	MSPE ($q = 1$)	1.812	1.812
	MSPE ($q = 2$)	1.159	1.159
	Groups	1 Group	1 Group
sVARMA	MSPE ($q = 1$)	1.649	1.649
	MSPE ($q = 2$)	2.975	2.975
	Groups	1 Group	1 Group
refVAR	MSPE ($q = 1$)	1.193	1.193
	MSPE ($q = 2$)	2.912	2.912

Table S10: Average MSPE for community mobility data in workplace category.

Method		$h_0 = 7$	$h_0 = 14$
	Groups	7 Groups (2,1,1,1,1,1)	7 Groups (2,1,1,1,1,1)
GMICA	MSPE ($q = 1$)	1.081	1.104
	MSPE ($q = 2$)	1.091	1.095
	Groups	8 Groups	8 Groups
MICA	MSPE ($q = 1$)	1.001	1.002
	MSPE ($q = 2$)	1.098	0.997
	Groups	5 Groups (4,1,1,1,1)	6 Groups (3,1,1,1,1,1)
TS-PCA	MSPE ($q = 1$)	1.876	1.792
	MSPE ($q = 2$)	1.933	1.786
	Groups	8 Groups	8 Groups
DOC	MSPE ($q = 1$)	1.235	1.182
	MSPE ($q = 2$)	1.112	1.158
	Groups	8 Groups	8 Groups
SOBI	MSPE ($q = 1$)	1.241	1.237
	MSPE ($q = 2$)	1.143	1.183
	Groups	8 Groups	8 Groups
sUARMA	MSPE ($q = 1$)	1.613	1.613
	MSPE ($q = 2$)	1.034	1.034
	Groups	1 Group	1 Group
sVARMA	MSPE ($q = 1$)	1.178	1.178
	MSPE ($q = 2$)	3.179	3.179
	Groups	1 Group	1 Group
refVAR	MSPE ($q = 1$)	1.593	1.593
	MSPE ($q = 2$)	1.299	1.299

Table S11: Average MSPE for community mobility data in residential areas category.

Method		$h_0 = 7$	$h_0 = 14$
	Groups	7 Groups (2,1,1,1,1,1,1)	7 Groups (2,1,1,1,1,1,1)
GMICA	MSPE ($q = 1$)	0.985	1.009
	MSPE ($q = 2$)	0.974	0.989
	Groups	8 Groups	8 Groups
MICA	MSPE ($q = 1$)	1.012	1.012
	MSPE ($q = 2$)	1.008	1.007
	Groups	7 Groups (2,1,1,1,1,1,1)	7 Groups (2,1,1,1,1,1,1)
TS-PCA	MSPE ($q = 1$)	1.200	1.178
	MSPE ($q = 2$)	1.197	1.186
	Groups	8 Groups	8 Groups
DOC	MSPE ($q = 1$)	1.127	1.128
	MSPE ($q = 2$)	1.118	1.119
	Groups	8 Groups	8 Groups
SOBI	MSPE ($q = 1$)	1.115	1.235
	MSPE ($q = 2$)	1.109	1.218
	Groups	8 Groups	8 Groups
sUARMA	MSPE ($q = 1$)	1.505	1.505
	MSPE ($q = 2$)	1.093	1.093
	Groups	1 Group	1 Group
sVARMA	MSPE ($q = 1$)	1.086	1.086
	MSPE ($q = 2$)	1.029	1.029
	Groups	1 Group	1 Group
refVAR	MSPE ($q = 1$)	1.590	1.590
	MSPE ($q = 2$)	2.571	2.571

