Supplement to “Asymptotic Distributions of High-Dimensional Distance Correlation Inference”

Lan Gao, Yingying Fan, Jinchi Lv and Qi-Man Shao

This Supplementary Material contains all the proofs and technical details. Section A presents the proofs of the main results in Theorems 1–6 and Propositions 1–3 in Section A.4. We provide the proofs of Propositions 1–3, some key lemmas with their proofs, and additional technical details in Sections B–F. In particular, Section D presents the parallel versions of Theorems 2 and 4 for the case of $1/2 < \tau \leq 1$ and their proofs, while Section E discusses the connections between the normal approximation for $T_n^*$ and the gamma approximation for $n V^*(X,Y)$. Moreover, we provide the proof of the asymptotic normality and associated rates of convergence for $T_R$ in Section F. Throughout the paper, $C$ stands for some positive constant whose value may change from line to line.

APPENDIX A: PROOFS OF MAIN RESULTS

A.1. Proof of Theorem 1. Note that Huo and Székely (2016) showed that $V_n^*(X,Y)$ is a U-statistic. The main idea of our proof is to apply the Hoeffding decomposition for U-statistics and the martingale central limit theorem. Lemmas 1–4 in Sections C.1–C.4 of Supplementary Material, respectively, draw an outline of the proof. In particular, Lemma 1 provides the ratio consistency of $V_n^*(X)$ and $V_n^*(Y)$. Thus by (A.48) and (A.49), the denominator of $T_n$ can be replaced with the corresponding population counterpart in Lemma 1. In consequence, by Slutsky’s lemma it suffices to analyze the limiting distribution of the following random variable

\begin{equation}
\tilde{T}_n = \sqrt{n(n-1)} \frac{V_n^*(X,Y)}{\sqrt{V^2(X)V^2(Y)}}.
\end{equation}

Moreover, we have the conclusion in Lemma 2 by the Hoeffding decomposition. In fact, Lemma 2 implies that under the independence of $X$ and $Y$, $\tilde{T}_n$ can be decomposed into two parts $W_n^{(1)}(X,Y)$ and $W_n^{(2)}(X,Y)$, where the former is the leading term and the latter is asymptotically negligible. Hence to obtain the limiting distribution of $\tilde{T}_n$, it suffices to focus on $W_n^{(1)}(X,Y)$ defined in (A.52).

Recall the definition of the double-centered distance $d(\cdot, \cdot)$ in (4). Define $\zeta_{n,1} = 0$ and for $k \geq 2$,

\begin{equation}
\zeta_{n,k} = \sqrt{\frac{2}{n(n-1)}} \sum_{i=1}^{k-1} d(X_i, X_k) d(Y_i, Y_k) / \sqrt{V^2(X)V^2(Y)}.
\end{equation}

It is easy to see that $W_n^{(1)}(X,Y) = \sum_{k=1}^{n} \zeta_{n,k}$. Then by Lemmas 3 and 4, (18) and (19) directly lead to

\[
\sum_{k=1}^{n} \mathbb{E}[\zeta_{n,k}^2 | \mathcal{F}_{k-1}] \to 1 \quad \text{in probability}
\]

with $\mathcal{F}_k$ a $\sigma$-algebra defined in Lemma 3, and for any $\varepsilon > 0$,

\[
\sum_{k=1}^{n} \mathbb{E}[\zeta_{n,k}^2 1\{|\zeta_{n,k}| > \varepsilon\}] \to 0.
\]

Therefore, by the Lindeberg-type central limit theorem for martingales (see, for example, Brown (1971)), we can obtain $W_n^{(1)}(X,Y) \overset{D}{\to} N(0,1)$. This completes the proof of Theorem 1.
A.2. Proof of Theorem 2. The main idea of the proof is based on the conclusion of Theorem 4. In view of the definitions of $E_x$ and $L_{x,\tau}$, by the Cauchy–Schwarz inequality we can obtain that

$$E_x \geq \frac{B_x^{-2\tau} L_{x,\tau}^{(2+\tau)/(1+\tau)}}{\{\mathbb{E}[(X_1^T X_2)^2]\}^{2}} \geq \frac{B_x^{-2\tau} L_{x,\tau}}{\mathbb{E}[(X_1^T X_2)^2]}.$$  

In the same manner, we can deduce

$$E_y \geq \frac{B_y^{-2\tau} L_{y,\tau}}{\mathbb{E}[(Y_1^T Y_2)^2]}.$$  

Note that $p + q \to \infty$ implies that at least one of $p$ and $q$ tends to infinity. First let us assume that both $p \to \infty$ and $q \to \infty$. Then by assumption, we have $E_x \to 0$ and $E_y \to 0$. Thus for sufficiently large $p$ and $q$, it holds that

$$B_x^{-2\tau} L_{x,\tau}/\mathbb{E}[(X_1^T X_2)^2] \leq \frac{1}{18} \quad \text{and} \quad B_y^{-2\tau} L_{y,\tau}/\mathbb{E}[(Y_1^T Y_2)^2] \leq \frac{1}{18}.$$  

It follows from Theorem 4 that if (20) holds, $E_x \to 0$, and $E_y \to 0$, then we have

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \Phi(x)| \to 0$$  

with $\Phi(x)$ the standard normal distribution function, which yields $T_n \overset{D}{\to} N(0, 1)$.

We now consider the scenario when only one of $p$ and $q$ tends to infinity. Without loss of generality, assume that $p$ is bounded and $q \to \infty$. Then by assumption, we have $E_y \to 0$. In addition, note that $L_{x,\tau} \geq (\mathbb{E}[(X_1^T X_2)^2])^{1+\tau}$. Thus it follows from (20) that

$$\frac{n^{-\tau} L_{y,\tau}}{\{\mathbb{E}[(Y_1^T Y_2)^2]\}^{1+\tau}} \to 0.$$  

Consequently, an application of bound (24) results in

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \Phi(x)| \to 0,$$  

which concludes the proof of Theorem 2.

A.3. Proof of Theorem 3. The key ingredient of the proof is to replace the denominator with the population counterpart and apply the convergence rate in the martingale central limit theorem. In light of the definition in (A.1), we can write

$$|\mathbb{P}(T_n \leq x) - \Phi(x)| = |\mathbb{P}(\hat{T}_n \cdot \sqrt{\frac{\mathcal{V}_n(X)\mathcal{V}_n(Y)}{\mathcal{V}_n^*(X)\mathcal{V}_n^*(Y)}} \leq x) - \Phi(x)|.$$  

Note that Lemma 1 entails that $\mathcal{V}_n^*(X)/\mathcal{V}_n^*(X)$ and $\mathcal{V}_n^*(Y)/\mathcal{V}_n^*(Y)$ converge to one in probability. Thus we can relate the distance between $\mathbb{P}(T_n \leq x)$ and $\Phi(x)$ to that between $\mathbb{P}(\hat{T}_n \leq x)$ and $\Phi(x)$. Specifically, for small quantities $\gamma_1 > 0$ and $\gamma_2 > 0$ it holds that

(A.3) $|\mathbb{P}(T_n \leq x) - \Phi(x)| \leq P_1 + P_2 + \mathbb{P}
left(\frac{|\mathcal{V}_n^*(X)\mathcal{V}_n^*(Y)}{\mathcal{V}_n^*(X)\mathcal{V}_n^*(Y)} - 1| > \gamma_1 \right) + \mathbb{P}
left(\left|\frac{\mathcal{V}_n^*(Y)}{\mathcal{V}_n^*(Y)} - 1\right| > \gamma_2 \right),$  

where

$$P_1 = \left|\mathbb{P}(\hat{T}_n \leq x(1 + \gamma_1)(1 + \gamma_2)) - \Phi(x)\right|,$$  

$$P_2 = \left|\mathbb{P}(\hat{T}_n \leq x(1 - \gamma_1)(1 - \gamma_2)) - \Phi(x)\right|. $$
Let us choose
\[ \gamma_1 = \left\{ \frac{\mathbb{E}[d(X_1, X_2)]^{2+2\tau}}{n^\tau V^2(X)} \right\}^{1/(2+\tau)}, \quad \gamma_2 = \left\{ \frac{\mathbb{E}[d(Y_1, Y_2)]^{2+2\tau}}{n^\tau V^2(Y)} \right\}^{1/(2+\tau)}. \]

Without loss of generality, assume that \( \gamma_1 \leq 1/2 \) and \( \gamma_2 \leq 1/2 \). Otherwise since
\[ (A.4) \quad \mathbb{E}[d(X_1, X_2)]^{2+2\tau} \geq \left\{ \mathbb{E}[d^2(X_1, X_2)] \right\}^{1+\tau} = \mathbb{E}[V^2(X)]^{1+\tau} \]
and similar result holds for \( Y \), we have
\[ \mathbb{E}[d(X_1, X_2)]^{2+2\tau} \mathbb{E}[d(Y_1, Y_2)]^{2+2\tau} \geq \max \left\{ \frac{\mathbb{E}[d(X_1, X_2)]^{2+2\tau}}{n^\tau V^2(X)} \mathbb{E}[d(Y_1, Y_2)]^{2+2\tau} \right\} \]
\[ \geq 2^{-(2+\tau)} \]
and thus the desired result (21) is trivial.

Now we bound the four terms on the right hand side of (A.3). By (A.51), it holds that
\[ (A.5) \quad \mathbb{P}\left( \left| \frac{\mathbb{E}^n(X)}{\mathbb{E}^n(Y)} - 1 \right| > \gamma_1 \right) \leq C \mathbb{E}[d(X_1, X_2)]^{2+2\tau} n^\tau V^2(X)^{1+\tau} = C \left\{ \frac{\mathbb{E}[d(X_1, X_2)]^{2+2\tau}}{n^\tau V^2(X)} \right\}^{1/(2+\tau)} \]
and similarly,
\[ (A.6) \quad \mathbb{P}\left( \left| \frac{\mathbb{E}^n(Y)}{\mathbb{E}^n(X)} - 1 \right| > \gamma_2 \right) \leq C \left\{ \frac{\mathbb{E}[d(Y_1, Y_2)]^{2+2\tau}}{n^\tau V^2(Y)} \right\}^{1/(2+\tau)}. \]

Then we deal with term \( P_1 \). By symmetry, term \( P_2 \) shares the same bound as term \( P_1 \). By Lemma 2, \( T_n \) can be decomposed into two parts, one being the dominating martingale array and the other being an asymptotically negligible error term. In details, for \( 0 < \gamma_3 = n^{-1/3}/4 \leq 1/4 \) we have
\[ P_1 \leq P_{11} + P_{12} + \mathbb{P}(\mathbb{E}^n(X, Y) > \gamma_3), \]
where
\[ P_{11} = \mathbb{P}(\mathbb{E}^n(X, Y) \leq x(1 + \gamma_1)(1 + \gamma_2)(1 + \gamma_3) - \Phi(x)), \]
\[ P_{12} = \mathbb{P}(\mathbb{E}^n(X, Y) \leq x(1 + \gamma_1)(1 + \gamma_2)(1 + \gamma_3) - \Phi(x)). \]

It follows from Lemma 2 that
\[ (A.7) \quad \mathbb{P}(\mathbb{E}^n(X, Y) > \gamma_3) \leq \frac{1}{n^{\gamma_3}} \leq 16n^{-1/3}. \]

Since terms \( P_{11} \) and \( P_{12} \) share the same bound, it suffices to show the analysis for term \( P_{11} \). It holds that
\[ (A.8) \quad P_{11} \leq \sup_{x \in \mathbb{R}} \left| \mathbb{P}(\mathbb{E}^n(X, Y) \leq x) - \Phi(x) \right| + \sup_{x \in \mathbb{R}} \left| \Phi(x(1 + \gamma_1)(1 + \gamma_2) - 1 - \Phi(x)) \right|. \]

Observe that by definitions, we have \( \gamma_1 \leq 1/2, \gamma_2 \leq 1/2, \) and \( \gamma_3 \leq 1/4 \). When \( |x| \leq 2 \), it is easy to see that
\[ |\Phi(x(1 + \gamma_1)(1 + \gamma_2) - \gamma_3) - \Phi(x)| \leq C(\gamma_1 + \gamma_2 + \gamma_3). \]

When \( |x| > 2 \), we have \( |x(1 + \gamma_1)(1 + \gamma_2)|/2 > \gamma_3 \) and thus
\[ |\Phi(x(1 + \gamma_1)(1 + \gamma_2) - \gamma_3) - \Phi(x)| \leq C(x \gamma_1 + x \gamma_2 + \gamma_3)e^{-x^2/128} \leq C(\gamma_1 + \gamma_2 + \gamma_3). \]
Consequently, it follows that
\[
\sup_{x \in \mathbb{R}} |\Phi(x(1 + \gamma_1)(1 + \gamma_2) + \gamma_3) - \Phi(x)| \leq C(\gamma_1 + \gamma_2 + \gamma_3).
\]

As for the bound of \( |P(W_n^{(1)}(X, Y) \leq x) - \Phi(x)| \), note that \( W_n^{(1)}(X, Y) = \sum_{k=1}^{n} \zeta_{n,k} \) and Lemma 3 states that \( \{\zeta_{n,k}, \mathcal{F}_k\}, k \geq 1 \) is a martingale difference array under the independence of \( X \) and \( Y \). Hence by Theorem 1 in Haeusler (1988) on the convergence rate of the martingale central limit theorem and Lemma 4, we can obtain
\[
\sup_{x \in \mathbb{R}} |P(W_n^{(1)}(X, Y) \leq x) - \Phi(x)| \\
\leq C \left\{ \sum_{k=1}^{n} E[|\zeta_{n,k}|^{2+2\tau}] + E\left( \left| \sum_{k=1}^{n} E[\zeta_{n,k}^2 | \mathcal{F}_{k-1}] - 1 \right|^{1+\tau} \right) \right\}^{1/(3+2\tau)} \\
\leq C \left\{ \left( \frac{E[g(X_1, X_2, X_3, X_4)]}{V^2(X)V^2(Y)} \right)^{(1+\tau)/2} \right. \\
+ \frac{E[|d(X_1, X_2)|^{2+2\tau}]E[|d(Y_1, Y_2)|^{2+2\tau}]}{n^\tau V^2(X)V^2(Y)} \right\}^{1/(3+2\tau)}.
\]
(A.10)

By (A.4) and \( 0 < \tau \leq 1 \), it holds that
\[
\frac{E[|d(X_1, X_2)|^{2+2\tau}]E[|d(Y_1, Y_2)|^{2+2\tau}]}{n^\tau V^2(X)V^2(Y)} \geq \frac{E[|d(X_1, X_2)|^{2+2\tau}]}{n^\tau V^2(X)}
\]
and
\[
\left\{ \frac{E[|d(X_1, X_2)|^{2+2\tau}]}{n^\tau V^2(X)} \right\}^{1/(2+\tau)} \geq n^{-\tau/(2+\tau)} \geq n^{-1/3}.
\]

Finally, the desired result (21) can be derived by plugging in (A.5)–(A.10) and noting that all the error terms can be absorbed into (A.10). This completes the proof of Theorem 3.

\subsection*{A.4. Proof of Theorem 4.} The proof is mainly based on the conclusion of Theorem 3. It is quite challenging to calculate the exact form of the moments that appear in conditions (18) and (19). Nevertheless, the bounds of these moments can be worked out in concise form under some general conditions. These bounds are summarized in the following three propositions, respectively.

**Proposition 1.** If \( E[||X||^{4+4\tau}] < \infty \) for some constant \( \tau > 0 \), then there exists some absolute positive constant \( C \), such that
\[
E(|d(X_1, X_2)|^{2+2\tau}) \leq C_x B_X^{-1(1+\tau)} L_{x,\tau}.
\]
(A.11)

**Proposition 2.** If \( E[||X||^{4+4\tau}] < \infty \) for some constant \( 0 < \tau \leq 1/2 \), then it holds that
\[
|V^2(X) - B_X^{-1}E[(X^T_1 X_2)^2]| \leq 9 B_X^{-1(1+2\tau)} L_{x,\tau}.
\]
(A.12)

**Proposition 3.** If \( E[||X||^{4+4\tau}] < \infty \) for some constant \( 0 < \tau \leq 1/2 \), then there exists some absolute positive constant \( C \) such that
\[
E[g(X_1, X_2, X_3, X_4)] \leq B_X^{-2} E[(X^T_1 \Sigma x X_2)^2] + C B_X^{-1(2+2\tau)} L_{x,\tau}^{2(2+\tau)/1(1+\tau)}.
\]
(A.13)
The proofs of Propositions 1–3 are presented in Sections A.7–A.9, respectively. We now proceed with the proof of Theorem 4. Note that condition (22) entails that
\[ 9B_X^{-(1+2r)} L_{x,r} \leq \frac{1}{2} B_X^{-1} \mathbb{E}[(X_1^T X_2)^2] \text{ and } 9B_Y^{-(1+2r)} L_{y,r} \leq \frac{1}{2} B_Y^{-1} \mathbb{E}[(Y_1^T Y_2)^2]. \]
Therefore, it follows from Proposition 2 that
\[ \sqrt{V^2(X)} \geq \frac{1}{2} B_X^{-1} \mathbb{E}[(X_1^T X_2)^2] \text{ and } \sqrt{V^2(Y)} \geq \frac{1}{2} B_Y^{-1} \mathbb{E}[(Y_1^T Y_2)^2], \]
which together with Propositions 1 and 3 yield the desired results (23) by Theorem 3. This concludes the proof of Theorem 4.

A.5. Proof of Theorem 5. Recall that
\[ T_n = \sqrt{\frac{n(n-1)}{2}} \frac{\mathbb{V}_n(X,Y)}{\sqrt{\mathbb{V}_n^2(X) \mathbb{V}_n^2(Y)}} \]
and it has been proved in Lemma 1 in Section C.1 that under condition (18), we have \( \mathbb{V}_n^2(X)/\mathbb{V}_n^2(X) \rightarrow 1 \) and \( \mathbb{V}_n^2(Y)/\mathbb{V}_n^2(Y) \rightarrow 1 \) in probability. Thus it suffices to show that for any arbitrarily large constant \( C > 0 \),
\[ \hat{T}_n := \sqrt{\frac{n(n-1)}{2}} \frac{\mathbb{V}_n^2(X,Y)}{\sqrt{\mathbb{V}_n^2(X) \mathbb{V}_n^2(Y)}} > C \text{ with asymptotic probability 1.} \]
Observe that
\[ |\hat{T}_n - \sqrt{\frac{n(n-1)}{2}} R^2(X,Y)| = \sqrt{\frac{n(n-1)}{2}} \frac{|\mathbb{V}_n^2(X,Y) - \mathbb{V}_n^2(X,Y)|}{\sqrt{\mathbb{V}_n^2(X) \mathbb{V}_n^2(Y)}}. \]
It follows from (A.50), (A.58) and Proposition 1 that there exists some absolute positive constant \( C \) such that
\[ \mathbb{E}[(\mathbb{V}_n^2(X,Y) - \mathbb{V}_n^2(X,Y))^2] \leq C n^{-1} \mathbb{E}[h^2((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))] \]
\[ \leq C n^{-1} (\mathbb{E}[d^4(X_1, X_2)] \mathbb{E}[d^4(Y_1, Y_2)])^{1/2} \]
\[ \leq C n^{-1} B_X^{-1} B_Y^{-1} L_{x,1}^{1/2} L_{y,1}^{1/2}. \]
Therefore, if \( \sqrt{n} \mathbb{V}(X,Y)/\left(B_X^{-1/2} B_Y^{-1/2} L_{x,1}^{1/4} L_{y,1}^{1/4}\right) \rightarrow \infty \), it holds that
\[ \frac{|\mathbb{V}_n^2(X,Y) - \mathbb{V}_n^2(X,Y)|}{\sqrt{\mathbb{V}_n^2(X) \mathbb{V}_n^2(Y)}} \rightarrow 0 \text{ in probability.} \]
This together with \( n R^2(X,Y) \rightarrow \infty \) yields for any arbitrarily large constant \( C > 0 \), \( \mathbb{P}(\hat{T}_n > C) \rightarrow 1 \) and hence as \( \mathbb{P}(T_n > C) \rightarrow 1 \), which completes the proof of Theorem 5.

A.6. Proof of Theorem 6. The main ingredient of the proof is bounding \( \mathbb{V}(X,Y) \) using the decomposition developed in Lemma 10 in Section C.10. We will calculate the orders of terms \( I_i, 1 \leq i \leq 5 \), introduced in Lemma 10. Let us begin with the first term
\[ I_1 = \frac{1}{4} B_X^{1/2} B_Y^{1/2} \left( \mathbb{E}[W_{12} V_{12}] - 2 \mathbb{E}[W_{12} V_{13}] \right), \]
where \( W_{12} = B_X^{-1}(\|X_1 - X_2\|^2 - B_X) \) and \( V_{12} = B_Y^{-1}(\|Y_1 - Y_2\|^2 - B_Y) \). Denote by \( \tilde{Y}_1 = Y_1 - \mathbb{E}Y = (Y_{1,1} - \mathbb{E}Y_{1,1}, \ldots, Y_{1,p} - \mathbb{E}Y_{1,p})^T \) and \( \tilde{Y}_2 = Y_2 - \mathbb{E}Y = (Y_{2,1} - \mathbb{E}Y_{2,1}, \ldots, Y_{2,p} - \mathbb{E}Y_{2,p})^T \) the centered random variables, and define

\[
\alpha_1(X) = \|X\|^2 - \mathbb{E}\|X\|^2, \quad \alpha_2(X_1, X_2) = X_1^T X_2,
\]

\[
\beta_1(Y) = \|Y\|^2 - \mathbb{E}\|Y\|^2, \quad \beta_2(Y_1, Y_2) = \tilde{Y}_1^T \tilde{Y}_2.
\]

Since \( \mathbb{E}[\alpha_1(X)] = \mathbb{E}[\beta_1(Y)] = 0 \) and \( \mathbb{E}[\alpha_2(X_1, X_2)] = \mathbb{E}[\beta_2(Y_1, Y_2)] = 0 \), it holds that

\[
\mathbb{E}[W_{12}V_{12}] = \mathbb{E}\left[ (\alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2)) (\beta_1(Y_1) + \beta_1(Y_2) - 2\beta_2(Y_1, Y_2)) \right]
\]

\[
= 2\mathbb{E}[\alpha_1(X_1)\beta_1(Y_1)] + 4\mathbb{E}[\alpha_2(X_1, X_2)\beta_2(Y_1, Y_2)].
\]

Similarly, we have \( \mathbb{E}[W_{12}V_{13}] = 2\mathbb{E}[\alpha_1(X_1)\beta_1(Y_1)] \). Thus it follows that

\[
I_1 = 4\mathbb{E}[\alpha_2(X_1, X_2)\beta_2(Y_1, Y_2)] = 4 \sum_{i,j=1}^{p} (\text{cov}(X_{1,i}, Y_{1,j}))^2.
\]

Observe that under the symmetry assumptions, there is no linear dependency between \( X \) and \( Y \); that is, \( \text{cov}(X_{1,i}, Y_{1,j}) = 0 \) for each \( 1 \leq i, j \leq p \). This together with the representation of \( I_1 \) above entails that \( I_1 = 0 \).

We now consider the second term \( I_2 \). Using similar arguments but much more tedious calculations, we can obtain

\[
I_2 = \frac{1}{4} B_X^{-1/2} B_Y^{-3/2} \left( 2\mathbb{E}[\alpha_2(X_1, X_2)\beta_2^2(Y_1, Y_2)] + \mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_1(Y_2)] 
\]

\[
- 4\mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_2(Y_1, Y_2)] \right)
\]

\[
+ \frac{1}{4} B_X^{-3/2} B_Y^{-1/2} \left( 2\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_2^2(X_1, X_2)] + \mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_1(X_2)] 
\]

\[
- 4\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_2(X_1, X_2)] \right).
\]

By assumption, we have \( c_2 p \leq B_X \leq c_1 p \) and \( c_2 p \leq B_Y \leq c_1 p \). Since \( X \) has a symmetric distribution and \( Y_{1,j} = g_j(X_{1,j}) \) with \( g_j(x) \), \( 1 \leq j \leq p \), symmetric functions, it holds that

\[
\mathbb{E}[\alpha_2(X_1, X_2)\beta_2^2(Y_1, Y_2)] = \mathbb{E}[(X_1^T X_2)(\tilde{Y}_1^T \tilde{Y}_2)^2] = \mathbb{E}((-X_1^T X_2)(\tilde{Y}_1^T \tilde{Y}_2)^2) = 0.
\]

Similarly, with the symmetry assumptions we can show that \( \mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_1(Y_2)] = 0 \), \( \mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_2(Y_1, Y_2)] = 0 \), and \( \mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_2(X_1, X_2)] = 0 \). Moreover, it holds that

\[
\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_2^2(X_1, X_2)] = \sum_{i,j,k=1}^{p} \left( \mathbb{E}[X_{1,i}X_{1,j}\tilde{Y}_{1,k}] \right)^2 \geq 0,
\]

\[
\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_1(X_2)] = \sum_{i=1}^{p} \left( \sum_{j=1}^{p} \mathbb{E}[\tilde{Y}_{1,i}(X_{1,j}^2 - \mathbb{E}X_{1,j}^2)] \right)^2 \geq 0.
\]

Thus it follows that \( I_2 \geq 0 \).

Let us proceed with terms \( I_3 \) and \( I_4 \). By some tedious calculations, we can deduce that

\[
I_3 = \frac{1}{8} B_X^{-1/2} B_Y^{-5/2} \left( 4\mathbb{E}[\alpha_2(X_1, X_2)\beta_2^2(Y_1, Y_2)] + 6\mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_1(Y_2)\beta_2(Y_1, Y_2)] 
\]

\[
+ 6\mathbb{E}[\alpha_2(X_1, X_2)\beta_1^2(Y_1)\beta_2(Y_1, Y_2)] - 3\mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_1^2(Y_2)] \right).
\]
Consequently, it follows that

\[-12\mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_2^2(Y_1, Y_2)]\]

\[+ \frac{1}{8}B_X^{-5/2}B_Y^{-1/2}\left(4\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_2^3(X_1, X_2)] + 6\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_1(X_2)\alpha_2(X_1, X_2)]
+ 6\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_2^2(X_1)\alpha_2(X_1, X_2)] - 3\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_1^2(X_2)]
- 12\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_2^2(X_1, X_2)]\right)\]

and

\[I_4 = \frac{1}{16}B_X^{-3/2}B_Y^{-3/2}\left(4\mathbb{E}[\alpha_2^3(X_1, X_2)\beta_2^2(Y_1, Y_2)] + (\mathbb{E}[\alpha_1(X)\beta_1(Y)])^2
+ 8\mathbb{E}[\alpha_2(X_1, X_2)\beta_2^2(Y_1, Y_3)] + 4\mathbb{E}[\alpha_2(X_1, X_2)]\mathbb{E}[\beta_2^2(Y_1, Y_2)]
+ 2\mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_1(Y_2)] - 8\mathbb{E}[\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_2(Y_1, Y_2)]
+ 2\mathbb{E}[\alpha_1(X_1)\alpha_1(X_2)\beta_2^2(Y_1, Y_2)] - 4\mathbb{E}[\alpha_1(X_1)\alpha_1(X_2)\beta_1(Y_1)\beta_2(Y_1, Y_2)]
- 8\mathbb{E}[\alpha_1(X_1)\alpha_2(X_1, X_2)\beta_2(Y_1, Y_2)] - 4\mathbb{E}[\alpha_1(X_1)\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_2(Y_1, Y_2)]
+ 8\mathbb{E}[\alpha_1(X_1)\alpha_2(X_1, X_2)\beta_1(Y_1)\beta_2(Y_1, Y_2)] + 8\mathbb{E}[\alpha_1(X_1)\alpha_2(X_1, X_2)\beta_1(Y_2)\beta_2(Y_1, Y_2)]
- 8\mathbb{E}[\alpha_2^2(X_1, X_2)\beta_1(Y_3)\beta_2(Y_1, Y_3)] - 8\mathbb{E}[\alpha_1(X_2)\alpha_2(X_1, X_2)\beta_2^2(Y_1, Y_3)]
+ 8\mathbb{E}[\alpha_1(X_2)\alpha_2(X_1, X_2)\beta_1(Y_3)\beta_2(Y_1, Y_3)]\right).

A useful observation is that under the assumptions that $X_1$ has a symmetric distribution and $g_j(x)$ with $1 \leq j \leq p$ are symmetric functions, many terms in $I_3$ and $I_4$ above in fact become zero. In particular, we can show that

\[I_3 = \frac{1}{8}B_X^{-5/2}B_Y^{-1/2}\left(-3\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_1^2(X_2)]
- 12\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_2^2(X_1, X_2)]\right).

Denote by \(D(i) = \{(j, k, l) : \max(|j - i|, |k - i|, |l - i|) \leq 3m + 1\}\). Since \(\{X_{1,i}, 1 \leq i \leq p\}\) are \(m\)-dependent, it holds that

\[
\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_1^2(X_2)]
= \sum_{i=1}^{p} \sum_{(j,k,l) \in D(i)} \mathbb{E}[\tilde{Y}_{1,i}X_{1,j}^2 - \mathbb{E}X_{1,j}^2] \mathbb{E}[\tilde{Y}_{1,i}X_{1,k}^2 - \mathbb{E}X_{1,k}^2] \mathbb{E}[\tilde{Y}_{1,i}X_{1,l}^2 - \mathbb{E}X_{1,l}^2] = O(c_1^8 m^3 p)
\]

and

\[
\mathbb{E}[\beta_2(Y_1, Y_2)\alpha_1(X_1)\alpha_2^2(X_1, X_2)]
= \sum_{i=1}^{p} \sum_{(j,k,l) \in D(i)} \mathbb{E}X_{1,k} \mathbb{E}X_{1,l} = O(c_1^8 m^3 p).
\]

Consequently, it follows that

\[|I_3| \lesssim (c_1/c_2)^8 m^3 p^{-2},\]
where \( \lesssim \) represents the asymptotic order. By the same token, the symmetry assumptions lead to

\[
I_4 = \frac{1}{16} B_X^{-3/2} B_Y^{-3/2} \left( 4 \mathbb{E}[\alpha_2^2(X_1, X_2) \beta_2^2(Y_1, Y_2)] + (\mathbb{E}[\alpha_1(X) \beta_1(Y)])^2 \right. \\
+ 8 \mathbb{E}[\alpha_2^2(X_1, X_2) \beta_2(Y_1, Y_3)] + 4 \mathbb{E}[\alpha_2^2(X_1, X_2)] \mathbb{E}[\beta_2^2(Y_1, Y_2)] \\
+ 2 \mathbb{E}[\alpha_2^2(X_1, X_2) \beta_1(Y_1) \beta_1(Y_2)] - 8 \mathbb{E}[\alpha_2^2(X_1, X_2) \beta_1(Y_1) \beta_2(Y_1, Y_2)] \\
+ 2 \mathbb{E}[\alpha_1(X_1) \alpha_1(X_2) \beta_2^2(Y_1, Y_2)] - 4 \mathbb{E}[\alpha_1(X_1) \alpha_1(X_2) \beta_1(Y_1) \beta_2(Y_1, Y_2)] \\
- 8 \mathbb{E}[\alpha_2^2(X_1, X_2) \beta_1(Y_1) \beta_2(Y_1, Y_3)].
\]

It is easy to see that \( (\mathbb{E}[\alpha_1(X) \beta_1(Y)])^2 \geq 0, \mathbb{E}[\alpha_2^2(X_1, X_2) \beta_2^2(Y_1, Y_3)] \geq 0, \)

\[
\mathbb{E}[\alpha_2^2(X_1, X_2) \beta_2^2(Y_1, Y_2)] = \sum_{i,j,k,l=1}^{p} (\mathbb{E}[X_{1,i} X_{1,j} \bar{Y}_{1,k} \bar{Y}_{1,l}])^2 \\
\geq c_2^8 p(p - 2m),
\]

and

\[
\mathbb{E}[\alpha_2^2(X_1, X_2)] \mathbb{E}[\beta_2^2(Y_1, Y_2)] = \sum_{i,j,k,l=1}^{p} (\mathbb{E}[X_{1,i} X_{1,j}])^2 (\mathbb{E}[\bar{Y}_{1,k} \bar{Y}_{1,l}])^2 \\
\geq \sum_{i,k} (\mathbb{E}[X_{1,i}])^2 (\mathbb{E}[\bar{Y}_{1,k}])^2 \geq c_2^8 p^2.
\]

Moreover, since \( \{X_{1,i}, 1 \leq i \leq p\} \) are \( m \)-dependent random variables, we can deduce

\[
\mathbb{E}[\alpha_2^2(X_1, X_2) \beta_1(Y_1) \beta_1(Y_2)] = \sum_{i=1}^{p} \sum_{(j,k) \in \bar{D}(i)} \mathbb{E} \left[ X_{1,i} X_{1,j} (Y_{1,k}^2 - \mathbb{E} Y_{1,k}^2) \right] \\
\times \mathbb{E} \left[ X_{1,i} X_{1,j} (Y_{1,k}^2 - \mathbb{E} Y_{1,k}^2) \right] = O(c_1^8 m^3 p),
\]

where \( \bar{D}(i) \) is defined similarly as for \( D(i) \). In the same fashion, we can show that

\[
\mathbb{E}[\alpha_2^2(X_1, X_2) \beta_1(Y_1) \beta_2(Y_1, Y_2)] = O(c_1^8 m^3 p), \\
\mathbb{E}[\alpha_1(X_1) \alpha_1(X_2) \beta_2^2(Y_1, Y_2)] = O(c_1^8 m^3 p), \\
\mathbb{E}[\alpha_1(X_1) \alpha_1(X_2) \beta_1(Y_1) \beta_2(Y_1, Y_2)] = O(c_1^8 m^3 p), \\
\mathbb{E}[\alpha_2^2(X_1, X_2) \beta_1(Y_3) \beta_2(Y_1, Y_3)] = O(c_1^8 m^3 p).
\]

As a result, there exists some positive constant \( A \) depending on \( c_1, c_2, \) and \( m \) such that

\[
I_4 \geq Ap^{-1} + O(p^{-2}).
\]

Finally, we deal with term \( I_5 \). In view of Lemma 10, the first term for the order of \( I_5 \) is

\[
B_X^{1/2} B_Y^{1/2} (\mathbb{E}[|W_{12}|^5])^{2/5} (\mathbb{E}[|V_{12}|^5])^{3/5} \\
= B_X^{-3/2} B_Y^{-5/2} (\mathbb{E}[||X_1 - X_2||^2 - B_X|^5])^{2/5} (\mathbb{E}[||Y_1 - Y_2||^2 - B_Y|^5])^{3/5}.
\]
Clearly, \( \{X_{1,i} : i \in E_u\} \) are independent random variables for each \( 1 \leq u \leq m + 1 \). Then it follows from the basic inequality \( |\sum_{i=1}^{n} a_i|^r \leq n^{r-1} \sum_{i=1}^{n} |a_i|^r \) for \( r \geq 1 \) and Rosenthal’s inequality for independent random variables that

\[
\mathbb{E}[\|X_1 - X_2\|^2 - B_X|^5] = \mathbb{E}\left[\left|\sum_{i=1}^{p} [(X_{1,i} - X_{2,i})^2 - \mathbb{E}(X_{1,i} - X_{2,i})^2]\right|^5\right]
\]

\[
\leq (m + 1)^4 \sum_{i=1}^{m+1} \mathbb{E}\left[\left|\sum_{i \in E_u} [(X_{1,i} - X_{2,i})^2 - \mathbb{E}(X_{1,i} - X_{2,i})^2]\right|^5\right]
\]

\[
\leq C(m + 1)^4 \sum_{i \in E_u} \left\{ \left[ \sum_{i \in E_u} \mathbb{E}\left((X_{1,i} - X_{2,i})^2 - \mathbb{E}(X_{1,i} - X_{2,i})^2\right) \right]^{5/2} + \sum_{i \in E_u} \mathbb{E}\left((X_{1,i} - X_{2,i})^2 - \mathbb{E}(X_{1,i} - X_{2,i})^2\right) \right\}.
\]

Note that by assumptions, there exists some absolute positive constant \( A \) such that \( \mathbb{E}\left((X_{1,i} - X_{2,i})^2 - \mathbb{E}(X_{1,i} - X_{2,i})^2\right)^2 \leq Ac_{10}^2 \) and \( \mathbb{E}\left((X_{1,i} - X_{2,i})^2 - \mathbb{E}(X_{1,i} - X_{2,i})^2\right)^2 \) \( \leq Ac_{1}^2 \), and we have \( B_X \geq 2c_{20}^2 p \). Then it follows that

\[
\mathbb{E}[\|X_1 - X_2\|^2 - B_X|^5] \leq c_{10}^4 m^4 \cdot m \cdot (p/m)^{5/2} = c_{10}^4 m^5/2 p^{5/2}.
\]

Similarly, we can obtain

\[
\mathbb{E}[\|Y_1 - Y_2\|^2 - B_Y|^5] \leq c_{10}^4 m^5/2 p^{5/2}.
\]

Hence it holds that

\[
B_X^{1/2} B_Y^{1/2}(\mathbb{E}|W_{12}|^{5})^{2/5}(\mathbb{E}|V_{12}|^{5})^{3/5} \leq m^{5/2} p^{-3/2}.
\]

In the same manner, we can deduce that

\[
B_X^{1/2} B_Y^{1/2}(\mathbb{E}|W_{12}|^{5})^{3/5}(\mathbb{E}|V_{12}|^{5})^{2/5} \leq m^{5/2} p^{-3/2},
\]

\[
B_X^{1/2} B_Y^{1/2}(\mathbb{E}|W_{12}|^{5})^{4/5}(\mathbb{E}|V_{12}|^{5})^{1/5} \leq m^{5/2} p^{-3/2},
\]

\[
B_X^{1/2} B_Y^{1/2}(\mathbb{E}|W_{12}|^{5})^{4/5}(\mathbb{E}|V_{12}|^{5})^{1/5} \leq m^{5/2} p^{-3/2},
\]

\[
B_X^{1/2} B_Y^{1/2}(\mathbb{E}|W_{12}|^{5})^{1/2}(\mathbb{E}|V_{12}|^{6})^{1/2} \leq m^3 p^{-2}.
\]

Thus substituting the above five inequalities into the order of \( I_5 \) in Lemma 10 yields that there exists some positive constant \( A \) depending on \( c_1, c_2, \) and \( m \) such that

\[
I_5 \leq Ap^{-3/2}.
\]

As a consequence, combining all the bounds above leads to

\[
(A.15) \quad \mathcal{V}^2(X, Y) \geq Ap^{-1} + O(p^{-3/2}).
\]
Hence this entails that when \( p = o(\sqrt{n}) \), it holds that \( \sqrt{n}V^2(X, Y) \rightarrow \infty \). Furthermore, it follows from Proposition 2 that \( V^2(X) = B_X^{-1}E[(X^T X)^2] + O(B_X^{-2}L_{x,1/2}) \). By the assumptions \( E(X_{1,i}^4) + E(Y_{1,i}^4) \leq c_1^2 \), \( \text{var}(X_{1,i}) \geq c_2^2 \), and \( \text{var}(Y_{1,i}) \geq c_3^2 \), it is easy to see that \( 2c_2^2p \leq B_X \leq 2c_3^2p \) and \( 2c_4^2p \leq B_Y \leq 2c_3^2p \). Since \( \{X_{1,i}, 1 \leq i \leq p\} \) are \( m \)-dependent, we have

\[
\sum_{i=1}^{p}(E[X_{1,i}])^2 \leq E[(X^T X)^2] = \sum_{i=1}^{p} \sum_{j=1}^{p}(E[X_{1,i}X_{1,j}])^2
\]

which yields \( c_2^2p \leq E[(X^T X)^2] \leq 2(m + 1)c_1^4p \). In the same manner, we can obtain \( L_{x,1/2} \leq Cz, \quad L_{y,1/2} \leq Cz, \quad L_X \leq Cz^2 \), and \( L_Y \leq Cz^2 \) with some positive constant \( C \) depending on \( c_1, c_2, \text{ and } m \). Consequently, there exist some positive constants \( C_1 \) and \( C_2 \) depending on \( c_1, c_2, \text{ and } m \) such that \( C_1 \leq V^2(X) \leq C_2 \). Similarly, \( C_1 \leq V^2(Y) \leq C_2 \). This along with (A.15) entails that \( R^2(X, Y) \geq A\rho^{-1} + O(\rho^{-3/2}) \), where \( A > 0 \) is some constant depending on \( c_1, c_2, \text{ and } m \).

From the above analysis, it holds that \( B_X^{-1/2}B_Y^{-1/2}L_x^{1/4}L_y^{1/4} \leq A_1 \) for some positive constant \( A_1 \) depending on \( c_1, c_2, \text{ and } m \). Thus we can obtain under the assumption of \( p = o(\sqrt{n}) \) that

\[
nR^2(X, Y) \geq An\rho^{-1} + O(n\rho^{-3/2}) \rightarrow \infty
\]

and

\[
\sqrt{n}V^2(X, Y)/(B_X^{-1/2}B_Y^{-1/2}L_x^{1/4}L_y^{1/4}) \geq A(\rho^{-1} + O(\rho^{-3/2})) \rightarrow \infty.
\]

Finally, it follows from Theorem 5 that for any arbitrarily large \( C > 0 \), \( P(T_n > C) \rightarrow 1 \) as \( n \rightarrow \infty \), which concludes the proof of Theorem 6.

### A.7. Proof of Proposition 1

In view of the definition \( B_X = E[\|X_1 - X_2\|^2] \), we can write

\[
d(X_1, X_2) = (\|X_1 - X_2\| - B_X^{1/2}) - E[(\|X_1 - X_2\| - B_X^{1/2})|X_1]
\]

\[
- E[(\|X_1 - X_2\| - B_X^{1/2})|X_2] + E[(\|X_1 - X_2\| - B_X^{1/2})].
\]

Thus it follows from Jensen’s inequality that for \( \tau > 0 \),

\[
E[\|d(X_1, X_2)\|^{2+2\tau}] \leq C_\tau E[\|X_1 - X_2\|^{2+2\tau}] - B_X^{1/2 - 2\tau} \leq C_\tau \left( \frac{\|X_1 - X_2\|^2 - B_X^{1/2 - 2\tau}}{\|X_1 - X_2\|^{2+2\tau}} \right)
\]

\[
\leq C_\tau B_X^{-(1+\tau)} E[\|X_1 - X_2\|^2 - B_X^{1+2\tau}].
\]

Moreover, we have

\[
E[\|X_1 - X_2\|^2 - B_X^{1+2\tau}]
\]

\[
\leq C_\tau \left\{ E[\|X_1\|^2 - E[\|X_1\|^2]]^{2+2\tau} + E[\|X_2\|^2 - E[\|X_2\|^2]]^{2+2\tau} + E[(X_1^T X_2)^{2+2\tau}] \right\}
\]

\[
\leq C_\tau L_{x,\tau},
\]

which completes the proof of Proposition 1.
A.8. Proof of Proposition 2. The essential idea of the proof is to conduct the Taylor expansion for function \((1 + x)^{1/2}\) to relate the \(L_1\)-norm to the \(L_2\)-norm. Let us define

\[
b(X_1, X_2) = \|X_1 - X_2\| - B_X^{1/2}, \quad b_1(X_1) = \mathbb{E}[b(X_1, X_2)|X_1], \quad b_1(X_2) = \mathbb{E}[b(X_1, X_2)|X_2].
\]

Since \(V^2(X) = \mathbb{E}[d^2(X_1, X_2)]\), it follows from (A.16) that

\[
V^2(X) = \mathbb{E}\{b(X_1, X_2) - b_1(X_1) - b_1(X_2) + \mathbb{E}[b(X_1, X_2)]\}^2.
\]

Then by expanding the square and the symmetry of \(X_1\) and \(X_2\), we can obtain

\[
V^2(X) = \mathbb{E}[b^2(X_1, X_2)] - 2\mathbb{E}[b^2(X_1)] + \{\mathbb{E}[b(X_1, X_2)]\}^2.
\]

Next we will bound the moments \(\mathbb{E}[b^2(X_1, X_2)], \mathbb{E}[b^2(X_1)], \text{ and } \mathbb{E}[b(X_1, X_2)]\) by resorting to the basic inequalities in Lemma 7 in Section C.7 of Supplementary Material. Denote by

\[
W_{12} = B_X^{-1}(\|X_1 - X_2\|^2 - B_X) \quad \text{and} \quad W_{13} = B_X^{-1}(\|X_1 - X_3\|^2 - B_X).
\]

Observe that \(W_{12} \geq -1, W_{13} \geq -1, \text{ and } \mathbb{E}[W_{12}] = \mathbb{E}[W_{13}] = 0\). For term \(\mathbb{E}[b(X_1, X_2)]\), by (A.59) and (A.60) we have

\[
\mathbb{E}[b(X_1, X_2)] = B_X^{1/2}\left[\mathbb{E}((1 + W_{12})^{1/2} - 1)\mathbf{1}\{W_{12} \leq 1\} + \mathbb{E}((1 + W_{12})^{1/2} - 1)\mathbf{1}\{W_{12} > 1\}\right]
\]

\[
= B_X^{1/2}\left[\frac{1}{2}\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} \leq 1\} + 2\mathbb{E}W_{12}^2\mathbf{1}\{W_{12} \leq 1\} + O_3\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} > 1\}\right]
\]