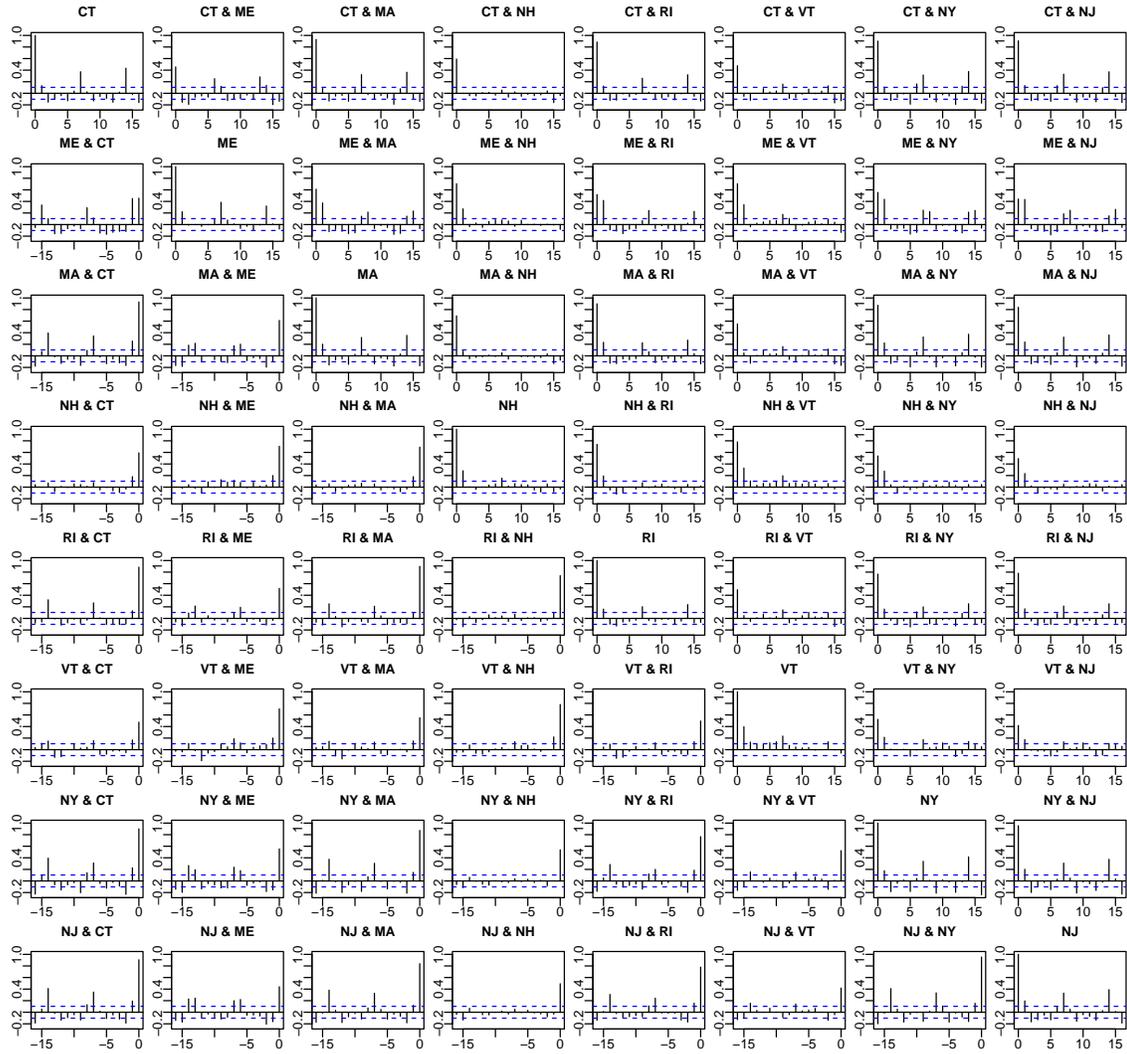


Figure S1: Auto and cross correlations of the 8 states of community mobility data.