(A.18)

\[
= O_1 B_X^{1/2}\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} \leq 1\} + O_3 B_X^{1/2}\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} > 1\},
\]

where \(\mathbf{1}\{\cdot\}\) denotes the indicator function and \(O_1, O_2, O_3\) are bounded quantities such that \(|O_1| \leq 1/2, |O_2| \leq 1, \text{ and } |O_3| \leq 3/2\). Thus it follows that

\[
\{\mathbb{E}[b(X_1, X_2)]\}^2 \leq B_X\left(\frac{1}{2}\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} \leq 1\} + \frac{3}{2}\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} > 1\}\right)^2
\]

(A.19)

\[
\leq B_X\left(\frac{1}{4}\mathbb{E}[W_{12}]^3\mathbf{1}\{W_{12} \leq 1\} + \frac{15}{4}\mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} > 1\}\right).
\]

If \(\mathbb{E}[\|X\|^{4+4\tau}] < \infty\) for some \(0 < \tau \leq 1/2\), then it holds that

(A.20)

\[
\{\mathbb{E}[b(X_1, X_2)]\}^2 \leq \frac{15}{4} B_X \mathbb{E}[|W_{12}|^{2+2\tau}].
\]

Similarly, by (A.59) and (A.60) again, for \(0 < \tau \leq 1/2\) we have

\[
\mathbb{E}[b^2(X_1, X_2)] = B_X\mathbb{E}((1 + W_{12})^{1/2} - 1)^2
\]

\[
= B_X\left(\mathbb{E}\left[\frac{1}{2}W_{12} + O_5 W_{12}^2\right]\mathbf{1}\{W_{12} \leq 1\}^2 + O_4 W_{12}^2\mathbf{1}\{W_{12} > 1\}\right),
\]

where \(|O_4| \leq 1\) and \(|O_5| \leq 1/2\). Hence for \(0 < \tau \leq 1/2\), it holds that

\[
|\mathbb{E}[b^2(X_1, X_2)] - \frac{1}{4} B_X \mathbb{E}[|W_{12}|^2]| \leq \mathbb{E}\left(\frac{5}{4} \mathbb{E}[W_{12}^2]\mathbf{1}\{W_{12} > 1\} + \frac{3}{4} \mathbb{E}[|W_{12}|^3]\mathbf{1}\{W_{12} \leq 1\}\right)
\]

(A.21)

\[
\leq \frac{5}{4} B_X \mathbb{E}[|W_{12}|^{2+2\tau}].
\]
Again it follows from (A.59) and (A.60) that for $0 < \tau \leq 1/2$, we have
\[
\mathbb{E}[b^2(X_1)] = \mathbb{E}[b(X_1, X_2)b(X_1, X_3)]
= B_X\mathbb{E}\{[(1 + W_{12})^{1/2} - 1][(1 + W_{13})^{1/2} - 1]1_{\max(W_{12}, W_{13}) \leq 1}\}
+ B_X\mathbb{E}\{[(1 + W_{12})^{1/2} - 1][(1 + W_{13})^{1/2} - 1]1_{\max(W_{12}, W_{13}) > 1}\}
= B_X\left(\frac{1}{4}\mathbb{E}[W_{12}W_{13}] + O_7\mathbb{E}[W_{12}^2]1_{\max(W_{12}, W_{13}) \leq 1}\right)
+ O_8\mathbb{E}[W_{12}W_{13}]1_{\max(W_{12}, W_{13}) > 1}\right)
\]

(A.22) \[ 12B_X \mathbb{E}[W_{12}W_{13}] + O_9B_X\mathbb{E}[\|W_{12}\|^{2+2\tau}], \]
where $O_7$, $O_8$, and $O_9$ are bounded quantities satisfying $|O_7| \leq 3/4$, $|O_8| \leq 5/4$, and $|O_9| \leq 4$.

Finally by combining (A.20)–(A.22) we can deduce
\[
\mathbb{E}\left|\mathcal{V}^2(X, X) - \frac{B_X}{4}(|\mathbb{E}[W_{12}^2] - 2\mathbb{E}[W_{12}W_{13}]|)\right| \leq 9B_X\mathbb{E}[\|W_{12}\|^{2+2\tau}].
\]

Moreover, Lemma 8 in Section C.8 of Supplementary Material yields
\[
\frac{B_X}{4}\mathbb{E}[W_{12}^2 - 2\mathbb{E}[W_{12}W_{13}]] = B^{-1}_X\mathbb{E}[(X_1^TX_2)^2].
\]

It follows from (A.17) that
\[
B_X\mathbb{E}[\|W_{12}\|^{2+2\tau}] \leq B_X^{-(1+2\tau)}L_{x,\tau}.
\]
Thus the desired result (A.12) can be derived. This concludes the proof of Proposition 2.

**A.9. Proof of Proposition 3.** Similar to the proof of Proposition 2, the main idea of the proof is to conduct the Taylor expansion to relate the $L_1$-norm to the $L_2$-norm. Denote by
\[
\Delta = \mathbb{E}[b(X_1, X_2)] = \mathbb{E}[\|X_1 - X_2\| - B_X^{1/2}].
\]
In light of (A.16), we have
\[
\mathbb{E}[g(X_1, X_2, X_3, X_4)]
= \mathbb{E}[(b(X_1, X_2) - b_1(X_1) - b_1(X_2) + \Delta)](b(X_1, X_3) - b_1(X_1) - b_1(X_3) + \Delta)
\times (b(X_2, X_4) - b_1(X_2) - b_1(X_4) + \Delta)(b(X_3, X_4) - b_1(X_3) - b_1(X_4) + \Delta)]
\]
Expanding the products and noting that $X_1, X_2, X_3, X_4$ are i.i.d. random variables, we can deduce
\[
\mathbb{E}[g(X_1, X_2, X_3, X_4)] = G_1 - 4G_2 + 2G_3^2 + 4AG_4 - 4\Delta^2G_3 + \Delta^4,
\]
where
\[
G_1 = \mathbb{E}[b(X_1, X_2)b(X_1, X_3)b(X_2, X_4)b(X_3, X_4)],
G_2 = \mathbb{E}[b(X_1, X_2)b_1(X_1, X_3)b(X_2, X_4)b(X_4, X_5)],
G_3 = \mathbb{E}[b(X_1, X_2)b(X_1, X_3)],
G_4 = \mathbb{E}[b(X_1, X_2)b(X_1, X_3)b(X_2, X_4)].
\]

Next we will analyze the six terms on the right hand side of (A.24) separately. The same technique as in the proof of Proposition 2 will be used. For any $i \neq j$, let us define
\[
W_{ij} = B_X^{-1}(\|X_i - X_j\|^2 - B_X),
\]
First for term $G_1$, by definition it holds that
\[ G_1 = B_X^2 \mathbb{E} \left[ \{(1 + W_{12})^{1/2} - 1\}\{(1 + W_{13})^{1/2} - 1\}\{(1 + W_{24})^{1/2} - 1\}\{(1 + W_{34})^{1/2} - 1\} \right]. \]

Denote by
\[ D_1 = \{\text{max}(W_{12}, W_{13}, W_{24}, W_{34}) \leq 1\} \]
and $D_1^c$ the complement of $D_1$. By separating the integration region into $D_1$ and $D_1^c$ and applying (A.59) and (A.60), we can deduce
\[
G_1 = B_X^2 \mathbb{E} \left( \left[ \frac{1}{2} W_{12} + O(1)(W_{12}^2) \right] \left[ \frac{1}{2} W_{13} + O(1)(W_{13}^2) \right] \right.
\times \left[ \frac{1}{2} W_{24} + O(1)(W_{24}^2) \right] \left[ \frac{1}{2} W_{34} + O(1)(W_{34}^2) \right] \mathbb{1}\{D_1\}
\] \[ + O(1)B_X^2 \mathbb{E}\left[ W_{12}W_{13}W_{24}W_{34} \mathbb{1}\{D_1^c\} \right], \]

where $O(1)$ represents a bounded quantity satisfying $|O(1)| \leq C$ for some absolute positive constant $C$. It follows from expanding the products and Chebyshev’s inequality that if $\mathbb{E}[|X|^{4+\tau}] < \infty$ for some $0 < \tau \leq 1/2$, then we have
\[
\left| G_1 - \frac{B_X^2}{16} \mathbb{E}[W_{12}W_{13}W_{24}W_{34}] \right| \leq CB_X^2 \mathbb{E}[|W_{12}|^{1+2\tau}|W_{13}|W_{24}|W_{34}|].
\]

Further, by conditioning on $X_2, X_3$, applying the Cauchy–Schwarz inequality, and noting that $X_1, X_2, X_3, X_4$ are i.i.d. random variables, it holds that
\[
\mathbb{E}[|W_{12}|^{1+2\tau}|W_{13}|W_{24}|W_{34}|] = \mathbb{E}\left\{ \mathbb{E}(|W_{12}|^{1+2\tau}|W_{13}|X_2, X_3) \mathbb{E}(|W_{24}|W_{34}|X_2, X_3) \right\}
\leq \mathbb{E}\left\{ \left( \mathbb{E}[|W_{12}|^{2+2\tau}|X_2] \right)^{1+2\tau} \right\}^{1/1+2\tau}
\times \left( \mathbb{E}[|W_{24}|^{2+2\tau}|X_2] \right)^{1/1+2\tau}
\leq \mathbb{E}\left\{ \left( \mathbb{E}[|W_{13}|^{2+2\tau}|X_3] \right)^{1/1+2\tau} \right\}
\leq \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{1/1+2\tau}. \tag{A.25}
\]

Consequently, we have
\[
\left| G_1 - \frac{B_X^2}{16} \mathbb{E}[W_{12}W_{13}W_{24}W_{34}] \right| \leq CB_X^2 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{1/1+2\tau}. \tag{A.26}
\]

An application of the similar argument as for the proof of (A.26) yields
\[
G_2 = \frac{B_X^2}{16} \mathbb{E}[W_{12}W_{13}W_{24}W_{45}] + O(1)B_X^2 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{1/1+2\tau}. \tag{A.27}
\]

As for term $G_3$, by the same token we can deduce
\[
G_3 = B_X \mathbb{E}\left\{ (1 + W_{12})^{1/2} - 1\right\}(1 + W_{13})^{1/2} - 1\mathbb{1}\{\text{max}(W_{12}, W_{13}) \leq 1) \right\}
+ B_X \mathbb{E}\left\{ (1 + W_{12})^{1/2} - 1\right\}(1 + W_{13})^{1/2} - 1\mathbb{1}\{\text{max}(W_{12}, W_{13}) > 1) \right\}
= \frac{B_X}{4} \mathbb{E}[W_{12}W_{13}] + O(1)B_X \delta_1,
\]

where $\delta_1 = \mathbb{E}[|W_{12}|^2|W_{13}|\mathbb{1}\{\text{max}(W_{12}, W_{13}) \leq 1\}] + \mathbb{E}[|W_{12}W_{13}|\mathbb{1}\{\text{max}(W_{12}, W_{13}) > 1\}]$. Observe that when $0 < \tau \leq 1/2$, we have
\[
\delta_1 \cdot \mathbb{E}[|W_{12}|W_{13}] \leq 2\mathbb{E}[|W_{12}|^{1+2\tau}|W_{13}|]\mathbb{E}[|W_{12}W_{13}|] \leq 2\left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{4+\tau}. \tag{A.28}
\]
and
\[ \delta^2 \leq 4(\mathbb{E}[|W_{12}|^{1+\tau}|W_{13}])^2 \leq 4(\mathbb{E}[|W_{12}|^{2+2\tau}])^{\frac{2+2\tau}{1+\tau}}. \]

As a consequence, it holds that
\[ (A.28) \quad \left| G_3^2 - \frac{B_X^2}{16} \mathbb{E}[|W_{12}W_{13}|]^2 \right| \leq CB_X^2 (\mathbb{E}[|W_{12}|^{2+2\tau}])^{\frac{2+2\tau}{1+\tau}}. \]

We next deal with term \( \Delta G_4 \). It follows from (A.59) and the Cauchy–Schwarz inequality that
\[ |G_4| = B_X^{3/2} \mathbb{E}\left\{ \left[ (1 + W_{12})^{1/2} - 1 \right] \left[ (1 + W_{13})^{1/2} - 1 \right] \left[ (1 + W_{24})^{1/2} - 1 \right] \right\} \]
\[ \leq B_X^{3/2} \mathbb{E}[|W_{12}W_{13}W_{24}|] = B_X^{3/2} \mathbb{E}\{\mathbb{E}(|W_{12}W_{13}|X_2, X_3)\mathbb{E}(|W_{24}|X_2)\} \]
\[ \leq B_X^{3/2} \mathbb{E}\{\mathbb{E}[|W_{12}|X_2]^2\mathbb{E}[|W_{13}|X_3]\mathbb{E}[|W_{24}|X_2]^{1/2}\} \]
\[ = B_X^{3/2} \mathbb{E}\{\mathbb{E}[|W_{12}|X_2]\mathbb{E}[|W_{13}|X_3]\mathbb{E}[|W_{24}|X_2]^{1/2}\} \]
\[ (A.29) \quad \leq B_X^{3/2} (\mathbb{E}[|W_{12}|^2])^{3/2}. \]

Moreover, (A.18) entails that for \( 0 < \tau \leq 1/2 \), we have
\[ |\Delta| = |\mathbb{E}[b(X_1, X_2)]| \leq CB_X^{1/2} \mathbb{E}[|W_{12}|^{1+2\tau}]. \]

As a result, it follows that
\[ |\Delta G_4| \leq CB_X^2 (\mathbb{E}[|W_{12}|^2])^{3/2} \mathbb{E}[|W_{12}|^{1+2\tau}] \]
\[ (A.30) \quad \leq CB_X^2 (\mathbb{E}[|W_{12}|^{2+2\tau}])^{\frac{2+2\tau}{1+\tau}}. \]

As for term \( \Delta^2 G_3 \), note that (A.59) leads to
\[ |G_3| = B_X \mathbb{E}\{\mathbb{E}\left\{ (1 + W_{12})^{1/2} - 1 \right\} \left[ (1 + W_{13})^{1/2} - 1 \right]\} \]
\[ \leq B_X \mathbb{E}[|W_{12}W_{13}|] \leq B_X \mathbb{E}[|W_{12}|^2]. \]

It follows from (A.20) that for \( 0 < \tau \leq 1/2 \), we have
\[ \Delta^2 \leq CB_X \mathbb{E}[|W_{12}|^{2+2\tau}]. \]

Hence it holds that
\[ (A.31) \quad \Delta^2 |G_3| \leq CB_X^2 \mathbb{E}[|W_{12}|^{2+2\tau}] \mathbb{E}[|W_{12}|^2] \leq CB_X^2 (\mathbb{E}[|W_{12}|^{2+2\tau}])^{\frac{2+2\tau}{1+\tau}}. \]

Furthermore, note that (A.19) implies that for \( 0 < \tau \leq 1/2 \), we have
\[ (A.32) \quad \Delta^4  \leq CB_X^2 (\mathbb{E}[|W_{12}|^{2+\tau}])^2 \leq CB_X^2 (\mathbb{E}[|W_{12}|^{2+2\tau}])^{\frac{2+2\tau}{1+\tau}}. \]

Therefore, by substituting (A.26)–(A.28) and (A.30)–(A.32) into (A.24) we can obtain that if \( \mathbb{E}[|X|^4] < \infty \) for some \( 0 < \tau \leq 1/2 \), then
\[ \mathbb{E}[g(X_1, X_2, X_3, X_4)] = \frac{B_X^2}{16} \left\{ \mathbb{E}[W_{12}W_{13}W_{24}W_{34}] - 4\mathbb{E}[W_{12}W_{13}W_{24}W_{45}] \right. \]
\[ + \left. 2(\mathbb{E}[|W_{12}W_{13}|]^2 + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}])^{\frac{2+2\tau}{1+\tau}} \right\}. \]

Finally, the desired result (A.13) can be derived from (A.17) and Lemma 9 given in Section C.9 of Supplementary Material. This completes the proof of Proposition 3.
APPENDIX B: PROOFS OF PROPOSITIONS 1–3

B.1. Proof of Proposition 1. The desired result follows from Theorem 4. By conditions (25)–(28), it holds that
\[ B_{X}^{2\tau}L_{x\tau}/\mathbb{E}[X_{2}^{2}] \leq c_{1}c_{2}^{-((1+2\tau)/p)} \] and \[ B_{Y}^{2\tau}L_{y\tau}/\mathbb{E}[Y_{2}^{2}] \leq c_{1}c_{2}^{-((1+2\tau)/q)}. \]
Thus by Theorem 4, the fact that \( p \to \infty \) and \( q \to \infty \), and substituting the bounds in (25)–(28) into (23), we can obtain
\[
\sup_{x \in \mathbb{R}} \|P(T_{n} \leq x) - \Phi(x)\| \leq A(c_{1}, c_{2}) \left[ (pq)^{-(1+\tau)/2} + n^{-\tau} \right]^{1/(3+2\tau)},
\]
which concludes the proof of Proposition 1.

B.2. Proof of Proposition 2. The proof is based on Theorem 4 in Section 3.3 for the case of \( 0 < \tau \leq 1/2 \) and Theorem 1 in Section D.1 for the case of \( 1/2 < \tau \leq 1 \). We need to calculate the moments involved therein. The main idea is to use the block technique to deal with the \( m \)-dependent structure so that the moment inequalities for independent random variables can be applied. For simplicity, assume that \( k = p/(m_{1} + 1) \) is an integer. For \( 1 \leq r \leq k \), we define
\[ H_{r} = \{ i : (k - 1)(m_{1} + 1) + 1 \leq i \leq k(m_{1} + 1) \} \]
and
\[ S_{1,r} = \sum_{i \in H_{r}} (X_{1,i}^{2} - \mathbb{E}[X_{1,i}^{2}]), \quad S_{2,r} = \sum_{i \in H_{r}} X_{1,i}X_{2,i}. \]
By the \( m_{1} \)-dependent component structure of random vector \( X \), the odd blocks are mutually independent and so are the even blocks. Hence \( \{S_{1,r}, r \text{ is odd}\}, \{S_{1,r}, r \text{ is even}\}, \{S_{2,r}, r \text{ is odd}\}, \text{ and } \{S_{2,r}, r \text{ is even}\} \) are sequences of independent random variables with zero mean, respectively.

Let us first analyze term \( L_{x\tau} \). It holds that
\[
\mathbb{E}(\|X\|^{2} - \mathbb{E}[|X|]^{2+2\tau}) = \mathbb{E}\left( \left| \sum_{r=1}^{k} S_{1,r} \right|^{2+2\tau} \right) \leq C\left( \mathbb{E}\left[ \sum_{r: \text{odd}} S_{1,r}^{2+2\tau} \right] + \mathbb{E}\left[ \sum_{r: \text{even}} S_{1,r}^{2+2\tau} \right] \right).
\]
Then it follows from Rosenthal’s inequality that
\[
\mathbb{E}(\|X\|^{2} - \mathbb{E}[|X|]^{2+2\tau}) \leq C\left\{ \left( \sum_{r: \text{odd}} \mathbb{E}[S_{1,r}^{2}] \right)^{1+\tau} + \left( \sum_{r: \text{even}} \mathbb{E}[S_{1,r}^{2}] \right)^{1+\tau} + \sum_{r=1}^{k} \mathbb{E}[|S_{1,r}|^{2+2\tau}] \right\}.
\]
Note that for positive numbers \( s > 1 \) and \( t > 1 \) with \( s^{-1} + t^{-1} = 1 \), we have
\[
(A.33) \quad \left( \sum_{i=1}^{n} a_{i}b_{i} \right) \leq \left( \sum_{i=1}^{n} |a_{i}|^{s} \right)^{1/s} \left( \sum_{i=1}^{n} b_{i}^{t} \right)^{1/t}.
\]
Thus we can deduce
\[
\mathbb{E}[S_{1,r}^{2}] = \mathbb{E}\left[ \sum_{i \in H_{r}} (X_{1,i}^{2} - \mathbb{E}[X_{1,i}^{2}]) \right]^{2} \leq (m_{1} + 1) \sum_{i \in H_{r}} \mathbb{E}[(X_{1,i}^{2} - \mathbb{E}[X_{1,i}^{2}])^{2}] \leq (m_{1} + 1) \sum_{i \in H_{r}} \mathbb{E}[X_{1,i}^{4}].
\]
and
\[
E(|S_{1,r}|^{2+2\tau}) \leq (m_1 + 1)^{1+2\tau} \sum_{i \in H_r} E(|X_{1,i}^2 - E[X_{1,i}^2]|^{2+2\tau}) \\
\leq C(m_1 + 1)^{1+2\tau} \sum_{i \in H_r} E(|X_{1,i}|^{4+4\tau}).
\]

By plugging in the above bounds and applying (A.33), it follows that
\[
E(||X||^2 - E||X||^2)^{2+2\tau}) \leq C \{(m_1 + 1)^{1+\tau}(p/2)^{\tau} \sum_{r: odd \ i \in H_r} (E[X_{1,i}^4])^{1+\tau} \\
+ (m_1 + 1)^{1+\tau}(p/2)^{\tau} \sum_{r: even \ i \in H_r} (E[X_{1,i}^4])^{1+\tau} \\
+ (m_1 + 1)^{1+2\tau} \sum_{r=1}^k \sum_{i \in H_r} E||X_{1,i}|^{4+4\tau}\} \]
\[
(A.34) \leq C(m_1 + 1)^{1+\tau}p^{\tau} \sum_{i=1}^p E||X_{1,i}|^{4+4\tau}.
\]

In a similar fashion, we have
\[
E||X_1^T X_2||^{2+2\tau}) = E\left[||\sum_{r=1}^k S_{2,r}||^{2+2\tau}\right] \\
\leq C \left\{ \left( \sum_{r: odd \ i \in H_r} E[S_{2,r}^4]\right)^{1+\tau} + \left( \sum_{r: even \ i \in H_r} E[S_{2,r}^4]\right)^{1+\tau} \\
+ \sum_{r=1}^k E||S_{2,r}|^{2+2\tau}\right\}.
\]

In addition, it follows from the basic inequality (A.33) that
\[
E[S_{2,r}^4] \leq (m_1 + 1) \sum_{i \in H_r} E[X_{1,i}^4 X_{2,i}^4] \leq (m_1 + 1) \sum_{i \in H_r} E[X_{1,i}^4], \\
E||S_{2,r}||^{2+2\tau} \leq (m_1 + 1)^{1+2\tau} \sum_{i \in H_r} E||X_{1,i} X_{2,i}||^{2+2\tau} \\
\leq (m_1 + 1)^{1+2\tau} \sum_{i \in H_r} E||X_{1,i}|^{4+4\tau}.
\]

Thus an application of the same argument as in (A.34) results in
\[
(A.35) E||X_1^T X_2||^{2+2\tau}) \leq C(m_1 + 1)^{1+\tau}p^{\tau} \sum_{i=1}^p E||X_{1,i}|^{4+4\tau},
\]

which together with (A.34) entails that under condition (29), we have
\[
(A.36) L_{x,\tau} = E(||X||^2 - E||X||^2)^{2+2\tau}) + E||X_1^T X_2||^{2+2\tau}) \leq C_k (m_1 + 1)^{1+\tau}p^{1+\tau}.
\]
Next we deal with term \( \mathbb{E}[(X_1^T \Sigma_x X_2)^2] \). Denote by \( \sigma_{ij} \) the \((i,j)\)th entry of matrix \( \Sigma_x \). By (32) and the \( m_1 \)-dependent structure, it holds that

\[
\mathbb{E}[(X_1^T \Sigma_x X_2)^2] = \mathbb{E} \left[ \left( \sum_{i=1}^{p} \sum_{|j-i| \leq m_1} \sigma_{ij} X_{1,i} X_{2,j} \right)^2 \right]
\]

\[
= \sum_{i=1}^{p} \sum_{u=1}^{p} \sum_{|j-i| \leq m_1} \sum_{|v-u| \leq m_1} \sigma_{ij} \sigma_{uv} \mathbb{E}(X_{1,i} X_{1,u}) \mathbb{E}(X_{2,j} X_{2,v})
\]

\[
\leq \kappa_2^2 \sum_{i=1}^{p} \sum_{u=1}^{p} \sum_{|j-i| \leq m_1} \sum_{|v-u| \leq m_1} |\mathbb{E}(X_{1,i} X_{1,u})||\mathbb{E}(X_{2,j} X_{2,v})|
\]

(A.37)

\[
\leq C \kappa_2^2 (m_1 + 1)^3 p \kappa_2^2 = C \kappa_4 (m_1 + 1)^3 p.
\]

Similar results as in (A.36) and (A.37) also hold for \( Y \). That is,

(A.38) \quad L_{Y,\tau} \leq C \kappa_1 (m_2 + 1)^{1+\tau} q^{1+\tau},

(A.39) \quad \mathbb{E}[(Y_1^T \Sigma_y Y_2)^2] \leq C \kappa_4^2 (m_2 + 1)^3 q.

As a consequence, under conditions \((29)-(32)\) there exists some positive constant \( C_\kappa \) depending on \( \kappa_1, \kappa_2, \kappa_3, \) and \( \kappa_4 \) such that

\[
B_X^{-2\tau} L_{x,\tau} / \mathbb{E}[(X_1^T X_2)^2] \leq \frac{C_\kappa (m_1 + 1)^{1+\tau}}{p^\tau} \rightarrow 0,
\]

\[
B_Y^{-2\tau} L_{y,\tau} / \mathbb{E}[(Y_1^T Y_2)^2] \leq \frac{C_\kappa (m_2 + 1)^{1+\tau}}{q^\tau} \rightarrow 0,
\]

and

\[
\frac{n^{-\tau} L_{x,\tau} L_{y,\tau}}{\mathbb{E}[(X_1^T X_2)^2] \mathbb{E}[(Y_1^T Y_2)^2]} \leq \frac{C_\kappa (m_1 + 1)^{1+\tau} (m_2 + 1)^{1+\tau}}{n^\tau},
\]

\[
\frac{\mathbb{E}[(X_1^T \Sigma_x X_2)^2] + B_X^{-2\tau} L_{x,\tau}^{(2+\tau)/(1+\tau)}}{\mathbb{E}[(X_1^T X_2)^2]^2} \leq \frac{C_\kappa (m_1 + 1)^{2+\tau} p^{-\tau}}{n^\tau},
\]

\[
\frac{\mathbb{E}[(Y_1^T \Sigma_y Y_2)^2] + B_Y^{-2\tau} L_{y,\tau}^{(2+\tau)/(1+\tau)}}{\mathbb{E}[(Y_1^T Y_2)^2]^2} \leq \frac{C_\kappa (m_2 + 1)^{2+\tau} q^{-\tau}}{n^\tau}.
\]

Hence by Theorem 4, we see that (34) holds for \( 0 < \tau \leq 1/2 \).

We next prove the result for the case of \( 1/2 < \tau \leq 1 \). By the previous analysis, it holds that

\[
B_X^{-1} L_{x,1/2} / \mathbb{E}[(X_1^T X_2)^2] \leq \frac{C_\kappa (m_1 + 1)^{3/2}}{p^{1/2}} \rightarrow 0,
\]

\[
B_Y^{-1} L_{y,1/2} / \mathbb{E}[(Y_1^T Y_2)^2] \leq \frac{C_\kappa (m_2 + 1)^{3/2}}{q^{1/2}} \rightarrow 0,
\]

where the convergence to zero is by the assumption of \( m_1 = o(p^{\tau/(2+\tau)}) \) and \( m_2 = o(q^{\tau/(2+\tau)}) \). In view of Theorem 1 in Section D.1, it suffices to calculate \( \sum_{i=1}^{3} \mathcal{G}_i(X) \) and
\[ \sum_{i=1}^{3} \mathcal{G}_i(Y) \], where

\[
\begin{align*}
\mathcal{G}_1(X) &= \|E[(X_1^T X_2)^2 X_1^T \Sigma_x X_2]\|, \\
\mathcal{G}_2(X) &= E[\|X_1\|^2 (X_1^T \Sigma_x X_2)^2], \\
\mathcal{G}_3(X) &= E[X^T X X^T \Sigma_x^2 E[X X^T X]].
\end{align*}
\]

Then we analyze term \( \mathcal{G}_1(X) \). Note that

\[
\mathcal{G}_1(X) = \|E[(X_1^T X_2)^2 X_1^T \Sigma_x X_2]\| \\
\leq (\|E[(X_1^T X_2)^2]\|^{1/(1+\tau)} (\|E[(X_1^T \Sigma_x^2 X_2)^{(1+\tau)/\tau}]\|)^{(1+\tau)}).
\]

It follows from (A.35) and assumption (29) that

\[
(A.40) \quad (\|E[(X_1^T X_2)^2]\|^{1/(1+\tau)}) \leq C \kappa_1^{1/(1+\tau)} (m_1 + 1)p.
\]

Then we analyze term \( E[(X_1^T \Sigma_x^2 X_2)^{(1+\tau)/\tau}] \). Denote by \( X_{1,r} \), the \( r \)th block of \( X_1 \) for \( 1 \leq r \leq k \), and \( \Sigma_{i,j} \) the \((i,j)\)th block of \( \Sigma_x \) for \( 1 \leq i,j \leq k \). In particular, let \( \Sigma_{1,0} \) and \( \Sigma_{k,k+1} \) be zero matrices. By the \( m_1 \)-dependent structure, \( \Sigma_x \) is a tridiagonal block matrix and thus

\[
E[(X_1^T \Sigma_x^2 X_2)^{(1+\tau)/\tau}] = E\left[ \left| \sum_{r=1}^{k} S_{3,r} \right|^{(1+\tau)/\tau} \right],
\]

where

\[
S_{3,r} = (\Sigma_{r-1,X_{1,H_{r-1}}} + \Sigma_{r,X_{1,H_r}} + \Sigma_{r+1,X_{1,H_{r+1}}})^T \\
\cdot (\Sigma_{r-1,X_{2,H_{r-1}}} + \Sigma_{r,X_{2,H_r}} + \Sigma_{r+1,X_{2,H_{r+1}}}).
\]

In addition, \( \{S_{3,r}, 1 \leq r \leq k\} \) is a 3-dependent sequence. For simplicity, assume that \( k/8 \) is an integer. Then it is easy to see that \( \{\sum_{r=8(l-1)+1}^{8l} S_{3,r}, 1 \leq l \leq k/8\} \) and \( \{\sum_{r=8(l-1)+1}^{8l} S_{3,r}, 1 \leq l \leq k/8\} \) are sequences of independent random variables. Since \( 2 \leq (1+\tau)/\tau < 3 \) when \( 1/2 < \tau \leq 1 \), it follows from Rosenthal’s inequality that

\[
E[(X_1^T \Sigma_x^2 X_2)^{(1+\tau)/\tau}] \leq C E\left( \sum_{l=1}^{k/8} \sum_{r=8(l-1)+1}^{8l} S_{3,r} \right)^{1+\tau/\tau} + C E\left( \sum_{l=1}^{k/8} \sum_{r=8(l-1)+1}^{8l} S_{3,r} \right)^{1+\tau/\tau}
\]

\[
\leq C \left\{ \left( E\left( \sum_{l=1}^{k/8} \sum_{r=8(l-1)+1}^{8l} S_{3,r} \right)^2 \right)^{1+\tau/\tau} + \sum_{l=1}^{k/8} E\left( \sum_{r=8(l-1)+1}^{8l} S_{3,r} \right)^{1+\tau/\tau} \right\}
\]

Then by inequality (A.33), we can obtain

\[
E[(X_1^T \Sigma_x^2 X_2)^{(1+\tau)/\tau}] \leq C \left\{ \left( \sum_{l=1}^{k/8} \sum_{r=8(l-1)+1}^{8l} E[S_{3,r}^2] \right)^{1+\tau/\tau} + \left( \sum_{l=1}^{k/8} \sum_{r=8(l-1)+1}^{8l} E[S_{3,r}^2] \right)^{1+\tau/\tau} \right\}
\]

\[
+ \sum_{r=1}^{k} E[|S_{3,r}|^{1+\tau/\tau}]
\]
Thus it holds that

\[ E[S_{3,r}] \leq C k^{\frac{12}{7}} \left\{ \sum_{l=1}^{k/8} \sum_{r=8(l-1)+1}^{8(l-1)+4} \mathbb{E}[|S_{3,r}|] + \sum_{l=1}^{k/8} \sum_{r=8(l-1)+5}^{8l} \mathbb{E}[|S_{3,r}|] \right\} \]

\[ \leq C \left\{ p/(m_1+1) \right\}^{\frac{12}{7}} \sum_{r=1}^{k} \mathbb{E}[|S_{3,r}|]. \]

Furthermore, it holds that

\[ \mathbb{E}[|S_{3,r}|] \leq C \mathbb{E}\left[ (\| \Sigma_{r,r-1} X_{1,H_{r-1}} \|^2 + \| \Sigma_{r,r} X_{1,H_r} \|^2 + \| \Sigma_{r,r+1} X_{1,H_{r+1}} \|^2 )^{\frac{12}{7}} \right] \]

\[ \leq C \left\{ \mathbb{E}[\| \Sigma_{r,r-1} X_{1,H_{r-1}} \|^{\frac{2+2\tau}{\tau}}] + \mathbb{E}[\| \Sigma_{r,r} X_{1,H_r} \|^{\frac{2+2\tau}{\tau}}] + \mathbb{E}[\| \Sigma_{r,r+1} X_{1,H_{r+1}} \|^{\frac{2+2\tau}{\tau}}] \right\}. \]

For \( 1 \leq i,j \leq m_1 + 1 \), denote by \( \Sigma_{r,r-1}^{(i,j)} \) the \((i,j)\)th entry of \( \Sigma_{r,r-1} \) and \( X_{1,H_{r-1}}^{(i)} \) the \(i\)th component of \( X_{1,H_{r-1}} \). Observe that by assumption (32), we have

\[ \mathbb{E}[\| \Sigma_{r,r-1} X_{1,H_{r-1}} \|^{\frac{2+2\tau}{\tau}}] \]

\[ = \mathbb{E}\left[ \left\{ \sum_{i=1}^{m_1+1} \sum_{j=1}^{m_1+1} \sum_{\tau=1}^{m_1+1} \Sigma_{r,r-1}^{(i,j)} \Sigma_{r,r-1}^{(l,j)} X_{1,H_{r-1}}^{(i)} X_{1,H_{r-1}}^{(j)} \right\}^{\frac{12}{7}} \right] \]

\[ \leq (m_1 + 1)^{\frac{12}{7}} \kappa_4^{\frac{3(2-2\tau)(1+\tau)}{\tau}} \mathbb{E}\left[ \left\{ \sum_{i=1}^{m_1+1} \mathbb{E}(X_{1,H_r}^{(i)}|^2)^{\frac{2+2\tau}{\tau}} \right\} \left| X_{1,H_{r-1}} \right| \right]. \]

Moreover, it follows from (A.33) that

\[ \mathbb{E}[\| \Sigma_{r,r-1} X_{1,H_{r-1}} \|^{\frac{2+2\tau}{\tau}}] \]

\[ \leq (m_1 + 1)^{\frac{12}{7}} \kappa_4^{\frac{3(2-2\tau)(1+\tau)}{\tau}} \mathbb{E}\left[ \left\{ \sum_{i=1}^{m_1+1} \mathbb{E}(X_{1,H_r}^{(i)}|^2)^{\frac{2+2\tau}{\tau}} \right\} \left| X_{1,H_{r-1}} \right| \right]. \]

\[ \leq (m_1 + 1)^{\frac{3+2\tau}{\tau}} \kappa_4^{\frac{3(2-2\tau)(1+\tau)}{\tau}} \sum_{i=1}^{m_1+1} \mathbb{E}(X_{1,H_r}^{(i)}|^2)^{\frac{2+2\tau}{2+2\tau} + \frac{2+2\tau}{\tau}} \mathbb{E}(|X_{1,H_{r-1}}|^2)^{\frac{2+2\tau}{\tau}}. \]

Note that for any \( a > 0, b > 0, \) and \( 0 < \alpha < 1, \) we have

(A.41) \[ a^{1-\alpha} b^{\alpha} \leq a + b. \]

Thus it holds that

\[ \mathbb{E}[\| \Sigma_{r,r-1} X_{1,H_{r-1}} \|^{\frac{2+2\tau}{\tau}}] \]

\[ \leq (m_1 + 1)^{\frac{3+2\tau}{\tau}} \kappa_4^{\frac{3(2-2\tau)(1+\tau)}{\tau}} \left( \sum_{i \in H_r} \mathbb{E}[|X_{i,i}|^4] + \sum_{i \in H_{r-1}} \mathbb{E}[|X_{i,i}|^4] \right). \]
In the same manner, we can deduce

\[
\mathbb{E}[\|\Sigma_{r,r}X_1.H_r\|^{\frac{2+2r}{r}}] \leq 2(m + 1)^{\frac{3+2r}{r}} \kappa_4^{\frac{(3-2r)(1+r)}{r}} \sum_{i \in H_r} \mathbb{E}[|X_{1,i}|^{4+4r}]
\]

\[
\mathbb{E}[\|\Sigma_{r+1,r+1}X_{1,H_{r+1}}\|^{\frac{2+2r}{r}}] \leq (m + 1)^{\frac{3+2r}{r}} \kappa_4^{\frac{(3-2r)(1+r)}{r}} \left( \sum_{i \in H_r} \mathbb{E}[|X_{1,i}|^{4+4r}] + \sum_{i \in H_{r+1}} \mathbb{E}[|X_{1,i}|^{4+4r}] \right).
\]

Thus by (29), it holds that

\[
\mathbb{E}[X_1^T \Sigma_2 X_2 \frac{1+\tau}{\tau}] \leq C[p/(m + 1)]^{\frac{1-r}{2}} (m + 1)^{\frac{3+2r}{2r}} \kappa_4^{\frac{(3-2r)(1+r)}{r}} \sum_{r=1}^{k} \sum_{i \in H_r} \mathbb{E}[X_{i,H_r}^{(i)}]^{4+4r}
\]

\[
= C \kappa_4^{\frac{(3-2r)(1+r)}{r}} (m + 1)^{\frac{3+5\tau}{2r}} p^{\frac{1-r}{2r}} \sum_{i=1}^{p} \mathbb{E}[|X_{1,i}|^{4+4r}]
\]

\[
\leq C \kappa_4^{\frac{(3-2r)(1+r)}{r}} (m + 1)^{\frac{3+5\tau}{2r}} p^{\frac{1-r}{2r}},
\]

which together with (A.40) leads to

(A.42) \[ G_1(X) \leq C \kappa_1 \kappa_4^{3-2r} (m + 1)^{7/2} p^{3/2}. \]

We proceed with bounding term \( G_2(X) \). Denote by \( \sigma_{i,j} \) the \((i,j)\)th entry of matrix \( \Sigma_x \). Under the \( m_1 \)-dependent structure, we have

\[
G_2(X) = \sum_{i=1}^{p} \sum_{i=1}^{p} \sum_{u=1}^{p} \sum_{|j-i| \leq m_1} \sum_{|v-u| \leq m_1} \sigma_{i,j} \sigma_{i,v} \mathbb{E}[X_{1,i}X_1.i X_{1,u}] \mathbb{E}[X_{2,j}X_2.v].
\]

Observe that \( \mathbb{E}[X_{2,j}X_2.v] = 0 \) if \(|j - v| > m_1 \). Thus it follows that

\[
G_2(X) \leq \sum_{i=1}^{p} \sum_{i=1}^{p} \sum_{u=1}^{p} \sum_{|j-i| \leq m_1} \sum_{|v-u| \leq m_1} \sigma_{i,j} \sigma_{i,v} \mathbb{E}[X_{1,i}X_{1,i} X_{1,u}] \mathbb{E}[X_{2,j}X_2.v].
\]

By the Cauchy–Schwarz inequality, we can obtain

\[
G_2(X) \leq \kappa_4^{3-2r} \sum_{i=1}^{p} \sum_{i=1}^{p} \sum_{u=1}^{p} \sum_{|j-i| \leq m_1} \sum_{|v-u| \leq m_1} \left[ \left( \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \left( \mathbb{E}[|X_{1,j}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right]^{\frac{r}{r+2}} \times \left( \mathbb{E}[|X_{1,u}|^{4+4r}] \right)^{\frac{1}{4+4r}} \left( \mathbb{E}[|X_{1,v}|^{4+4r}] \right)^{\frac{1}{4+4r}} \times \left( \sum_{|j-i| \leq m_1} \left( \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right)^{\frac{r}{r+2}} \times \left( \sum_{|v-u| \leq m_1} \left( \mathbb{E}[|X_{1,u}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right)^{\frac{r}{r+2}}
\]

\[
\leq \kappa_4^{3-2r} \left( \sum_{i=1}^{p} \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \left( \sum_{i=1}^{p} \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \times \left( \sum_{|j-i| \leq m_1} \left( \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right)^{\frac{r}{4}} \times \left( \sum_{|v-u| \leq m_1} \left( \mathbb{E}[|X_{1,u}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right)^{\frac{r}{4+4r}}
\]

Further, by the basic inequality (A.33) it holds that

\[
G_2(X) \leq \kappa_4^{3-2r} \left( \sum_{i=1}^{p} \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \left( \sum_{i=1}^{p} \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \times \left( \sum_{|j-i| \leq m_1} \left( \mathbb{E}[|X_{1,i}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right)^{\frac{r}{4}} \times \left( \sum_{|v-u| \leq m_1} \left( \mathbb{E}[|X_{1,u}|^{4+4r}] \right)^{\frac{1}{4+4r}} \right)^{\frac{r}{4+4r}}
\]
Thus the \( g_r \) is given by

\[
\sum_{i=1}^p \left( \sum_{|j-i| \leq 2m_1} (\mathbb{E}[|X_{1,j}|^{4+4\tau}]) \frac{1}{2+\tau} \right) \left( \sum_{|u-i| \leq 3m_1} (\mathbb{E}[|X_{1,u}|^{4+4\tau}]) \frac{1}{2+\tau} \right) \leq C \kappa_4^{3-2\tau} (m_1 + 1)^{\frac{1+\tau}{1+2\tau}} \left( \sum_{i=1}^p (\mathbb{E}[|X_{1,i}|^{4+4\tau}]) \right)^{1/2} \frac{1+\tau}{2+\tau}
\]

Hence it follows from the basic inequality (A.41) that

\[
(\mathbb{E}[|X_{1,j}|^{4+4\tau}]) \frac{1}{2+\tau} (\mathbb{E}[|X_{1,u}|^{4+4\tau}]) \frac{1}{2+\tau} \leq (\mathbb{E}[|X_{1,j}|^{4+4\tau}] + \mathbb{E}[|X_{1,u}|^{4+4\tau}]) \frac{1+\tau}{2+\tau}
\]

which together with (A.33) and assumption (29) yields

\[
\mathcal{G}_2(X) \leq C \kappa_2^{3-2\tau} (m_1 + 1)^{\frac{1+\tau}{1+2\tau}} \left( \sum_{i=1}^p (\mathbb{E}[|X_{1,i}|^{4+4\tau}]) \right)^{1/2} \frac{1+\tau}{2+\tau}
\]

\[
\leq C \kappa_2^{3-2\tau} (m_1 + 1)^{3} \left( \sum_{i=1}^p (\mathbb{E}[|X_{1,i}|^{4+4\tau}]) \right)^{1/2} \frac{1+\tau}{2+\tau}
\]

(A.43) \[\kappa_2^{3-2\tau} (m_1 + 1)^{3} \sum_{i=1}^p (\mathbb{E}[|X_{1,i}|^{4+4\tau}]) \leq C \kappa_4^{3-2\tau} (m_1 + 1)^3 p^2.
\]

As for term \( \mathcal{G}_3(X) \), we exploit similar arguments. It is easy to see that the \( r \)th block of \( \mathbb{E}[X_1^T X_1 X_1^T] \) is given by

\[
(\mathbb{E}[X_1^T X_1 X_1^T])^{(r)} = \mathbb{E} [X_1 H_r (\|X_1 H_{r-1}\|^2 + \|X_1 H_r\|^2 + \|X_1 H_{r+1}\|^2)].
\]

Thus the \( r \)th block of \( \Sigma_x \mathbb{E}[X_1^T X_1 X_1^T] \) is

\[
(\Sigma_x \mathbb{E}[X_1^T X_1 X_1^T])^{(r)} = \Sigma_{r,r-1} \mathbb{E} [X_1 H_{r-1} (\|X_1 H_{r-2}\|^2 + \|X_1 H_{r-1}\|^2 + \|X_1 H_r\|^2)] + \Sigma_{r,r} \mathbb{E} [X_1 H_r (\|X_1 H_{r-1}\|^2 + \|X_1 H_r\|^2 + \|X_1 H_{r+1}\|^2)] + \Sigma_{r,r+1} \mathbb{E} [X_1 H_{r+1} (\|X_1 H_r\|^2 + \|X_1 H_{r+1}\|^2 + \|X_1 H_{r+2}\|^2)].
\]

Then it follows that

\[
\mathcal{G}_3(X) = \|\Sigma_x \mathbb{E}[X_1 X_1^T X_1]\|_2^2 = \sum_{r=1}^k \| (\Sigma_x \mathbb{E}[X_1 X_1^T X_1])^{(r)} \|_2^2
\]

\[
\leq C \sum_{r=1}^k \sum_{u \in \{r-1, r, r+1\}} \left\{ \| \Sigma_{r,u} \mathbb{E}(X_1 H_u \|X_1 H_{u-1}\|^2) \|_2^2 + \| \Sigma_{r,u} \mathbb{E}(X_1 H_u \|X_1 H_{u+1}\|^2) \|_2^2 \right\}.
\]

(A.44)
In fact, the terms on the right hand side of the above inequality share the same bounds. Thus we show the analysis only for the first term.