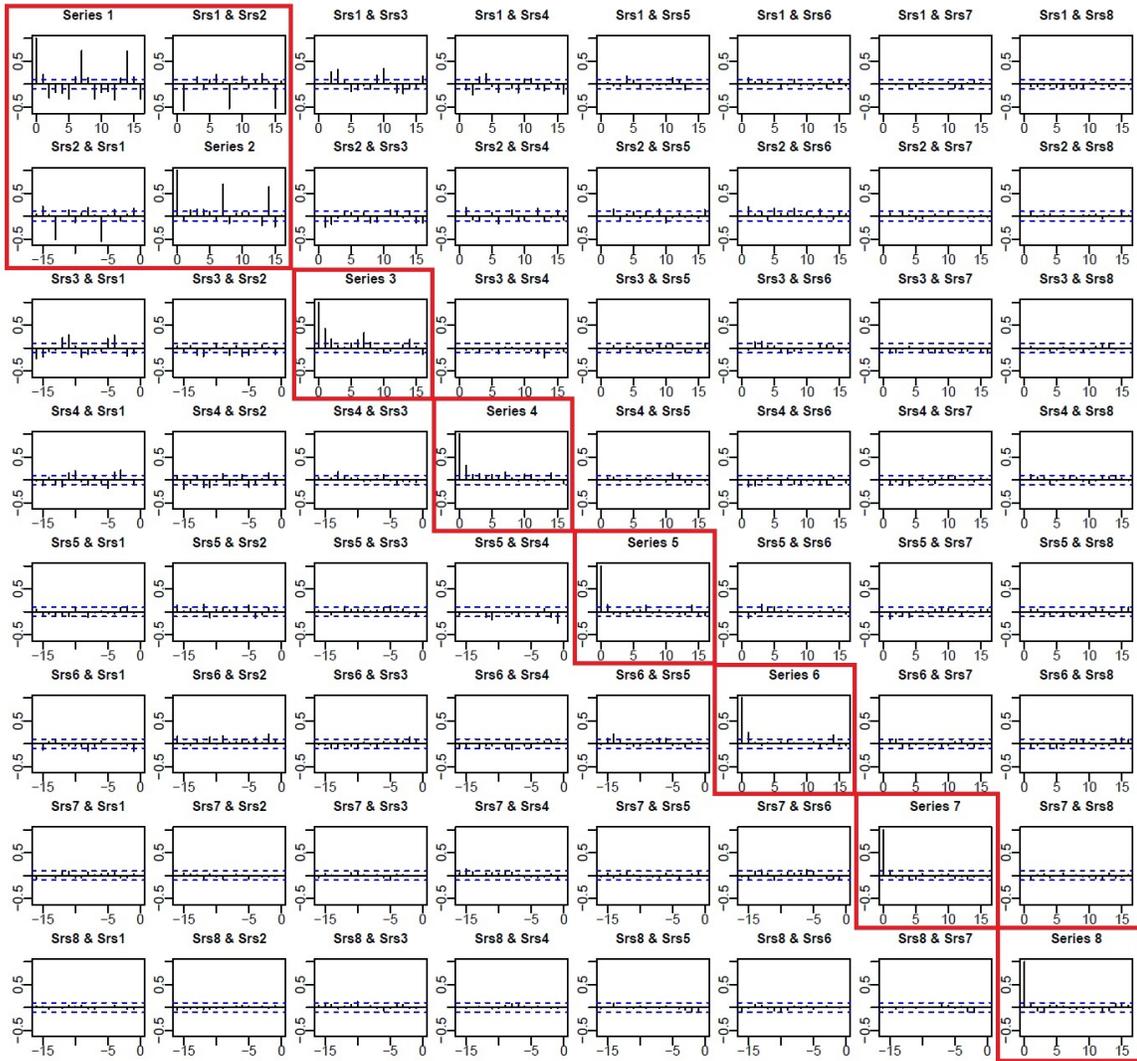


Figure S2: Auto and cross correlations of the uncorrelated components obtained by TS-PCA for community mobility data. The subseries inside the red box belong to the same group.