Observe that
\[
\sum_{r=1}^{k} \| \Sigma_{r,r-1} \mathbb{E}(X_{1,H_{r-1}} \| X_{1,H_{r-2}} \|^2) \|^2
\]
\[
= \sum_{r=1}^{k} \sum_{i \in H_r} \sum_{j \in H_r} \sum_{l \in H_{r-1}} \Sigma_{r,r-1}^{(i,l)} \Sigma_{r-1,r}^{(l,j)} \mathbb{E}[X_{1,H_{r-1}} \| X_{1,H_{r-2}} \|^2] \mathbb{E}[X_{1,H_{r-1}} \| X_{1,H_{r-2}} \|^2].
\]

Then it follows from the Cauchy–Schwarz inequality, assumption (32), and the basic inequality (A.33) that
\[
\sum_{r=1}^{k} \| \Sigma_{r,r-1} \mathbb{E}(X_{1,H_{r-1}} \| X_{1,H_{r-2}} \|^2) \|^2 \\
\leq \kappa^3 \sum_{r=1}^{k} \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^2] \right)^{2r-1} \left( \sum_{j \in H_r} \mathbb{E}[X_{1,j}^2] \right)^{2r-1} \mathbb{E}[\|X_{1,H_r}\|^2] \mathbb{E}[\|X_{1,H_{r-2}}\|^4]
\]
\[
\leq \kappa^3 \sum_{r=1}^{k} \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^2] \right)^{2r-1} \left( \sum_{j \in H_r} \mathbb{E}[X_{1,j}^2] \right)^{2r-1} \mathbb{E}[\|X_{1,H_r}\|^2] \mathbb{E}[\|X_{1,H_{r-2}}\|^4].
\]

Moreover, note that for any \( a, b > 0 \), we have
\[
ab^2 \leq a^3 + b^3.
\]

Thus in light of (A.33), we can obtain
\[
\mathbb{E}[\|X_{1,H_r}\|^2] \leq \mathbb{E}[\|X_{1,H_r}\|^6] + \mathbb{E}[\|X_{1,H_{r-2}}\|^6]
\]
\[
\leq (m_1 + 1)^2 \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^6] + \sum_{i \in H_{r-2}} \mathbb{E}[X_{1,i}^6] \right).
\]

Furthermore, it follows from (A.33) that
\[
\left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^2] \right)^{2r-1} \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^6] \right)
\]
\[
\leq 2(m_1 + 1) \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^{4+4r}] \right)^{2r-1} \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^{4+4r}] \right)^{2r-1}
\]
\[
\leq 4(m_1 + 1) \left( \sum_{i \in H_r} \mathbb{E}[X_{1,i}^{4+4r}] \right) + \sum_{i \in H_r} \mathbb{E}[X_{1,i}^{4+4r}].
\]

Similarly, we can deduce
\[
\left( \sum_{i \in H_{r-1} \cup H_r} \mathbb{E}[X_{1,i}^2] \right)^{2r-1} \left( \sum_{i \in H_{r-2}} \mathbb{E}[X_{1,i}^6] \right)
\]
Combining (A.42), (A.43), and (A.45), and noting that 
\[ m \]
similar way. Thus we have
\[ \text{In the same manner, we can also show that} \]
\[ (A.46) \]
with the transformed random vectors
\[ \tilde{X} \]
Euclidean norm is invariant to orthogonal transformations. Thus
\[ X \]
Proposition 2.
\[ \text{Consequently, it holds that} \]
\[ \sum_{r=1}^{k} \|\Sigma_{r,r-1} E(X_{1,H_{r-1}} \|X_{1,H_{r-2}}\|^2)\|^{2} \]
\[ \leq C \kappa_4^{3-2\tau} (m_1 + 1)^4 \sum_{r=1}^{k} \sum_{i \in H_r} E[|X_{1,i}|^{4+4\tau}] \]
\[ = C \kappa_4^{3-2\tau} (m_1 + 1)^4 \sum_{i=1}^{p} E[|X_{1,i}|^{4+4\tau}] \leq C \kappa_1 \kappa_4^{3-2\tau} (m_1 + 1)^4 p. \]

For the other terms on the right hand side of (A.44), the same bound can be derived in a similar way. Thus we have
\[ (A.45) \]
\[ G_3(X) \leq C \kappa_1 \kappa_4^{3-2\tau} (m_1 + 1)^4 p. \]
Combining (A.42), (A.43), and (A.45), and noting that \( m_1 + 1 \leq p \), we can obtain
\[ (A.46) \]
\[ G_1(X) + G_2(X) + G_3(X) \leq C \tau \kappa_1 \kappa_4^{3-2\tau} (m_1 + 1)^3 p^2. \]
In the same manner, we can also show that
\[ (A.47) \]
\[ G_1(Y) + G_2(Y) + G_3(Y) \leq C \tau \kappa_1 \kappa_4^{3-2\tau} (m_2 + 1)^3 q^2. \]
Hence (34) follows from substituting (A.36)–(A.39) and (A.46)–(A.47) into Theorem 1. Then we can see that when \( m_1 \) and \( m_2 \) satisfy (33), \( T_n \xrightarrow{p} N(0,1) \). This completes the proof of Proposition 2.

B.3. Proof of Proposition 3. Assume that \( \Sigma_x = \Gamma_1^T \text{diag}(\lambda_{1X}^1, \ldots, \lambda_{pX}^X) \Gamma_1 \) and \( \Sigma_y = \Gamma_2^T \text{diag}(\lambda_{1Y}^1, \ldots, \lambda_{qY}^Y) \Gamma_2 \) for some orthogonal matrices \( \Gamma_1 \) and \( \Gamma_2 \). A useful fact is that the Euclidean norm is invariant to orthogonal transformations. Thus \( X \) and \( Y \) can be replaced with the transformed random vectors \( \tilde{X} = \Gamma_1 X \) and \( \tilde{Y} = \Gamma_2 Y \), respectively. Clearly the transformed random vectors are distributed as
\[ \tilde{X} \sim N(0, \text{diag}(\lambda_{1X}^1, \ldots, \lambda_{pX}^X)) \] and \( \tilde{Y} \sim N(0, \text{diag}(\lambda_{1Y}^1, \ldots, \lambda_{qY}^Y)) \).

It is equivalent to analyze the distance correlation between the new multivariate normal random variables \( \tilde{X} \) and \( \tilde{Y} \). It is easy to show that
\[ \max_{1 \leq i \leq p} E[\tilde{X}_{1,i}^2] \leq a_2, \quad p^{-1} \sum_{i=1}^{p} E[|\tilde{X}_{1,i}|^{8}] \leq C a_2^4, \quad p^{-1} E[(\tilde{X}_{1,i}^T \tilde{X}_{2})^2] \geq a_1^2, \quad p^{-1} B_X \geq a_1. \]

Similar bounds also hold for \( Y \). Then the conditions of Proposition 2 are satisfied and the independence of coordinates entails that \( m_1 = m_2 = 0 \). Therefore, the desired result can be derived by applying Proposition 2 with \( \tau = 1 \) and \( m_1 = m_2 = 0 \). This concludes the proof of Proposition 3.
APPENDIX C: SOME KEY LEMMAS AND THEIR PROOFS

C.1. Lemma 1 and its proof.

**Lemma 1.** Under condition (18), we have
\[
\mathbb{V}_n^*(X)/\mathbb{V}^2(X) \to 1 \quad \text{in probability}
\]
and
\[
\mathbb{V}_n^*(Y)/\mathbb{V}^2(Y) \to 1 \quad \text{in probability}
\]
as \( n \to \infty. \)

**Proof.** For any \( X \) and \( Y \), since \( \mathbb{V}_n^*(X,Y) \) is a U-statistic and noting that \( \mathbb{E}[\mathbb{V}_n^*(X,Y)] = \mathbb{V}^2(X,Y) \) by (15), it follows from the moment inequality of U-statistics (Koroljuk and Borovskich, 1994, p. 72) and conditional Jensen’s inequality that for \( 0 < \tau \leq 1 \),
\[
\mathbb{E}\left[|\mathbb{V}_n^*(X,Y) - \mathbb{V}^2(X,Y)|^{1+\tau}\right] 
\leq C \sum_{i=1}^{4} \left( \begin{array}{c} 4 \\ i \end{array} \right)^{1+\tau} \left( \frac{n}{i} \right)^{-\tau} \mathbb{E}\left[|h((X_1,Y_1),(X_2,Y_2),(X_3,Y_3),(X_4,Y_4))|^{1+\tau}\right]
\]
(\[A.50\])
\[
\leq Cn^{-\tau}\mathbb{E}\left[|h((X_1,Y_1),(X_2,Y_2),(X_3,Y_3),(X_4,Y_4))|^{1+\tau}\right].
\]
In fact, the moment of \( h((X_1,Y_1),(X_2,Y_2),(X_3,Y_3),(X_4,Y_4)) \) can be dominated by that of
\[
d(X_1,X_2)d(Y_1,Y_2)
\]
based on the expression given in Lemma 5 in Section C.5.

By choosing \( X = Y \) in (A.58) and the Cauchy–Schwarz inequality, we can obtain that for
\[ 0 < \tau \leq 1, \]
\[
\mathbb{E}[|h((X_1,X_1),(X_2,X_2),(X_3,X_3),(X_4,X_4))|^{1+\tau}] \leq C\mathbb{E}(|d(X_1,X_2)|^{2+2\tau}].
\]
Thus it follows from (A.50) that
\[
\mathbb{E}[|\mathbb{V}_n^*(X)/\mathbb{V}^2(X) - 1|^{1+\tau}] \leq \frac{C\mathbb{E}[|d(X_1,X_2)|^{2+2\tau}]}{n^{\tau}|\mathbb{V}^2(X)|^{1+\tau}},
\]
(\[A.51\])
Moreover, since \( \mathbb{E}[|d(Y_1,Y_2)|^{2+2\tau}] \geq (\mathbb{E}[\mathbb{V}^2(Y_1,Y_2)])^{1+\tau} = |\mathbb{V}^2(Y)|^{1+\tau} \), it follows from condition (18) that
\[
\frac{\mathbb{E}[|d(X_1,X_2)|^{2+2\tau}]}{n^{\tau}|\mathbb{V}^2(X)|^{1+\tau}} \to 0,
\]
which yields the ratio consistency (A.48). The result in (A.49) can be obtained similarly. This completes the proof of Lemma 1.

C.2. Lemma 2 and its proof.

**Lemma 2.** If \( \mathbb{E}[||X||^2] + \mathbb{E}[||Y||^2] < \infty \) and \( X \) is independent of \( Y \), then we have
\[
\tilde{T}_n = W_n^{(1)}(X,Y) + W_n^{(2)}(X,Y),
\]
where
\[
W_n^{(1)}(X,Y) = \sqrt{\frac{2}{n(n-1)}} \sum_{1 \leq i < j \leq n} \frac{d(X_i,X_j)d(Y_i,Y_j)}{\sqrt{\mathbb{V}^2(X)\mathbb{V}^2(Y)}},
\]
and \( W_n^{(2)}(X,Y) \) satisfies \( \mathbb{E}(|W_n^{(2)}(X,Y)|^2) \leq Cn^{-1}. \)
Proof. Recall that $V_n(X, Y)$ is a U-statistic and

$$V_n(X, Y) = \left( \binom{n}{2} \right)^{-1} \sum_{1 \leq i_1 < i_2 < i_3 < i_4 \leq n} h((X_{i_1}, Y_{i_1}), \ldots, (X_{i_4}, Y_{i_4})).$$

It has been shown in Huang and Huo (2017) that under the independence of $X$ and $Y$,

$$\mathbb{E}[h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))|(X_1, Y_1)] = 0,$$

$$\mathbb{E}[h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))|(X_1, Y_1), (X_2, Y_2)] = \frac{1}{6} d(X_1, X_2) d(Y_1, Y_2).$$

Thus by the Hoeffding decomposition (e.g. Koroljuk and Borovskich, 1994, p. 23) and dispersion for U-statistics (Koroljuk and Borovskich, 1994, p. 31), when $X$ is independent of $Y$ we have

$$V_n(X, Y) = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} d(X_i, X_j) d(Y_i, Y_j) + U_n(X, Y),$$

where

$$\mathbb{E}[U_n^2(X, Y)] \leq \frac{C}{n^3} \mathbb{E}[h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))]^2.$$

Furthermore, Lemma 6 given in Section C.6 yields

$$\mathbb{E}[U_n^2(X, Y)] \leq \frac{CV^2(X)V^2(Y)}{2n^3}.$$

Hence it follows from (A.1) that

$$\mathbb{E}[W_n^{(2)}(X, Y)]^2 = \frac{n(n - 1) \mathbb{E}[U_n^2(X, Y)]}{2V^2(X)V^2(Y)} \leq \frac{C}{n},$$

which concludes the proof of Lemma 2.

C.3. Lemma 3 and its proof.

Lemma 3. Let $\mathscr{F}_k = \sigma\{(X_1, Y_1), \ldots, (X_k, Y_k)\}$ be a $\sigma$-algebra. Then $\{(\zeta_{n,k}, \mathscr{F}_k), k \geq 1\}$ forms a martingale difference array under the independence of $X$ and $Y$, where $\zeta_{n,k}$ is defined in (A.2).

Proof. It is easy to see that $\zeta_{n,k} \in \mathscr{F}_k$ and when $X$ is independent of $Y$,

$$\mathbb{E} \left[ \sum_{i=1}^{k-1} d(X_i, X_k) d(Y_i, Y_k) \big| \mathscr{F}_{k-1} \right] = \sum_{i=1}^{k-1} \mathbb{E} \left[ d(X_i, X_k) \big| X_i \right] \mathbb{E} \left[ d(Y_i, Y_k) \big| Y_i \right] = 0,$$

where the last equality is due to $\mathbb{E} \left[ d(X_1, X_2) \big| X_1 \right] = 0$ and $\mathbb{E} \left[ d(X_1, X_2) \big| X_2 \right] = 0.$

C.4. Lemma 4 and its proof.

Lemma 4. If $\mathbb{E}[\|X\|^{2+2\tau}] + \mathbb{E}[\|Y\|^{2+2\tau}] < \infty$ for some constant $0 < \tau \leq 1$ and $X$ is independent of $Y$, then we have

$$\mathbb{E} \left( \sum_{k=1}^{n} \mathbb{E}[\zeta_{n,k}^2] \big| \mathscr{F}_{k-1} \right] - \left( 1 + \tau \right) \leq C \left( \frac{\mathbb{E}[g(X_1, X_2, X_3, X_4)] \mathbb{E}[g(Y_1, Y_2, Y_3, Y_4)]}{[V^2(X)V^2(Y)]^{1+\tau}} \right)^{(1+\tau)/2}$$

$$+ \frac{C \mathbb{E}[d(X_1, X_2)]^{2+2\tau} \mathbb{E}[d(Y_1, Y_2)]^{2+2\tau}}{n^\tau [V^2(X)V^2(Y)]^{1+\tau}}$$

(A.53)
and

\[
\sum_{k=1}^{n} \mathbb{E}[|\zeta_{n,k}|^{2+2\tau}] \leq \frac{C\mathbb{E}[|d(X_1, X_2)|^{2+2\tau}]\mathbb{E}[|d(Y_1, Y_2)|^{2+2\tau}]}{n^{\tau}|V^2(X)V^2(Y)|^{1+\tau}}.
\]

**Proof.** (i) We first prove (A.53). Recall the definition of \(\zeta_{n,k}\) in (A.2). Note that under the independence of \(X\) and \(Y\), we have

\[
\sum_{k=1}^{n} \mathbb{E}[|\zeta_{n,k}|^{2+2\tau}] = 2\sum_{k=1}^{n} \mathbb{E}\left( \frac{\left( \sum_{i=1}^{n-1} d(X_i, X_k) d(Y_i, Y_k) \right)^2}{n(n-1)|V^2(X)V^2(Y)|} \right) := R_n^{(1)} + R_n^{(2)},
\]

where \(R_n^{(1)}\) is the sum of squared terms given by

\[
R_n^{(1)} = 2\sum_{k=1}^{n} \sum_{i=1}^{k-1} \mathbb{E} \left[ d^2(X_i, X_k) d(Y_i, Y_k) | X_i, X_k \right] \mathbb{E} \left[ d^2(Y_i, Y_k) | Y_i \right],
\]

and \(R_n^{(2)}\) is the sum of cross-product terms given by

\[
R_n^{(2)} = 4\sum_{k=1}^{n} \sum_{1 \leq i < j \leq k-1} \mathbb{E}\left[ d(X_i, X_j) d(X_j, X_k) | X_i, X_j \right] \mathbb{E}\left[ d(Y_j, Y_k) | Y_j \right].
\]

Thus it holds that

\[
\mathbb{E}\left( \sum_{k=1}^{n} \mathbb{E}[|\zeta_{n,k}|^{2+2\tau}] \right)^{1+\tau} \leq C(\mathbb{E}[|R_n^{(1)}|^{1+\tau}] + \mathbb{E}[|R_n^{(2)}|^{1+\tau}]).
\]

We first bound term \(\mathbb{E}[|R_n^{(2)}|^{1+\tau}]\). Let \((X', Y')\) be an independent copy of \((X, Y)\) that is independent of \((X_1, Y_1), \ldots, (X_n, Y_n)\). For notational simplicity, define

\[
\eta_1(X_i, X_j) = \mathbb{E}\left[ d(X_i, X)d(X_j, X) | X_i, X_j \right] \quad \text{and} \quad \eta_2(Y_i, Y_j) = \mathbb{E}\left[ d(Y_i, Y)d(Y_j, Y) | Y_i, Y_j \right].
\]

By changing the order of summation, we can obtain

\[
\mathbb{E}[|R_n^{(2)}|^2] = \frac{16\mathbb{E}\left[ \sum_{1 \leq i < j \leq n} \sum_{k \geq j+1} \eta_1(X_i, X_j) \eta_2(Y_i, Y_j) \right]^2}{n^2(n-1)^2|V^2(X)V^2(Y)|^2} = \frac{16\mathbb{E}\left[ \sum_{1 \leq i < j \leq n} \eta_1(X_i, X_j) \eta_2(Y_i, Y_j) \right]^2}{n^2(n-1)^2|V^2(X)V^2(Y)|^2}.
\]

In addition, for pairwisely nonequal \(i, j, l\), it holds that

\[
\mathbb{E}[\eta_1(X_i, X_j) \eta_1(X_i, X_l)] = \mathbb{E}\{ \mathbb{E}[d(X_i, X)d(X_j, X)d(X_i, X')d(X_i, X')d(X_i, X_l)] \} = \mathbb{E}[d(X_i, X)d(X_j, X)d(X_i, X')d(X_i, X')]
\]

\[
= \mathbb{E}\left\{ \mathbb{E}[d(X_i, X)d(X_j, X)d(X_i, X')d(X_i, X') | X, X'] \right\}
\]

\[
= \mathbb{E}\left\{ \mathbb{E}[d(X_i, X)d(X_i, X') | X, X'] \mathbb{E}[d(X_j, X') | X'] \mathbb{E}[d(X_i, X') | X'] \right\}
\]

\[
= 0,
\]

where we have used the fact that \(\mathbb{E}[d(X_j, X') | X'] = \mathbb{E}[d(X_i, X') | X'] = 0\).

It is easy to see that \(\mathbb{E}[\eta_1(X_i, X_j)] = \mathbb{E}[\eta_2(Y_i, Y_j)] = 0\) for \(i \neq j\). Thus for pairwisely nonequal \(i, j, k, l\), it holds that

\[
\mathbb{E}[\eta_1(X_i, X_j) \eta_1(X_k, X_l)] = 0.
\]
Then the cross-product terms in the numerator of $\mathbb{E}([R_n^{(2)}]^2)$ vanish. Moreover, in view of the definition of $g(X_1, X_2, X_3, X_4)$ in (17), we have

$$
\mathbb{E}[\eta_1(X_i, X_j)]^2 = \mathbb{E}\{\mathbb{E}[d(X_i, X) d(X_j, X) d(X_j, X')] d(X_j, X') | X_i, X_j]\}
= \mathbb{E}\{d(X_1, X_2) d(X_1, X_3) d(X_2, X_4) d(X_3, X_4)\} = \mathbb{E}[g(X_1, X_2, X_3, X_4)].
$$

Consequently, it follows that

$$
\mathbb{E}([R_n^{(2)}]^2) = \frac{16 \sum_{1 \leq i < j \leq n} (n - j)^2 \mathbb{E}[\eta_1(X_i, X_j)]^2 \mathbb{E}[\eta_2(Y, Y_j)]^2}{n^2 (n - 1)^2 \mathbb{V}^2(X) \mathbb{V}^2(Y)} \leq \frac{16 \sum_{j=1}^n (j - 1)(n - j)^2 \mathbb{E}[g(X_1, X_2, X_3, X_4)] \mathbb{E}[g(Y_1, Y_2, Y_3, Y_4)]}{n^2 (n - 1)^2 \mathbb{V}^2(X) \mathbb{V}^2(Y)} \leq C \mathbb{E}[g(X_1, X_2, X_3, X_4)] \mathbb{E}[g(Y_1, Y_2, Y_3, Y_4)] \frac{1}{\mathbb{V}^2(X) \mathbb{V}^2(Y)}. $$

Hence we can obtain

$$(A.56) \quad \mathbb{E}([R_n^{(2)}]|^{1 + \tau}) \leq \left( \frac{C \mathbb{E}[g(X_1, X_2, X_3, X_4)] \mathbb{E}[g(Y_1, Y_2, Y_3, Y_4)]}{\mathbb{V}^2(X) \mathbb{V}^2(Y)} \right)^{(1 + \tau)/2}.$$

Next we deal with term $\mathbb{E}([R_n^{(1)}] - 1|^{1 + \tau})$. Since $\mathbb{E}[d^2(X_1, X_2)] = \mathbb{V}^2(X)$, clearly when $X$ is independent of $Y$, we have

$$
\mathbb{E}[R_n^{(1)}] = \frac{2 \sum_{k=1}^n \sum_{i=1}^{k-1} \mathbb{V}^2(X) \mathbb{V}^2(Y)}{n(n - 1) \mathbb{V}^2(X) \mathbb{V}^2(Y)} = 1.
$$

For simplicity, denote by $\eta_3(X_i, Y_i) = \mathbb{E}[d^2(X_i, X) | X_i] \mathbb{E}[d^2(Y_i, Y) | Y_i]$. Then by changing the order of summation, we deduce

$$
\mathbb{E}([R_n^{(1)}] - 1|^{1 + \tau}) = \frac{\mathbb{E}\left[\left[ 2 \sum_{k=1}^n \sum_{i=1}^{k-1} [\eta_3(X_i, Y_i) - \mathbb{E}\eta_3(X_i, Y_i)] \right]_{1 + \tau} \right]}{\mathbb{V}^2(X) \mathbb{V}^2(Y)}
= \frac{\mathbb{E}\left[\left[ 2 \sum_{i=1}^n (n - i) [\eta_3(X_i, Y_i) - \mathbb{E}\eta_3(X_i, Y_i)] \right]_{1 + \tau} \right]}{\mathbb{V}^2(X) \mathbb{V}^2(Y)}.
$$

Then it follows from the von Bahr–Esseen inequality (Lin and Bai, 2010, p. 100) for independent random variables that when $0 < \tau \leq 1$,

$$
\mathbb{E}([R_n^{(1)}] - 1|^{1 + \tau}) \leq \frac{C \sum_{i=1}^n (n - i)^{1 + \tau} \mathbb{E}[\eta_3(X_i, Y_i)]^{1 + \tau}}{n^2 \mathbb{V}^2(X) \mathbb{V}^2(Y)}
\leq \frac{C \mathbb{E}[\eta_3(X_i, Y_i)]^{1 + \tau}}{n \mathbb{V}^2(X) \mathbb{V}^2(Y)^{1 + \tau}} \leq \frac{C \mathbb{E}[d(X_1, X_2)^{2 + 2\tau}] \mathbb{E}[d(Y_1, Y_2)^{2 + 2\tau}]}{n \mathbb{V}^2(X) \mathbb{V}^2(Y)^{1 + \tau}}
$$

which along with (A.55) and (A.56) leads to (A.53).

(ii) We now show (A.54). Note that

$$
(A.57) \quad \sum_{k=1}^n \mathbb{E}[|\zeta_{n,k}|^{2 + 2\tau}] = \frac{2^{1 + \tau} \sum_{k=1}^n \mathbb{E}\left[\left| \sum_{i=1}^{k-1} d(X_i, X_k) d(Y_i, Y_k) \right|^{2 + 2\tau} \right]}{n(n - 1) \mathbb{V}^2(X) \mathbb{V}^2(Y)^{1 + \tau}}.
$$
Given \((X_k, Y_k), \{d(X_i, X_k)d(Y_i, Y_k), 1 \leq i \leq k - 1\}\) is a sequence of independent random variables and under the independence of \(X\) and \(Y\),
\[
\mathbb{E}[d(X_i, X_k)d(Y_i, Y_k)|X_k, Y_k] = \mathbb{E}[d(X_i, X_k)|X_k]\mathbb{E}[d(Y_i, Y_k)|Y_k] = 0.
\]
Thus it follows from Rosenthal’s inequality for independent random variables that
\[
\mathbb{E} \left[ \left| \sum_{i=1}^{k-1} d(X_i, X_k)d(Y_i, Y_k) \right|^{2+2\tau} \right] = \mathbb{E} \left[ \left( \sum_{i=1}^{k-1} d(X_i, X_k)d(Y_i, Y_k) \right)^{2+2\tau} \mathbb{E}[X_k, Y_k] \right] 
\leq C\mathbb{E} \left[ \left( \sum_{i=1}^{k-1} d(X_i, X_k)d(Y_i, Y_k) \right)^{2} \mathbb{E}[X_k, Y_k] \right]^{1+\tau} 
+ C(k-1)\mathbb{E}(\|d(X_1, X_2)\|^{2+2\tau})\mathbb{E}(\|d(Y_1, Y_2)\|^{2+2\tau}).
\]
Since given \((X_k, Y_k), \{d(X_i, X_k)d(Y_i, Y_k), 1 \leq i \leq k - 1\}\) is a sequence of independent random variables with zero means under the independence of \(X\) and \(Y\), it is easy to see that
\[
\mathbb{E} \left( \left| \sum_{i=1}^{k-1} d(X_i, X_k)d(Y_i, Y_k) \right|^{2} \mathbb{E}[X_k, Y_k] \right) = \sum_{i=1}^{k-1} \mathbb{E}[d^2(X_i, X_k)d^2(Y_i, Y_k)|X_k, Y_k],
\]
\[
= (k-1)\mathbb{E}[d^2(X, X_k)|X_k]\mathbb{E}[d^2(Y, Y_k)|Y_k].
\]
Then it follows from the conditional Jensen’s inequality that when \(X\) is independent of \(Y\),
\[
\mathbb{E} \left[ \left( \sum_{i=1}^{k-1} d(X_i, X_k)d(Y_i, Y_k) \right)^{2+2\tau} \mathbb{E}[X_k, Y_k] \right]^{1+\tau} 
\leq (k-1)^{1+\tau}\mathbb{E}[\|d(X_1, X_2)\|^{2+2\tau}]\mathbb{E}[\|d(Y_1, Y_2)\|^{2+2\tau}].
\]
Finally we can obtain
\[
\sum_{k=1}^{n} \mathbb{E} \left[ \left| \sum_{i=1}^{k-1} d(X_i, X_k)d(Y_i, Y_k) \right|^{2+2\tau} \right] \leq C\sum_{k=1}^{n}(k-1)^{1+\tau}\mathbb{E}[\|d(X_1, X_2)\|^{2+2\tau}]\mathbb{E}[\|d(Y_1, Y_2)\|^{2+2\tau}]
\leq Cn^{2+\tau}\mathbb{E}[\|d(X_1, X_2)\|^{2+2\tau}]\mathbb{E}[\|d(Y_1, Y_2)\|^{2+2\tau}].
\]
Substituting the above bound into (A.57) results in (A.54). This completes the proof of Lemma 4.

C.5. Lemma 5 and its proof. The following lemma provides a useful representation of the kernel function \(h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))\) in terms of the double-centered distance \(d(\cdot, \cdot)\).

**Lemma 5.** For any random vectors \(X\) and \(Y\) with finite first moments, we have
\[
h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))
= \frac{1}{4} \sum_{1 \leq i,j \leq 4, \ i \neq j} d(X_i, X_j)d(Y_i, Y_j) - \frac{1}{4} \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, \ j \neq i} d(X_i, X_j) \sum_{1 \leq j \leq 4, \ j \neq i} d(Y_i, Y_j) \right)
+ \frac{1}{24} \sum_{1 \leq i,j \leq 4, \ i \neq j} d(X_i, X_j) \sum_{1 \leq i,j \leq 4, \ i \neq j} d(Y_i, Y_j).
\]
Proof. Let us define
\[ a_1(X_1, X_2) = \|X_1 - X_2\| - E[\|X_1 - X_2\|], \quad a_1(Y_1, Y_2) = \|Y_1 - Y_2\| - E[\|Y_1 - Y_2\|], \]
\[ a_2(X_1) = E[a_1(X_1, X_2)|X_1], \quad a_3(Y_1) = E[a_1(Y_1, Y_2)|Y_1]. \]

We divide the proof into two steps.

Step 1. Recall the definition of \( h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4)) \) given in (16). It is easy to show that
\[
\frac{1}{4} \sum_{1 \leq i,j \leq 4, i \neq j} \|X_i - X_j\|\|Y_i - Y_j\| = \frac{1}{4} \sum_{1 \leq i,j \leq 4, i \neq j} a_1(X_i, X_j) a_1(Y_i, Y_j) + \frac{1}{4} E[\|X_1 - X_2\|] \sum_{1 \leq i,j \leq 4, i \neq j} a_1(Y_i, Y_j)
\]
\[
+ \frac{1}{4} E[\|Y_1 - Y_2\|] \sum_{1 \leq i,j \leq 4, i \neq j} a_1(X_i, X_j) + 3E[\|X_1 - X_2\|]E[\|Y_1 - Y_2\|],
\]
\[
\frac{1}{4} \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, j \neq i} \|X_i - X_j\| \sum_{1 \leq j \leq 4, j \neq i} \|Y_i - Y_j\| \right)
\]
\[
= \frac{1}{4} \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, j \neq i} a_1(X_i, X_j) \sum_{1 \leq j \leq 4, j \neq i} a_1(Y_i, Y_j) \right) + \frac{3}{4} E[\|X_1 - X_2\|] \sum_{1 \leq i,j \leq 4, i \neq j} a_1(Y_i, Y_j)
\]
\[
+ \frac{3}{4} E[\|Y_1 - Y_2\|] \sum_{1 \leq i,j \leq 4, i \neq j} a_1(X_i, X_j) + 9E[\|X_1 - X_2\|]E[\|Y_1 - Y_2\|],
\]
and
\[
\frac{1}{24} \sum_{1 \leq i,j \leq 4, i \neq j} \|X_i - X_j\| \sum_{1 \leq i,j \leq 4, i \neq j} \|Y_i - Y_j\|
\]
\[
= \frac{1}{24} \sum_{1 \leq i,j \leq 4, i \neq j} a_1(X_i, X_j) \sum_{1 \leq i,j \leq 4, i \neq j} a_1(Y_i, Y_j) + \frac{1}{2} E[\|X_1 - X_2\|] \sum_{1 \leq i,j \leq 4, i \neq j} a_1(Y_i, Y_j)
\]
\[
+ \frac{1}{2} E[\|X_1 - X_2\|] \sum_{1 \leq i,j \leq 4, i \neq j} a_1(Y_i, Y_j) + 6E[\|X_1 - X_2\|]E[\|Y_1 - Y_2\|].
\]

By these equalities and (16), we can obtain
\[
h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))
\]
\[
= \frac{1}{4} \sum_{1 \leq i,j \leq 4, i \neq j} a_1(X_i, X_j) a_1(Y_i, Y_j) - \frac{1}{4} \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, j \neq i} a_1(X_i, X_j) \sum_{1 \leq j \leq 4, j \neq i} a_1(Y_i, Y_j) \right)
\]
Combining the above three equalities yields (A.58). This concludes the proof of Lemma 5.

Step 2. Since \( d(X_1, X_2) = a_1(X_1, X_2) - a_2(X_1) - a_2(X_2) \), it holds that
\[
\begin{align*}
\sum_{1 \leq i, j \leq 4, \ i \neq j} a_1(X_i, X_j) a_1(Y_i, Y_j) &= \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) d(Y_i, Y_j) + 2 \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) a_3(Y_i) + 2 \sum_{1 \leq i, j \leq 4, \ i \neq j} d(Y_i, Y_j) a_2(X_i) + 4 \sum_{i=1}^4 a_2(X_i) a_3(Y_i) + 2 \sum_{1 \leq i, j \leq 4, \ i \neq j} d(Y_i, Y_j) a_2(X_i) + 8 \sum_{i=1}^4 a_2(X_i) \left( \sum_{i=1}^4 a_3(Y_i) \right) + 4 \sum_{i=1}^4 a_2(X_i) a_3(Y_i),
\end{align*}
\]
and
\[
\begin{align*}
\sum_{1 \leq i, j \leq 4, \ i \neq j} a_1(X_i, X_j) a_1(Y_i, Y_j) &= \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) d(Y_i, Y_j) + 6 \sum_{i=1}^4 a_2(X_i) \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(Y_i, Y_j) \right) + 6 \left( \sum_{i=1}^4 a_3(Y_i) \right) \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) \right) + 36 \sum_{i=1}^4 a_2(X_i) \left( \sum_{i=1}^4 a_3(Y_i) \right).
\end{align*}
\]
Combining the above three equalities yields (A.58). This concludes the proof of Lemma 5.


**LEMMA 6.** If \( X \) is independent of \( Y \), it holds that
\[
\mathbb{E}[h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))]^2 = \frac{1}{2} \mathbb{V}^2(X) \mathbb{V}^2(Y).
\]
Proof. From Lemma 5, we can deduce

\[ \mathbb{E}[h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))^2] = \sum_{k=1}^{6} I_k, \]

where

\[ I_1 = \frac{1}{16} \mathbb{E}\left[ \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) d(Y_i, Y_j) \right)^2 \right], \]

\[ I_2 = \frac{1}{16} \mathbb{E}\left[ \left( \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, \ j \neq i} d(X_i, X_j) \sum_{1 \leq j \leq 4, \ j \neq i} d(Y_i, Y_j) \right) \right)^2 \right], \]

\[ I_3 = \frac{1}{576} \mathbb{E}\left[ \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) \sum_{1 \leq i, j \leq 4, \ i \neq j} d(Y_i, Y_j) \right)^2 \right], \]

\[ I_4 = -\frac{1}{8} \mathbb{E}\left\{ \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) d(Y_i, Y_j) \right) \left[ \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, \ j \neq i} d(X_i, X_j) \sum_{1 \leq j \leq 4, \ j \neq i} d(Y_i, Y_j) \right) \right]\right\}, \]

\[ I_5 = \frac{1}{48} \mathbb{E}\left[ \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) d(Y_i, Y_j) \right) \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) \right) \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(Y_i, Y_j) \right) \right], \]

and

\[ I_6 = -\frac{1}{48} \mathbb{E}\left\{ \left[ \sum_{i=1}^{4} \left( \sum_{1 \leq j \leq 4, \ j \neq i} d(X_i, X_j) \sum_{1 \leq j \leq 4, \ j \neq i} d(Y_i, Y_j) \right) \right] \times \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(X_i, X_j) \right) \left( \sum_{1 \leq i, j \leq 4, \ i \neq j} d(Y_i, Y_j) \right) \right\}. \]

Since \( \mathbb{E}[d(X_1, X_2)d(X_1, X_3)] = 0 \) and \( \mathbb{E}[d(X_1, X_2)] = 0 \), under the independence of \( X \) and \( Y \) we have

\[ I_1 = \frac{3}{2} \mathbb{E}[d^2(X_1, X_2)] \mathbb{E}[d^2(Y_1, Y_2)] \]

and

\[ I_2 = \frac{1}{16} \sum_{i=1}^{4} \left[ \mathbb{E}\left( \sum_{1 \leq j \leq 4, \ j \neq i} d(X_i, X_j) \right)^2 \mathbb{E}\left( \sum_{1 \leq j \leq 4, \ j \neq i} d(Y_i, Y_j) \right)^2 \right] \]

\[ + \frac{1}{16} \sum_{1 \leq i, k \leq 4, \ i \neq k} \left\{ \mathbb{E}\left[ \left( \sum_{1 \leq j \leq 4, \ j \neq i} d(X_i, X_j) \right) \left( \sum_{1 \leq l \leq 4, \ l \neq k} d(X_k, X_l) \right) \right] \times \mathbb{E}\left[ \left( \sum_{1 \leq j \leq 4, \ j \neq i} d(Y_i, Y_j) \right) \left( \sum_{1 \leq l \leq 4, \ l \neq k} d(Y_k, Y_l) \right) \right] \right\}. \]
\[
I_3 = \frac{16}{576} \mathbb{E} \left( \sum_{1 \leq i < j \leq 4} d(X_i, X_j) \right)^2 \mathbb{E} \left( \sum_{1 \leq i < j \leq 4} d(Y_i, Y_j) \right)^2
= \mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)],
\]

\[
I_4 = -\frac{1}{4} \sum_{i=1}^{4} \mathbb{E} \left\{ \left( \sum_{1 \leq k < l \leq 4} d(X_k, X_l) d(Y_k, Y_l) \right) \left( \sum_{1 \leq j \leq 4, j \neq i} d(X_i, X_j) \right) \left( \sum_{1 \leq j \leq 4, j \neq i} d(Y_i, Y_j) \right) \right\}
= -\frac{1}{4} \times 4 \times 3 \mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)] = -3\mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)],
\]

\[
I_5 = \frac{8}{48} \sum_{1 \leq i < j \leq 4} \mathbb{E} \left[ d(X_i, X_j) d(Y_i, Y_j) \left( \sum_{1 \leq k < l \leq 4} d(X_k, X_l) \right) \left( \sum_{1 \leq k < l \leq 4} d(Y_k, Y_l) \right) \right]
= \mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)],
\]

and
\[
I_6 = -\frac{4}{48} \sum_{i=1}^{4} \left\{ \mathbb{E} \left[ \left( \sum_{1 \leq j \leq 4, j \neq i} d(X_i, X_j) \right) \left( \sum_{1 \leq k < l \leq 4} d(X_k, X_l) \right) \right] \right\}
\times \mathbb{E} \left[ \left( \sum_{1 \leq j \leq 4, j \neq i} d(Y_i, Y_j) \right) \left( \sum_{1 \leq k < l \leq 4} d(Y_k, Y_l) \right) \right]
= -\frac{4 \times 4 \times 9}{48} \mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)] = -3\mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)].
\]

Consequently, it follows that
\[
\mathbb{E}[h((X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4))]^2 = \sum_{k=1}^{6} I_k = \frac{1}{2} \mathbb{E}[d^2(X_1, X_2)]\mathbb{E}[d^2(Y_1, Y_2)]
= \frac{1}{2} \nu^2(X)\nu^2(Y),
\]

which completes the proof of Lemma 6.

**C.7. Lemma 7 and its proof.** The following lemma provides some basic inequalities that are based on the Taylor expansion and serve as the fundamental ingredients for the proofs of Propositions 1–3.
Lemma 7. For $x \geq -1$, it holds that

\begin{align}
(A.59) \quad |(1 + x)^{1/2} - 1| & \leq |x|, \\
(A.60) \quad |(1 + x)^{1/2} - (1 + x/2)| & \leq x^2/2, \\
(A.61) \quad |(1 + x)^{1/2} - (1 + x/2 - x^2/8)| & \leq 3|x^2|/8, \\
(A.62) \quad |(1 + x)^{1/2} - (1 + x/2 - x^2/8 + x^3/16)| & \leq x^4.
\end{align}

Proof. (i) We first prove (A.59). It is evident that $1 + x \leq (1 + x)^{1/2} \leq 1$ for $x \in [-1, 0]$ and $1 < (1 + x)^{1/2} < 1 + x$ for $x \in (0, \infty)$. Thus we can obtain (A.59) directly.

(ii) We next show (A.60). Define $u_1(x) = (1 + x)^{1/2} - (1 + x/2)$. Then we have the derivative

$$u_1'(x) = [(1 + x)^{-1/2} - 1]/2,$$

$u_1'(x) > 0$ for $x \in [-1, 0)$, and $u_1'(x) < 0$ for $x \in (0, \infty)$. Since $u_1(0) = 0$, it holds that $u_1(x) \leq 0$ for $x \geq -1$. It remains to show that for $x \geq -1$,

$$u_1(x) \geq -x^2/2.$$

Denote by $u_2(x) = (1 + x)^{1/2} - (1 + x/2) + x^2/2$. Then we have

$$u_2'(x) = \frac{1}{2}(1 + x)^{1/2} - \frac{1}{2} + x,$$

$$u_2''(x) = -\frac{1}{4}(1 + x)^{-3/2} + 1,$$

$u_2''(x) \leq 0$ for $-1 \leq x \leq 4^{-2/3} - 1$, and $u_2''(x) > 0$ for $x > 4^{-2/3} - 1$. In addition, it holds that $u_2'(-1) = +\infty$, $u_2'(0) = 0$, and $u_2'(\infty) = +\infty$, which lead to $u_2(x) \geq \min\{u_2(-1), u_2(0)\} = 0$. Hence the proof of (A.60) is completed.

(iii) We now prove (A.61). First, the result is trivial when $x = 0$. Define $u_3(x) = (1 + x)^{1/2} - (1 + x/2 - x^2/8)$. Then we have

$$u_3'(x) = \frac{1}{2}[(1 + x)^{-1/2} - 1 + x/2]$$

$$= \frac{1}{2}(1 + x)^{-1}[(1 + x)^{1/2} - (1 + x/2 - x^2/2)].$$

It has been shown in the proof above that $(1 + x)^{1/2} - (1 + x/2) + x^2/2 \geq 0$ for $x \geq 1$. Thus $u_3'(x) \geq 0$ for $x \geq -1$. It follows that $u_3(x) \leq 0$ for $-1 \leq x \leq 0$ and $u_3(x) > 0$ for $x > 0$. Now it remains to show that for $x \in [-1, 0) \cup (0, \infty)$,

$$u_3(x)/x^3 \leq 3/8.$$

It is easy to show that

$$\left(\frac{u_3(x)}{x^3}\right)' = \frac{u_4(x)}{2x^4},$$

where $u_4(x) = -5(1 + x)^{1/2} - (1 + x)^{-1/2} + 2x - x^2/4 + 6$.

Observe that

$$u_4'(x) = -\frac{5}{2}(1 + x)^{-1/2} + \frac{1}{2}(1 + x)^{-3/2} + 2 - \frac{x}{2},$$

$$u_4''(x) = \frac{5}{4}(1 + x)^{-3/2} - \frac{3}{4}(1 + x)^{-5/2} - \frac{1}{2},$$

$$u_4'''(x) = -\frac{15}{8}x(1 + x)^{-7/2},$$

which lead to

$$u_4'(x) = -\frac{5}{2}(1 + x)^{-1/2} + \frac{1}{2}(1 + x)^{-3/2} + 2 - \frac{x}{2},$$

$$u_4''(x) = \frac{5}{4}(1 + x)^{-3/2} - \frac{3}{4}(1 + x)^{-5/2} - \frac{1}{2},$$

$$u_4'''(x) = -\frac{15}{8}x(1 + x)^{-7/2},$$

which lead to

$$u_4'(x) = -\frac{5}{2}(1 + x)^{-1/2} + \frac{1}{2}(1 + x)^{-3/2} + 2 - \frac{x}{2},$$

$$u_4''(x) = \frac{5}{4}(1 + x)^{-3/2} - \frac{3}{4}(1 + x)^{-5/2} - \frac{1}{2},$$

$$u_4'''(x) = -\frac{15}{8}x(1 + x)^{-7/2},$$

which lead to
$u'''(x) > 0$ for $x \in [-1, 0)$, and $u''''(x) < 0$ for $x > 0$. Furthermore, $u''(0) = 0$ and thus $u''(x) \leq 0$ for any $x \in [-1, 0) \cup (0, \infty)$. In addition, $u'(0) = 0$ and thus $u'(x) > 0$ for $x \in [-1, 0)$ and $u'(x) < 0$ for $x \in (0, \infty)$. Since $u_4(0) = 0$, it follows that $u_4(x) < 0$ for any $x \in [-1, 0) \cup (0, \infty)$, which entails that
\[
\frac{u_3(x)}{x^3} \leq \frac{u_3(x)}{x^3} \bigg|_{x=-1} = \frac{3}{8}.
\]
Similay by taking derivatives, (A.62) can be proved. We omit its proof to avoid redundancy. This concludes the proof of Lemma 7.

C.8. Lemma 8 and its proof.