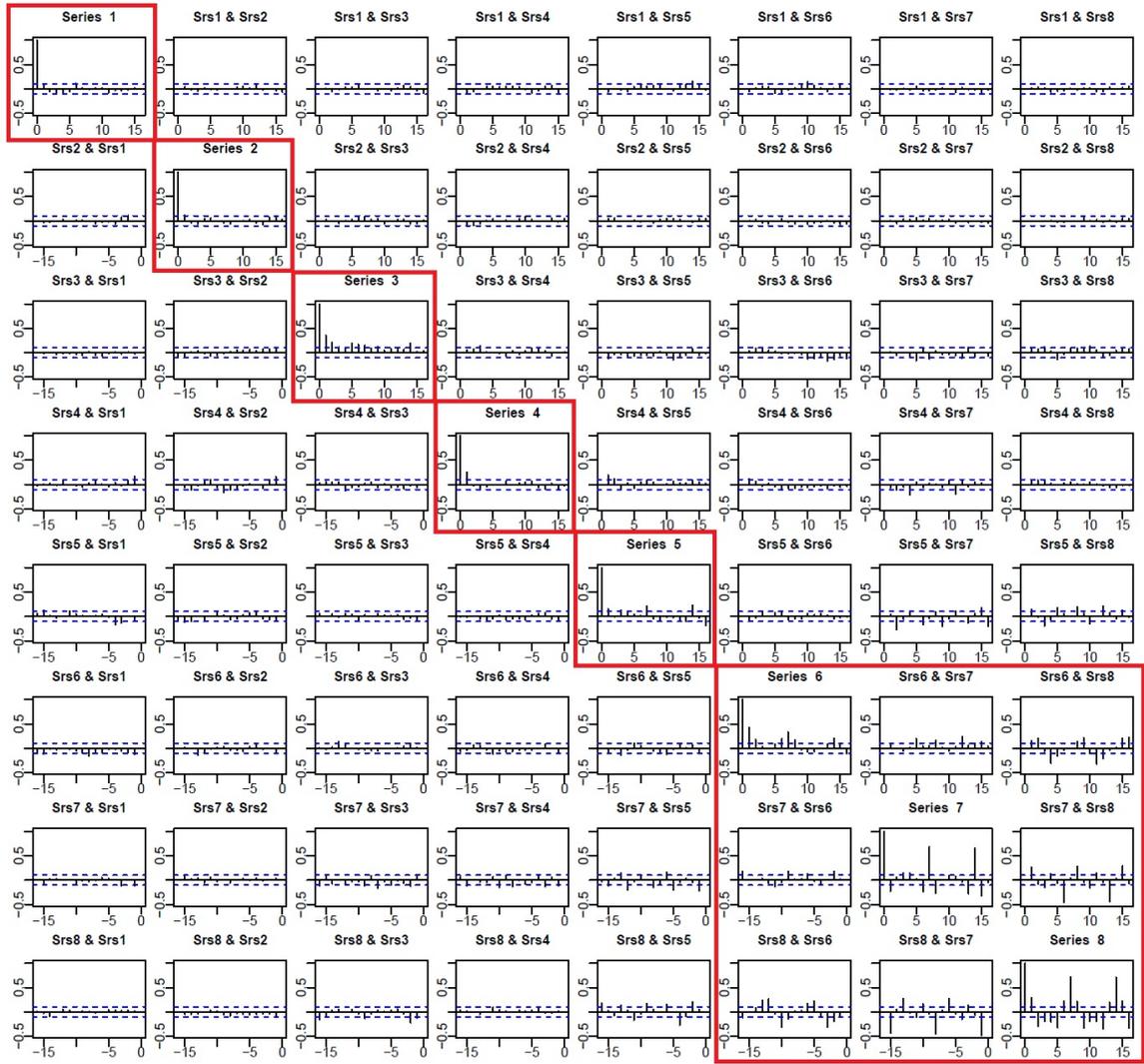


Figure S3: Auto and cross correlations of the group mean independent components obtained by GMICA for community mobility data. The subseries inside the red box belong to the same group.