**Lemma 8.** If $\mathbb{E}[\|X\|^2] < \infty$, then we have
\begin{align*}
(A.63) & \quad \mathbb{E}[W_{12}^2] = B_X^{-2} \left( 2 \mathbb{E}[\|X\|^4] - (\mathbb{E}[\|X\|^2])^2 \right) + 4 \mathbb{E}[X_1^2 X_2^2], \\
(A.64) & \quad \mathbb{E}[W_{12} W_{13}] = B_X^{-2} \left( \mathbb{E}[\|X\|^4] - (\mathbb{E}[\|X\|^2])^2 \right).
\end{align*}

**Proof.** Define $\alpha_1(X) = \|X\|^2 - \mathbb{E}[\|X\|^2]$ and $\alpha_2(X_1, X_2) = X_1^2 X_2$. By the definition of $W_{12}$ and $W_{13}$, we have
\[
\mathbb{E}[W_{12}^2] = B_X^{-2} \mathbb{E}\left\{ [\alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2)]^2 \right\},
\]
\[
\mathbb{E}[W_{12} W_{13}] = B_X^{-2} \mathbb{E}\left\{ [\alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2)] [\alpha_1(X_1) + \alpha_1(X_3) - 2\alpha_2(X_1, X_3)] \right\}.
\]

Since $\mathbb{E}[\alpha_1(X)] = 0$ and $\mathbb{E}[X] = 0$, by expanding the products above we can deduce
\begin{align*}
(A.65) & \quad \mathbb{E}[W_{12}^2] = B_X^{-2} \left( 2 \mathbb{E}[\alpha_1^2(X_1)] + 4 \mathbb{E}[\alpha_2^2(X_1, X_2)] \right) \\
& \quad = B_X^{-2} \left( 2 [\mathbb{E}[\|X\|^4] - (\mathbb{E}[\|X\|^2])^2 + 4 \mathbb{E}[X_1^2 X_2^2] \right)
\end{align*}
and
\begin{align*}
(A.66) & \quad \mathbb{E}[W_{12} W_{13}] = B_X^{-2} \mathbb{E}[\alpha_1^2(X_1)] = B_X^{-2} \left( \mathbb{E}[\|X\|^4] - (\mathbb{E}[\|X\|^2])^2 \right).
\end{align*}
The desired result then follows immediately. This completes the proof of Lemma 8.


**Lemma 9.** If $\mathbb{E}[\|X\|^4] < \infty$, then we have
\begin{align*}
(A.67) & \quad \mathbb{E}[W_{12} W_{13} W_{24} W_{34}] - 4 \mathbb{E}[W_{12} W_{13} W_{24} W_{45}] + 2 \mathbb{E}[W_{12} W_{13}]^2 \\
& \quad = 16 B_X^{-4} \mathbb{E}[X_1^2 \Sigma_x X_2]^2.
\end{align*}

**Proof.** By the definition of $W_{ij}$, we have
\begin{align*}
\mathbb{E}[W_{12} W_{13} W_{24} W_{34}] & \quad = B_X^{-4} \mathbb{E}\left\{ [\alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2)] [\alpha_1(X_1) + \alpha_1(X_3) - 2\alpha_2(X_1, X_3)] \right. \\
& \quad \times [\alpha_1(X_2) + \alpha_1(X_4) - 2\alpha_2(X_2, X_4)] [\alpha_1(X_3) + \alpha_1(X_4) - 2\alpha_2(X_3, X_4)] \left. \right\}.
\end{align*}
Noting that $\mathbb{E}[\alpha_1(X_1)] = 0$ and $\mathbb{E}[X] = 0$, it follows from expanding the above product and the symmetry of $X_1, \ldots, X_4$ that
\begin{align*}
\mathbb{E}[W_{12} W_{13} W_{24} W_{34}] & \quad = B_X^{-4} \left\{ 2 [\mathbb{E}[\alpha_1^2(X)]]^2 + 16 \mathbb{E}[\alpha_2(X_1, X_2) \alpha_2(X_1, X_3) \alpha_2(X_2, X_4)] \right. \\
& \quad \left. + 16 \mathbb{E}[\alpha_2(X_1, X_2) \alpha_2(X_1, X_3) \alpha_2(X_2, X_4) \alpha_2(X_3, X_4)] \right\}.
\end{align*}
By the same token, we can deduce
\[
\mathbb{E}[W_{12}W_{13}W_{24}W_{45}]
= B_{X}^{-4}\mathbb{E}\left\{ \left[ \alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2) \right] \left[ \alpha_1(X_1) + \alpha_1(X_3) - 2\alpha_2(X_1, X_3) \right] \times \left[ \alpha_1(X_2) + \alpha_1(X_4) - 2\alpha_2(X_2, X_4) \right] \left[ \alpha_1(X_4) + \alpha_1(X_5) - 2\alpha_2(X_4, X_5) \right] \right\},
\]
\[
= B_{X}^{-4} \left\{ \left\langle \mathbb{E}[\alpha_1^2(X)] \right\rangle^2 + 4\mathbb{E}[\alpha_2(X_1, X_2)\alpha_2(X_1, X_3)\alpha_2(X_2, X_4)\alpha_2(X_3, X_4)] \right\}.
\]
Therefore, combining the above expressions with (A.66) results in
\[
\mathbb{E}[W_{12}W_{13}W_{24}W_{34}] - 4\mathbb{E}[W_{12}W_{13}W_{24}W_{45}] + 2(\mathbb{E}[W_{12}W_{13}])^2
= 16B_{X}^{-4}\mathbb{E}[\alpha_2(X_1, X_2)\alpha_2(X_1, X_3)\alpha_2(X_2, X_4)\alpha_2(X_3, X_4)] = 16B_{X}^{-4}\mathbb{E}[X^T\Sigma_x X]^2,
\]
which concludes the proof of Lemma 9.

C.10. Lemma 10 and its proof.

**Lemma 10.** For any random vectors \(X \in \mathbb{R}^p\) and \(Y \in \mathbb{R}^q\) satisfying \(\mathbb{E}[\|X\|^2] + \mathbb{E}[\|Y\|^2] < \infty\), we have
\[
\mathcal{V}^2(X, Y) = I_1 + I_2 + I_3 + I_4 + I_5,
\]
where
\[
I_1 = \frac{1}{4}B_{X}^{1/2}B_{Y}^{1/2}(\mathbb{E}[W_{12}V_{12}] - 2\mathbb{E}[W_{12}V_{13}]),
\]
\[
I_2 = -\frac{1}{16}B_{X}^{1/2}B_{Y}^{1/2} \left( \mathbb{E}[W_{12}V_{12}^2] - 2\mathbb{E}[W_{12}V_{13}^2] + \mathbb{E}[W_{12}^2V_{12}] - 2\mathbb{E}[W_{12}^2V_{13}] \right),
\]
\[
I_3 = \frac{1}{32}B_{X}^{1/2}B_{Y}^{1/2} \left( \mathbb{E}[W_{12}V_{13}^3] - 2\mathbb{E}[W_{12}^2V_{13}] + \mathbb{E}[W_{12}V_{12}^3] - 2\mathbb{E}[W_{12}^2V_{12}] \right),
\]
\[
I_4 = \frac{1}{64}B_{X}^{1/2}B_{Y}^{1/2} \left( \mathbb{E}[W_{12}^2V_{12}^2] - 2\mathbb{E}[W_{12}^2V_{13}^2] + \mathbb{E}[W_{12}^3V_{13}] - 2\mathbb{E}[W_{12}^3V_{12}] \right),
\]
\[
I_5 = O \left\{ B_{X}^{1/2}B_{Y}^{1/2} \left[ (\mathbb{E}[W_{12}^5])^{2/5}(\mathbb{E}[V_{12}^5])^{3/5} + (\mathbb{E}[W_{12}^5])^{3/5}(\mathbb{E}[V_{12}^5])^{2/5} + (\mathbb{E}[W_{12}^5])^{4/5}(\mathbb{E}[V_{12}^5])^{1/5} + (\mathbb{E}[W_{12}^5])^{1/5}(\mathbb{E}[V_{12}^5])^{1/5} \right] \right\}.
\]

**Proof.** We will conduct the Taylor expansion to \(\mathcal{V}^2(X, Y) = \mathbb{E}[d(X_1, X_2)d(Y_1, Y_2)]\). In light of (A.16), some straightforward calculations lead to
\[
\mathcal{V}^2(X, Y) = \mathbb{E}[b(X_1, X_2)b(Y_1, Y_2)] - 2\mathbb{E}[b(X_1, X_2)b(Y_1, Y_3)] + \mathbb{E}[b(X_1, X_2)]\mathbb{E}[b(Y_1, Y_2)],
\]
where \(b(X_1, X_2) = \|X_1 - X_2\| - B_{X}^{1/2}\) and \(b(Y_1, Y_2) = \|Y_1 - Y_2\| - B_{Y}^{1/2}\). Define
\[
W_{ij} = B_{X}^{-1}(\|X_i - X_j\|^2 - B_{X}) \text{ and } V_{ij} = B_{Y}^{-1}(\|Y_i - Y_j\|^2 - B_{Y}).
\]
Observe that \(b(X_1, X_2) = B_{X}^{1/2}(1 + W_{12})^{1/2} - 1\). An application of similar arguments as those in the proof of Proposition 2 by resorting to (A.62) in Lemma 7 yields
\[
\mathbb{E}[b(X_1, X_2)b(Y_1, Y_2)] = B_{X}^{1/2}B_{Y}^{1/2} \left\{ \frac{1}{4}\mathbb{E}[W_{12}V_{12}] - \frac{1}{16}\mathbb{E}[W_{12}V_{12}^2] + \mathbb{E}[W_{12}^2V_{13}] \right\}
+ \frac{1}{64}\mathbb{E}[W_{12}^2V_{12}^2] + \frac{1}{32}\left( \mathbb{E}[W_{12}^3V_{12}] + \mathbb{E}[W_{12}^3V_{13}] \right) + O \left( \left( \mathbb{E}[W_{12}^3] + \mathbb{E}[W_{12}^3V_{12}] \right) \right).
By the same token, we can deduce that

\[
\mathbb{E}[b(X_1, X_2)b(Y_1, Y_3)] = B_X^{1/2}B_Y^{1/2}\left\{ \frac{1}{4}\mathbb{E}[W_{12}V_{13}] - \frac{1}{16}\left( \mathbb{E}[W_{12}V_{13}^2] + \mathbb{E}[W_{12}^2V_{13}] \right) \right. \\
+ \frac{1}{64}\mathbb{E}[W_{12}^2V_{13}^2] + \frac{1}{32}\left( \mathbb{E}[W_{12}V_{13}^3] + \mathbb{E}[W_{12}^3V_{13}] \right) + O\left( \mathbb{E}[W_{12}^4V_{13}] + \mathbb{E}[W_{12}^3V_{13}^2] \right) + \mathbb{E}[W_{12}^2V_{13}^3] + \mathbb{E}[W_{12}V_{13}^4] \right\},
\]

and

\[
\mathbb{E}[b(X_1, X_2)|\mathbb{E}[b(Y_1, Y_2)] = B_X^{1/2}B_Y^{1/2}\left\{ \frac{1}{64}\mathbb{E}[W_{12}^2\mathbb{E}[V_{13}^2] + O\left( \mathbb{E}[W_{12}^4\mathbb{E}[V_{13}^3] \right) + \mathbb{E}[W_{12}^3\mathbb{E}[V_{13}^2] \right) \right\}.
\]

Therefore, the desired decomposition follows from a combination of the above three representations and the Cauchy–Schwarz inequality, which completes the proof of Lemma 10.

APPENDIX D: THEORETICAL RESULTS FOR THE CASE OF 1/2 < \tau \leq 1

D.1. Theory. In this section, we introduce our parallel results of Theorems 2 and 4 for the case of 1/2 < \tau \leq 1. When \mathbb{E}[\|X\|^{2+2\tau}] + \mathbb{E}[\|Y\|^{2+2\tau}] < \infty \text{ for a larger value of } \tau \text{ with } 1/2 < \tau \leq 1, \text{ the key ingredient is that higher-order Taylor expansions can be applied while bounding } \mathbb{E}[g(X_1, X_2, X_3, X_4)]. \text{ We start with presenting the expansion of } \mathbb{E}[g(X_1, X_2, X_3, X_4)] \text{ for } 1/2 < \tau \leq 1. \text{ Let us define}

\[
\mathcal{G}_1(X) = |\mathbb{E}[(X_1^T X_2)^2 X_1^T \Sigma_x^2 X_2^2]|, \quad \mathcal{G}_2(X) = \mathbb{E}[\|X_1\|^2(X_1^T \Sigma_x X_2)^2],
\]

\[
\mathcal{G}_3(X) = \mathbb{E}[X^T X X^T] \Sigma_x^2 \mathbb{E}[X X^T X],
\]

\[
N_\tau(X) = \frac{\mathbb{E}[(X_1^T \Sigma_x X_2)^2] + B_X^{-2}\tau L_{x,\tau}^{(2+\tau)/(1+\tau)} + \sum_{i=1}^{3} \mathcal{G}_i(X)}{(\mathbb{E}[(X_1^T X_2)^2])^2}.
\]

We also have \mathcal{G}_1(Y), \mathcal{G}_2(Y), \mathcal{G}_3(Y), \text{ and } N_\tau(Y) \text{ that are defined in a similar way.}

PROPOSITION 4. \textit{If } \mathbb{E}[\|X\|^{4+4\tau}] < \infty \text{ for some } 1/2 < \tau \leq 1, \text{ then there exists some absolute positive constant } C \text{ such that}

\[
|\mathbb{E}[g(X_1, X_2, X_3, X_4)]| \leq C \left\{ B_X^{-2}\mathbb{E}[(X_1^T \Sigma_x X_2)^2] + B_X^{-3}\sum_{i=1}^{3} \mathcal{G}_i(X) + B_X^{-2}\tau L_{x,\tau}^{(2+\tau)/(1+\tau)} \right\}.
\]

The proof of Proposition 4 is given in Section D.3. We can obtain the following central limit theorem and the associated rate of convergence for the case of 1/2 < \tau \leq 1 by substituting the bounds in Propositions 1–2 and 4 into Theorem 3.

THEOREM 1. \textit{Assume that } \mathbb{E}[\|X\|^{4+4\tau}] + \mathbb{E}[\|Y\|^{4+4\tau}] < \infty \text{ for some } 1/2 < \tau \leq 1 \text{ and}

\[
B_X^{-1} L_{x,1/2}/\mathbb{E}[(X_1^T X_2)^2] \leq \frac{1}{18},
\]

(A.69)

\[
B_Y^{-1} L_{y,1/2}/\mathbb{E}[(X_1^T X_2)^2] \leq \frac{1}{18}.
\]

(A.70)
Then under the independence of $X$ and $Y$, we have

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(T_n \leq x) - \Phi(x)|$$

(A.71) $\leq C \left\{ [N_\tau(X)N_\tau(Y)]^{(1+\tau)/2} + \frac{n^{-\tau}L_{x,\tau}L_{y,\tau}}{(\mathbb{E}[(X^T X_2)^2]\mathbb{E}[(Y^T Y_2)^2])^{1+\tau}} \right\}^{1/(3+2\tau)}$.

The proof of Theorem 1 is provided in Section D.2. Theorem 2 below is a direct corollary of Theorem 1.

**Theorem 2.** Assume that $\mathbb{E}[\|X\|^{4+4\tau}] + \mathbb{E}[\|Y\|^{4+4\tau}] < \infty$ for some $1/2 < \tau \leq 1$ and (20) holds as $n \to \infty$ and $p + q \to \infty$. In addition, assume that (A.69) and $N_\tau(X) \to 0$ are satisfied as $p \to \infty$, and that (A.70) and $N_\tau(Y) \to 0$ are satisfied as $q \to \infty$. Then under the independence of $X$ and $Y$, we have

$$T_n \overset{D}{\to} N(0,1).$$

**D.2. Proof of Theorem 1.** Note that (A.69), (A.70), and Proposition 2 entail that

(A.72) $V^2(X) \geq B_X^{-1}\mathbb{E}[(X^T X_2)^2]/2$ and $V^2(Y) \geq B_Y^{-1}\mathbb{E}[(Y^T Y_2)^2]/2$,

which together with (A.68) leads to

$$\frac{|\mathbb{E}[g(X_1, X_2, X_3, X_4)]|}{|\mathbb{E}[g(Y_1, Y_2, Y_3, Y_4)]|} \leq N_\tau(X)N_\tau(Y).$$

It follows from Proposition 1 and (A.72) that

$$\frac{\mathbb{E}[|d(X_1, X_2)|^{2+2\tau}]\mathbb{E}[|d(Y_1, Y_2)|^{2+2\tau}]}{n^\tau|\mathbb{E}[V^2(X)V^2(Y)]} \leq \frac{n^{-\tau}L_{x,\tau}L_{y,\tau}}{(\mathbb{E}[(X^T X_2)^2]\mathbb{E}[(Y^T Y_2)^2])^{1+\tau}}.$$ 

Therefore, we can obtain the desired result (A.71) by Theorem 3, which concludes the proof of Theorem 1.

**D.3. Proof of Proposition 4.** It suffices to analyze the terms on the right hand side of (A.24). Compared to Proposition 3, we assume higher moments and thus we can conduct higher-order Taylor expansions for term $(1 + W_{12})^{1/2}$.

Let us first deal with term $G_1$. Denote by $D_1 = \{\max(W_{12}, W_{13}, W_{24}, W_{34}) \leq 1\}$ and $D_I^c$ the complement of $D_1$. Following the notation in the proof of Proposition 3, by (A.59) and (A.61) we can deduce

$$G_1 = B_X^2\mathbb{E}\left[\left[\frac{1}{2}W_{12} - \frac{1}{8}W_{12}^2 + O(1)(|W_{12}|^3)\right]\left[\frac{1}{2}W_{13} - \frac{1}{8}W_{13}^2 + O(1)(|W_{13}|^3)\right]\right.\left.\times \left[\frac{1}{2}W_{24} - \frac{1}{8}W_{24}^2 + O(1)(|W_{24}|^3)\right]\left[\frac{1}{2}W_{34} - \frac{1}{8}W_{34}^2 + O(1)(|W_{34}|^3)\right]1\{D_1\}\right)$$

$$+ O(1)B_X^2\mathbb{E}[W_{12}W_{13}W_{24}W_{34}1\{D_I^c\}].$$

By expanding the products and reorganizing the terms, it holds that

$$G_1 = \frac{B_X^2}{16}\left(\mathbb{E}[W_{12}W_{13}W_{24}W_{34}] - \mathbb{E}[W_{12}W_{13}W_{24}W_{34}] + O(1)\mathbb{E}[W_{12}^2W_{13}^2|W_{24}W_{34}1\{D_1\}] + O(1)\mathbb{E}[W_{12}^3|W_{13}W_{24}W_{34}1\{D_1\}] + O(1)\mathbb{E}[W_{12}^2W_{34}1\{D_1\}] + O(1)\mathbb{E}[W_{12}W_{13}W_{24}W_{34}1\{D_I^c\}].\right)$$
Furthermore, if $\mathbb{E}[\|X\|^{4+4\tau}] < \infty$ for some $1/2 < \tau \leq 1$, then an application of Chebyshev’s inequality and the Cauchy–Schwarz inequality results in
\[
\begin{align*}
|\mathbb{E}[W_{12}^2 W_{13}^2 | W_{24} W_{34} | 1(D_1)]| & \leq \mathbb{E}[|W_{12}^{1+\tau} | W_{13}^{1+\tau} | W_{24} W_{34} |] \\
& = \mathbb{E}\left\{ \mathbb{E}[|W_{12}|^{1+\tau} | W_{13}|^{1+\tau} | X_2, X_3] \mathbb{E}[|W_{24} W_{34}| | X_2, X_3]\right\} \\
& = \mathbb{E}\left\{ \left( \mathbb{E}[|W_{12}|^{2+2\tau} | X_2]\right)^{\frac{2+\tau}{2+\tau}} \right\} \mathbb{E}\left\{ \left( \mathbb{E}[|W_{13}|^{2+2\tau} | X_3]\right)^{\frac{2+\tau}{2+\tau}} \right\} \\
& \leq \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}} \mathbb{E}[|W_{13}|^{2+2\tau}]^{\frac{2+\tau}{2+\tau}} = \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}}.
\end{align*}
\]
By the same token, we can obtain
\[
\begin{align*}
\mathbb{E}[W_{12}^2 W_{13}^2 | W_{13} W_{24} | 1(D_1)] & \leq \mathbb{E}[|W_{12}|^{1+\tau} | W_{13}|^{1+\tau} | W_{13} W_{24}] \\
& \leq \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}}, \\
\mathbb{E}[|W_{12}|^{3} | W_{13} W_{24} | 1(D_1)] & \leq \mathbb{E}[|W_{12}|^{1+2\tau} | W_{13} W_{24} W_{34}] \\
& \leq \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}},
\end{align*}
\]
and
\[
\begin{align*}
\mathbb{E}[|W_{12} W_{13} W_{24} W_{34}| 1(D_1)] & \leq 4 \mathbb{E}[|W_{12}|^{1+2\tau} | W_{13} W_{24} W_{34}] \\
& \leq 4 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}}.
\end{align*}
\]
In consequence, it follows that
\begin{equation}
(A.73)
G_1 = \frac{B_X^2}{16} \left( \mathbb{E}[W_{12} W_{13} W_{24} W_{34}] - \mathbb{E}[W_{12}^2 W_{13} W_{24} W_{34}] + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2+\tau}{2+\tau}}) \right).
\end{equation}

As for term $G_2$, let $D_3 = \{\max(W_{12}, W_{13}, W_{24}, W_{45}) \leq 1\}$ and $D_3^c$ be its complement. Similarly, by (A.59) and (A.61) we can obtain
\begin{equation}
(A.74)
G_2 = \frac{B_X^2}{64} \left( 4 \mathbb{E}[W_{12} W_{13} W_{24} W_{45}] - \mathbb{E}[W_{12}^2 W_{13} W_{24} W_{45}] - \mathbb{E}[W_{12} W_{13} W_{24} W_{45}^2] - \mathbb{E}[W_{12} W_{13} W_{24} W_{45}] + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2+\tau}{2+\tau}}) \right).
\end{equation}

We now consider term $G_3$. Define $D_3 = \{\max(W_{12}, W_{13}) \leq 1\}$ and $D_3^c$ its complement. Similarly, we can show that
\[
G_3 = B_X \left( \frac{1}{4} \mathbb{E}[W_{12} W_{13}] - \frac{1}{8} \mathbb{E}[W_{12}^2 W_{13} 1\{D_3\}] + O(1)\delta_2 \right),
\]
where $\delta_2 = \mathbb{E}[W_{12}^2 W_{13} 1\{D_3\}] + \mathbb{E}[W_{12} W_{13}^2 1\{D_3\}] + \mathbb{E}[W_{12} W_{13}^2 1\{D_3^c\}]$. Note that when $\mathbb{E}[\|X\|^{4+4\tau}] < \infty$ for some $1/2 < \tau \leq 1$, it follows from Chebyshev’s inequality that
\[
\begin{align*}
\delta_2 \cdot \mathbb{E}[W_{12} W_{13}] & \leq \mathbb{E}[|W_{12}|^{1+\tau} | W_{13}|^{1+\tau} | W_{12} W_{13}] + 3 \mathbb{E}[|W_{12}| | W_{13}|^{1+2\tau} | W_{12} W_{13}] \\
& \leq 4 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}}, \\
\delta_2 \mathbb{E}[W_{12}^2 W_{13}] & \leq 4 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}}, \\
\delta_2 \mathbb{E}[W_{12}^2 W_{13}] & \leq 4 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2+\tau}},
\end{align*}
\]
\[ \delta_2^2 \leq 3(\mathbb{E}[W_{12}^2W_{13}^21(D_3)])^2 + 3(\mathbb{E}[W_{12}W_{13}^31(D_3)])^2 + 3(\mathbb{E}[W_{12}W_{13}1(D_3)])^2 \]
\[ \leq 18(\mathbb{E}[W_{12}|W_{13}|^{1+\gamma})^2 \leq 18(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}), \]
\[ \mathbb{E}[W_{12}W_{13}]|\mathbb{E}[W_{12}W_{13}1(D_3)]| \leq \mathbb{E}[W_{12}W_{13}] (\mathbb{E}[W_{12}^2|W_{13}|^{2\tau}] + \mathbb{E}[W_{12}|W_{13}|^{1+2\tau}]) \]
\[ \leq 2(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}), \]
and
\[ (\mathbb{E}[W_{12}^2W_{13}1(D_3)])^2 \leq (\mathbb{E}[W_{12}^2|W_{13}|^{\gamma}])^2 \leq (\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}). \]