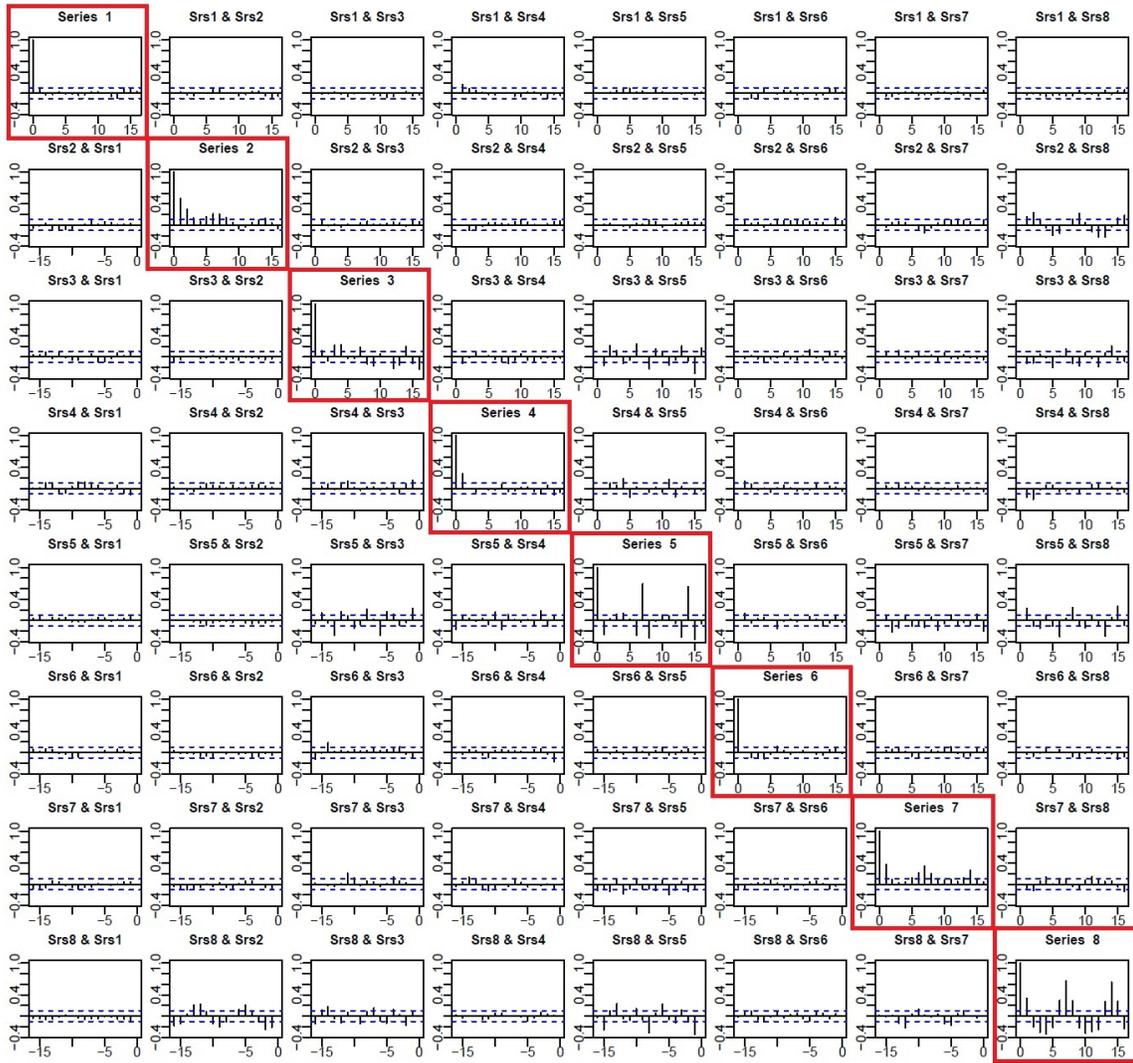


Figure S4: Auto and cross correlations of the uncorrelated components components obtained by DOC for community mobility data. The series inside the red box forms one group.

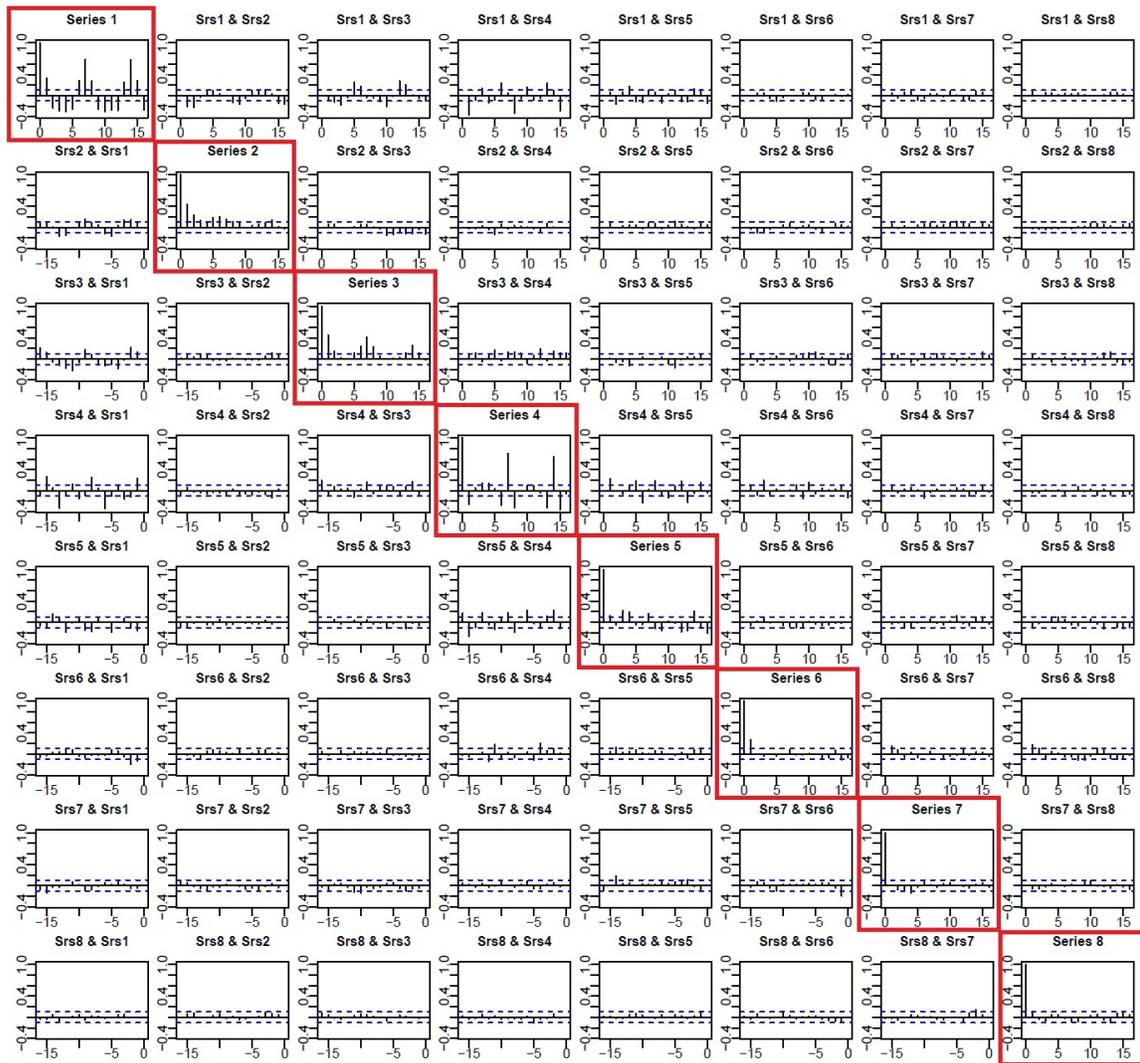


Figure S5: Auto and cross correlations of the uncorrelated components obtained by SOBI for community mobility data. The series inside the red box forms one group.