Thus we can deduce
\[ G_3^2 = \frac{B_3^2}{16} (\mathbb{E}[|W_{12}|^{1/2}] - \mathbb{E}[W_{12}W_{13}]|\mathbb{E}[W_{12}W_{13}] + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}). \]

Then we deal with term \( \Delta G_4 \). Denote by \( D_4 = \{\max(W_{12}, W_{13}, W_{24}) \leq 1\} \) and \( D_5 \) its complement. By (A.59) and (A.60), we have for \( 1/2 < \tau \leq 1, \)
\[ G_4 = B_3^{3/2} \mathbb{E}\{(1 + W_{12})^{1/2}(1 + W_{13})^{1/2} - 1\} \]
\[ = B_3^{3/2} \mathbb{E}\{[W_{12}/2 + O(W_{12})][W_{13}/2 + O(W_{13})][W_{24}/2 + O(W_{24})]1(D_4)\} \]
\[ + O(1)\mathbb{E}[W_{12}W_{13}W_{24}]1(D_5) \]
\[ = B_3^{3/2} \left( \frac{1}{8} \mathbb{E}[W_{12}W_{13}W_{24}] + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}W_{13}W_{24}]) \right). \]

Moreover, it holds that
\[ \Delta = B_3^{1/2}\mathbb{E}(1 + W_{12})^{1/2} - 1 \]
\[ = B_3^{1/2}\mathbb{E}\{(1 + W_{12}^2/2 + O(1)(W_{12})^2)1\{W_{12} \leq 1\}\} + O(1)B_3^{1/2}\mathbb{E}[W_{12}]1\{W_{12} > 1\} \]
\[ = B_3^{1/2}\left( \frac{1}{8} \mathbb{E}[W_{12}^2] + O(1)(\mathbb{E}[|W_{12}|^{2}1\{W_{12} \leq 1\}] + \mathbb{E}[|W_{12}|^{2}1\{W_{12} > 1\}) \right). \]

Observe from (A.29) that for \( 1/2 < \tau \leq 1, \) we have
\[ \mathbb{E}[W_{12}W_{13}W_{24}]1\{W_{12} \leq 1\} + \mathbb{E}[W_{12}^21\{W_{12} > 1\}) \leq C(\mathbb{E}[W_{12}^2])^{3/2}\mathbb{E}[|W_{12}|^{1+2\tau}] \]
\[ \leq C(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}. \]

and
\[ (\mathbb{E}[W_{12}^2W_{13}]W_{24}) + \mathbb{E}[W_{13}^2W_{12}W_{24}])(\mathbb{E}[W_{12}^31\{W_{12} \leq 1\}] + \mathbb{E}[W_{12}^31\{W_{12} > 1\}) \]
\[ \leq (\mathbb{E}[W_{12}^2W_{13}]W_{24}) + \mathbb{E}[W_{13}^2W_{12}W_{24}]\mathbb{E}[W_{12}^2] \]
\[ \leq 2(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}. \]

Hence it follows that
\[ \Delta G_4 = \frac{B_3^2}{64} \left( \mathbb{E}[W_{12}^2]\mathbb{E}[W_{12}W_{13}W_{24}] + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}. \right) \]

As for term \( \Delta^2 G_3 \), by (A.20) and (A.59) we have for \( 1/2 < \tau \leq 1, \)
\[ |\Delta^2 G_3| \leq C B_3^2 \mathbb{E}[W_{12}^2]\mathbb{E}[W_{12}W_{13}] \]
\[ \leq C B_3^2 (\mathbb{E}[|W_{12}|^{2+2\tau}]^{\frac{2}{3\gamma+\pi\gamma}}. \)
Note that (A.19) entails that
\[ |\Delta^4| \leq CB_X^2 \left( \mathbb{E}[|W_{12}|^3 1\{W_{12} \leq 1\}] + \mathbb{E}[W_{12}^2 1\{W_{12} > 1\}] \right)^2 \]
(A.78)
\[ \leq CB_X^2 \left( \mathbb{E}[|W_{12}|^{2+\tau}] \right)^2 \leq CB_X^2 \left( \mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2}} . \]

Finally, substituting (A.73)–(A.78) into (A.24) yields
\[ \mathbb{E}[g(X_1, X_2, X_3, X_4)] = \frac{B_X^2}{16} \left( E_1 + E_2 + O(1)(\mathbb{E}[|W_{12}|^{2+2\tau}] \right)^{\frac{2+\tau}{2}} , \]
where
\[ E_1 = \mathbb{E}[W_{12}W_{13}W_{24}W_{34}] - 4\mathbb{E}[W_{12}W_{13}W_{24}W_{45}] + 2(\mathbb{E}[W_{12}W_{13}])^2 \]
and
\[ E_2 = -\mathbb{E}[W_{12}^2W_{13}W_{24}W_{34}] + \mathbb{E}[W_{12}^2W_{13}W_{24}W_{45}] + \mathbb{E}[W_{12}W_{13}W_{24}W_{45}]
+ \mathbb{E}[W_{12}W_{13}W_{24}W_{45}] + \mathbb{E}[W_{12}W_{13}W_{45}] - 2\mathbb{E}[W_{12}W_{13}]\mathbb{E}[W_{12}W_{13}] \]
(A.80)
\[ - \mathbb{E}[W_{12}^2]\mathbb{E}[W_{12}W_{13}W_{24}] . \]

By some algebra, we can obtain Lemma 11 in Section D.4. Recall that \( B_X = 2\mathbb{E}[\|X\|^2] \). Then the equalities obtained above along with Lemma 9 lead to
\[ E_1 + E_2 = 16B_X^{-5} \left( 6\mathbb{E}[\|X\|^2]\mathbb{E}[(X_1^T \Sigma_x X_2)^2] + \mathbb{E}[X_1^T XX^T] \Sigma_x^2 \mathbb{E}[XX^T X]
+ 2\mathbb{E}[(X_1^T X_2)^2 X_1^T \Sigma_x X_2] - 4\mathbb{E}[\|X_1\|^2 (X_1^T \Sigma_x X_2)^2] \right) . \]

Therefore, we can obtain the desired result (A.68), which completes the proof of Proposition 4.


**Lemma 11.** It holds that
\[ E_2 = 16B_X^{-5} \left( \mathbb{E}[X_1^T XX^T] \Sigma_x^2 \mathbb{E}[XX^T X] + 2\mathbb{E}[(X_1^T X_2)^2 X_1^T \Sigma_x X_2]
- 4\mathbb{E}[\|X_1\|^2 (X_1^T \Sigma_x X_2)^2] + 4\mathbb{E}[\|X\|^2] \mathbb{E}[(X_1^T \Sigma_x X_2)^2] \right) . \]
(A.81)

**Proof.** In view of the notation in the proofs of Lemmas 8 and 9, it holds that \( \alpha_1(X) = \|X\|^2 - \mathbb{E}[\|X\|^2] \) and \( \alpha_2(X_1, X_2) = X_1^T X_2 \). Thus we have
\[ W_{12} = B_X^{-1}[\alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2)] . \]

Then it follows that
\[ \mathbb{E}[W_{12}W_{13}W_{24}W_{34}]
= B_X^{-5} \mathbb{E} \left\{ \left[ \alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_2) \right]^2 \left[ \alpha_1(X_1) + \alpha_1(X_2) - 2\alpha_2(X_1, X_3) \right]
\times \left[ \alpha_1(X_2) + \alpha_1(X_4) - 2\alpha_2(X_2, X_4) \right] \left[ \alpha_1(X_3) + \alpha_1(X_4) - 2\alpha_2(X_3, X_4) \right] \right\} . \]

The idea of the proof is to expand the products. Since \( X_1, X_2, X_3, \) and \( X_4 \) are i.i.d., we can deduce
\[ \mathbb{E}[W_{12}^2W_{13}W_{24}W_{34}]
= B_X^{-5}(2D_1 + 8D_2 - 20D_3 - 16D_4 - 8D_5 + 24D_6 + 32D_7 + 16D_8 - 48D_9 - 32D_{10} + 64D_{11}) , \]
It is easy to see that
\[
D_1 = \mathbb{E}[\alpha_1^2(X)\mathbb{E}[\alpha_1^2(X)],
\]
\[
D_2 = \mathbb{E}[\alpha_1^2(X)\mathbb{E}[\alpha_2^2(X_1, X_2)\alpha_1(X_2)],
\]
\[
D_3 = \mathbb{E}[\alpha_1^2(X)\mathbb{E}[\alpha_1(X_1)\alpha_1(X_2)\alpha_2(X_1, X_2)],
\]
\[
D_4 = \mathbb{E}[\alpha_1(X_3)\alpha_2(X_1, X_3)\alpha_2^2(X_1, X_2)\alpha_1(X_2)],
\]
\[
D_5 = \mathbb{E}[\alpha_2^2(X_1, X_2)\mathbb{E}[\alpha_1(X_1)\alpha_1(X_2)\alpha_2(X_1, X_2)],
\]
\[
D_6 = \mathbb{E}[\alpha_1(X_3)\alpha_2(X_1, X_3)\alpha_2(X_1, X_2)\alpha_1^2(X_2)],
\]
\[
D_7 = \mathbb{E}[\alpha_2(X_3, X_4)\alpha_2(X_1, X_3)\alpha_1(X_4)\alpha_2^2(X_1, X_2)],
\]
\[
D_8 = \mathbb{E}[\alpha_1(X_1)\alpha_1(X_2)\alpha_1(X_3)\alpha_2(X_1, X_3)\alpha_2(X_1, X_2)],
\]
\[
D_9 = \mathbb{E}[\alpha_2(X_3, X_4)\alpha_1(X_4)\alpha_1(X_3)\alpha_2(X_1, X_2)\alpha_1(X_2)],
\]
\[
D_{10} = \mathbb{E}[\alpha_2(X_3, X_4)\alpha_2(X_1, X_4)\alpha_2(X_1, X_3)\alpha_2^2(X_1, X_2)],
\]
\[
D_{11} = \mathbb{E}[\alpha_2(X_3, X_4)\alpha_2(X_1, X_3)\alpha_2(X_2, X_4)\alpha_1(X_1)\alpha_2(X_1, X_2)].
\]

Similarly, we can show that
\[
\mathbb{E}[W_{12}^2W_{13}W_{24}W_{45}] = B_X^{-5}(D_1 + 4D_2 - 8D_3 - 8D_4 + 8D_6 + 8D_8),
\]
\[
\mathbb{E}[W_{12}W_{13}W_{24}W_{45}] = B_X^{-5}(D_1 + 4D_2 - 8D_3 - 8D_5 + 4D_6 + 16D_7 - 16D_9),
\]
\[
\mathbb{E}[W_{12}W_{13}W_{24}^2W_{45}] = B_X^{-5}(D_1 + 4D_2 - 8D_3 - 8D_4 + 8D_6 + 8D_8),
\]
\[
\mathbb{E}[W_{12}W_{13}W_{24}W_{45}^2] = B_X^{-5}(D_1 + 4D_2 - 8D_3 - 8D_5 + 4D_6 + 16D_7 - 16D_9),
\]
\[
\mathbb{E}[W_{12}W_{13}]\mathbb{E}[W_{12}^2W_{13}] = B_X^{-5}(D_1 + 4D_2 - 4D_3),
\]
\[
\mathbb{E}[W_{12}^2]\mathbb{E}[W_{12}W_{13}W_{24}] = -B_X^{-5}(4D_3 + 8D_5).
\]
Thus by plugging the above equalities into (A.80), it holds that
\[
E_2 = 16B_X^{-5}(D_9 + 2D_{10} - 4D_{11}).
\]

It is easy to see that
\[
D_9 = \mathbb{E}[X^TX \Sigma^2 \mathbb{E}[XX^TX]],
\]
\[
D_{10} = \mathbb{E}[(X_1^TX_2)^2X_1^T \Sigma^2 X_2],
\]
\[
D_{11} = \mathbb{E}[\|X_1\|^2(X_1^T \Sigma X_2)^2] - \mathbb{E}[\|X\|^2]\mathbb{E}[(X_1^T \Sigma X_2)^2].
\]
Therefore, we can obtain the desired result (A.81), which concludes the proof of Lemma 11.

APPENDIX E: CONNECTIONS BETWEEN NORMAL APPROXIMATION AND GAMMA APPROXIMATION

For the test of independence based on the sample distance covariance, empirically one can use the gamma approximation to calculate the limiting p-values. Huo and Székely (2016) showed that under some moment conditions and the independence of \(X\) and \(Y\), it holds that
\[
n\mathcal{W}_n^\ast(X, Y) \xrightarrow{d} \sum_{i=1}^{\infty} \lambda_i(Z_i^2 - 1),
\]
where \( \{\lambda_i\}_{i\geq 1} \) are some values depending on the underlying distribution and \( \{Z_i\}_{i\geq 1} \) are i.i.d. standard normal random variables. In practice, it is infeasible to apply this limiting distribution directly and thus the gamma approximation can serve as a surrogate. By Huang and Huo (2017), it follows that
\[
\sum_{i=1}^{\infty} \lambda_i = \mathbb{E}[\|X - X'\|]\mathbb{E}[\|Y - Y'\|] \quad \text{and} \quad \sum_{i=1}^{\infty} \lambda_i^2 = \mathbb{V}^2(X)\mathbb{V}^2(Y),
\]
and hence \( \sum_{i=1}^{\infty} \lambda_i (Z_i^2 - 1) \) can be approximated by a centered gamma distribution \( \Gamma(\beta_1, \beta_2) - \beta_1 \beta_2^{-1} \), where the shape and rate parameters \( \beta_1 \) and \( \beta_2 \) are determined by matching the first two moments. To this end, we define
\[
\beta_1 = \left( \frac{\sum_{i=1}^{\infty} \lambda_i}{2 \sum_{i=1}^{\infty} \lambda_i^2} \right)^2 = \frac{\left( \mathbb{E}[\|X - X'\|]\mathbb{E}[\|Y - Y'\|]\right)^2}{2\mathbb{V}^2(X)\mathbb{V}^2(Y)}
\]
and
\[
\beta_2 = \frac{\sum_{i=1}^{\infty} \lambda_i^2}{2 \sum_{i=1}^{\infty} \lambda_i^3} = \frac{\mathbb{E}[\|X - X'\|]\mathbb{E}[\|Y - Y'\|]}{2\mathbb{V}^2(X)\mathbb{V}^2(Y)}.
\]
For a simple illustration, let us consider a specific case when both \( X \) and \( Y \) consist of i.i.d. components. Then it holds that \( \mathbb{E}[\|X - X'\|] = O(\sqrt{pq}) \) and \( \mathbb{E}[\|Y - Y'\|] = O(\sqrt{pq}) \). Moreover, it follows from Proposition 2 that \( \mathbb{V}^2(X) \) and \( \mathbb{V}^2(Y) \) are bounded from above and below by some positive constants, which entails that \( \beta_1 = O(pq) \) and \( \beta_1 \to \infty \) as \( \max\{p, q\} \to \infty \). Recall the fact that the gamma random variable can be represented as a sum of certain i.i.d. exponential random variables. Thus by the central limit theorem, we have
\[
\frac{\Gamma(\beta_1, \beta_2) - \beta_1 \beta_2^{-1}}{\sqrt{\beta_1 \beta_2^{-2}}} \xrightarrow{D} N(0, 1)
\]
as \( \max\{p, q\} \to \infty \). Since \( \beta_1 \beta_2^{-2} = 2\mathbb{V}^2(X)\mathbb{V}^2(Y) \) and Lemma 1 has provided the consistency of \( \mathbb{V}_n^*(X) \) and \( \mathbb{V}_n^*(Y) \), it holds that
\[
\mathbb{P}(T_n \leq x) = \mathbb{P} \left( \sqrt{\frac{n(n-1)}{2}} \frac{\mathbb{V}_n^*(X,Y)}{\sqrt{\mathbb{V}^2(X)\mathbb{V}^2(Y)}} \leq x \right)
\approx \mathbb{P} \left( \frac{n\mathbb{V}_n^*(X,Y)}{\sqrt{2\mathbb{V}^2(X)\mathbb{V}^2(Y)}} \leq x \right) \approx \mathbb{P} \left( \frac{\Gamma(\beta_1, \beta_2) - \beta_1 \beta_2^{-1}}{\sqrt{\beta_1 \beta_2^{-2}}} \leq x \right) \to \Phi(x),
\]
where \( \Phi(x) \) stands for the standard normal distribution function. Therefore, the gamma approximation for \( n\mathbb{V}_n^*(X,Y) \) may be asymptotically equivalent to the normal approximation to \( T_n \) under certain scenarios. It is worth mentioning that the above analysis intends to build some connections between the normal approximation and the gamma approximation, but is not a rigorous proof. A rigorous theoretical foundation for the gamma approximation still remains undeveloped.

**APPENDIX F: ASYMPTOTIC NORMALITY OF \( T_R \)**

An anonymous referee asked a great question on whether similar asymptotic normality as in Theorem 1 and associated rates of convergence as in Theorem 3 hold for the studentized sample distance correlation \( T_R \). The answer is affirmative as shown in the following proposition.
**Proposition 5.** Under the same conditions of Theorem 1, we have \( T_R \overset{D}{\rightarrow} N(0, 1) \). Moreover, under the conditions of Theorem 3, the same rate of convergence as in (21) holds for \( T_R \).

**Proof.** By Lemma 1, we have \( \mathcal{V}_n^*(X) / \mathcal{V}_n^2(X) \overset{p}{\rightarrow} 1 \) and \( \mathcal{V}_n^*(Y) / \mathcal{V}_n^2(Y) \overset{p}{\rightarrow} 1 \) under condition (18). In addition, it follows from (A.50) and Lemma 5 that for \( 0 < \tau \leq 1 \),

\[
E[|\mathcal{V}_n^*(X, Y) - \mathcal{V}^2(X, Y)|^{1+\tau}] \leq Cn^{-\tau} [E(|d(X_1, X_2)|^{2+2\tau})E(|d(Y_1, Y_2)|^{2+2\tau})]^{1/2}.
\]

Hence under condition (18), it holds that

\[
E\left[\left| \frac{\mathcal{V}_n^*(X, Y)}{\sqrt{\mathcal{V}_n^2(X, Y)}} - \mathcal{R}^2(X, Y) \right|^{1+\tau}\right] 
\leq \frac{C}{n^{\tau/2}} \left( \frac{E(|d(X_1, X_2)|^{2+2\tau})E(|d(Y_1, Y_2)|^{2+2\tau})}{n^{\tau} [\mathcal{V}^2(X, Y)]^{1+\tau}} \right)^{1/2} \rightarrow 0.
\]

This entails that \( \frac{\mathcal{V}_n^*(X, Y)}{\sqrt{\mathcal{V}_n^2(X, Y)}} \overset{p}{\rightarrow} \mathcal{R}^2(X, Y) \) and thus \( \mathcal{R}_n^*(X, Y) \overset{p}{\rightarrow} \mathcal{R}^2(X, Y) \) as well. Under the null hypothesis, it holds that \( \mathcal{R}^2(X, Y) = 0 \) and hence \( \mathcal{R}_n^*(X, Y) \overset{p}{\rightarrow} 0 \). In light of the definition of \( T_n^* \), it holds that

\[
T_R = T_n \cdot \sqrt{\frac{n(n-3) - 2}{n(n-1)}} \cdot \frac{1}{\sqrt{1 - \mathcal{R}_n^*(X, Y)}}.
\]

By Theorem 1, we have \( T_n \overset{D}{\rightarrow} N(0, 1) \). As a consequence, under the conditions of Theorem 1, it holds that \( T_R \overset{D}{\rightarrow} N(0, 1) \) as well.

Next we proceed to show that the rates of convergence in Theorem 3 also apply to \( T_R \). It follows from the definitions of \( T_n \) and \( T_R \) that for \( x > 0 \) (similar analysis applies for \( x \leq 0 \)),

\[
\mathbb{P}(T_R > x) = \mathbb{P}\left( T_n > x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} \right).
\]

Thus it holds that for \( x > 0 \),

\[
\mathbb{P}(T_R > x) - [1 - \Phi(x)] 
\leq \mathbb{P}\left( T_n > x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} \right) - \left[ 1 - \Phi\left( x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} \right) \right] 
\]

\[
+ \left[ 1 - \Phi\left( x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} \right) \right] - [1 - \Phi(x)].
\]

Note that the first term on the right hand side of the above inequality is bounded by the convergence rate in Theorem 3. As for the second term, observe that when \( 0 < x \leq cn \) for some small constant \( c > 0 \), we have

\[
\left| x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} - x \right| 
= \frac{x}{1 + \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}}} \left| \frac{2x^2 - 2n - 2}{2x^2 + n(n-3) - 2} \right| 
\leq x \cdot \left| \frac{2x^2 - 2n - 2}{2x^2 + n(n-3) - 2} \right| = O\left( x\frac{n^2}{n^2 + 1} \right).
\]
By the properties of normal distribution function, we can obtain that for $0 < x \leq cn$,

\begin{equation}
(A.82) \quad \left| 1 - \Phi \left( x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} \right) - [1 - \Phi(x)] \right| = O\left( \frac{1}{n} \right).
\end{equation}

When $x > cn$, it is easy to see that $1 - \Phi(x) \leq e^{-x^2/2} \leq C_1 n^{-1}$ for some constant $C > 0$ depending on $c$. In addition, it holds that

\[
x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} = \sqrt{\frac{n(n-1)}{2 + n(n-3)/x^2 - 2/x^2}} \geq C_2 n,
\]

where $C_2 > 0$ is some constant depending on $c$. Then it follows that for some positive constant $C_3$ depending on $c$,

\[
1 - \Phi \left( x \cdot \sqrt{\frac{n(n-1)}{2x^2 + n(n-3) - 2}} \right) \leq C_3 n^{-1}.
\]

Thus (A.82) still holds for the case of $x > cn$. In view of the convergence rate in Theorem 3, it is easy to see that $O\left( \frac{1}{n} \right)$ is of a smaller order. Finally, we obtain that the same convergence rate as stated in Theorem 3 also applies to $T_R$, which completes the proof of Proposition 5.

REFERENCES