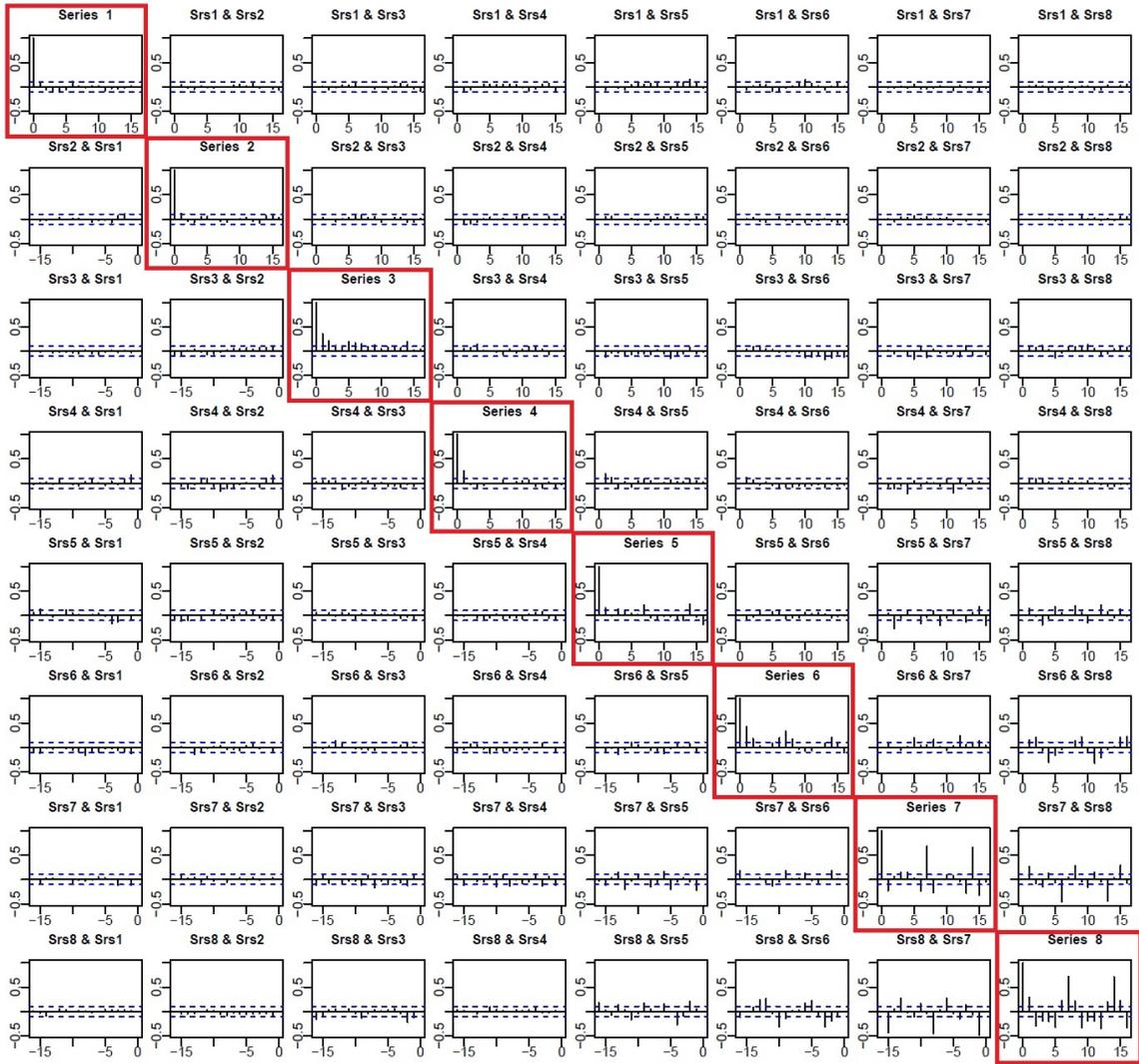


Figure S6: Auto and cross correlations of the mean independent components obtained by MICA for community mobility data. The series inside the red box forms one group.

References

- Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F. & Moulines, E. (1997), ‘A blind source separation technique using second-order statistics’, *IEEE Transactions on signal processing* **45**(2), 434–444.
- Chang, J., Guo, B. & Yao, Q. (2018), ‘Principal component analysis for second-order stationary vector time series’, *The Annals of Statistics* **46**(5), 2094–2124.
- Chen, X. (2007), ‘Large sample sieve estimation of semi-nonparametric models’, *Handbook of econometrics* **6**, 5549–5632.
- Dette, H. & Spreckelsen, I. (2004), ‘Some comments on specification tests in nonparametric absolutely regular processes’, *Journal of Time Series Analysis* **25**(2), 159–172.
- Matteson, D. S. & Tsay, R. S. (2011), ‘Dynamic orthogonal components for multivariate time series’, *Journal of the American Statistical Association* **106**(496), 1450–1463.
- Merlevède, F., Peligrad, M. & Rio, E. (2011), ‘A bernstein type inequality and moderate deviations for weakly dependent sequences’, *Probability Theory and Related Fields* **151**, 435–474.
- Vaart, A. V. D. & Wellner, J. A. (1997), ‘Weak convergence and empirical processes with applications to statistics’, *Journal of the Royal Statistical Society-Series A Statistics in Society* **160**(3), 596–608.
- White, H. & Wooldridge, J. (1991), ‘Some results on sieve estimation with dependent observations’, *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proc. 5th Int. Symp. in Economic Theory and Econometrics* pp. 459–493.

Yoshihara, K.-i. (1976), 'Limiting behavior of u-statistics for stationary, absolutely regular processes', *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **35**(3), 237–252.