Supplementary material for "Testing mutual independence in high dimension via distance covariance"

By Shun Yao, Xianyang Zhang, Xiaofeng Shao *

1 Technical appendix

1.1 Hoeffding decomposition

For the kernel h defined in Lemma 2.1, define that $h_c(z_1, \ldots, z_c) = \mathbb{E}h(z_1, \ldots, z_c, Z_{c+1}, \cdots, Z_4)$, where $Z_i = (X_i, Y_i) =^D (X, Y)$ for $c = 1, 2, 3, 4$. Let $z = (x, y)$, $z' = (x', y')$, $z'' = (x'', y'')$ and $z''' = (x''', y''')$. Let (X', Y'), (X'', Y'') and (X''', Y''') be independent copies of (X, Y). Direct calculation yields that

$$
\begin{aligned}
h_1(z) =& \frac{1}{2}\Big\{\mathbb{E}|x - X|(|Y' - Y''| + |y - Y| - |y - Y'| - |Y - Y'|) \\
& + \mathbb{E}|X - X'|(|y - Y''| + |Y - Y'| - |y - Y| - |Y' - Y''|)\Big\} \\
=& \frac{1}{2}\Big\{\mathbb{E}|x - X|[V(Y', Y'') + V(y, Y) - V(y, Y') - V(Y, Y'')] \\
& + \mathbb{E}|X - X'|[V(y, Y'') + V(Y, Y') - V(y, Y) - V(Y', Y'')]\Big\} \\
=& \frac{1}{2}\Big\{\mathbb{E}(|x - X| - |X - X'|)[V(y, Y) - V(Y, Y'')] + \mathbb{E}|X - X'|V(Y, Y')\Big\} \\
=& \frac{1}{2}\Big\{\mathbb{E}U(x, X)V(y, Y) + dCov^2(X, Y)\Big\}.
\end{aligned}
$$

*Address correspondence to Xianyang Zhang (zhangxiany@tamu.edu), Assistant Professor, Department of Statistics, Texas A&M University. Shun Yao (shunyao2@illinois.edu) is Ph.D. candidate and Xiaofeng Shao (xshao@illinois.edu) is Professor at Department of Statistics, University of Illinois at Urbana-Champaign.

Similarly, we obtain

$$\begin{aligned}h_2(z,z') =& \frac{1}{6}\Big\{U(x,x')V(y,y') + dCov^2(X,Y)\\ &+ \mathbb{E}U(x,X)(2V(y,Y) - V(y',Y))\\ &+ \mathbb{E}U(x',X)(2V(y',Y) - V(y,Y))\Big\},\end{aligned}$$

and

$$\begin{aligned}h_3(z,z',z'') =& \frac{1}{12}\Big\{(2U(x,x') - U(x',x'') - U(x,x''))V(y,y')\\ &+ (2U(x,x'') - U(x,x') - U(x',x''))V(y,y'')\\ &+ (2U(x',x'') - U(x,x') - U(x,x''))V(y',y'')\\ &+ \mathbb{E}(2U(x,X) - U(x',X) - U(x'',X))V(y,Y)\\ &+ \mathbb{E}(2U(x',X) - U(x,X) - U(x'',X))V(y',Y)\\ &+ \mathbb{E}(2U(x'',X) - U(x,X) - U(x',X))V(y'',Y)\Big\},\end{aligned}$$

and

$$\begin{aligned}&h_4(z,z',z'',z''')\\ =&\frac{1}{12}\Big\{(2U(x,x') + 2U(x'',x''') - U(x,x'') - U(x,x''') - U(x',x'') - U(x',x'''))(V(y,y') + V(y'',y'''))\\ &+ (2U(x,x'') + 2U(x',x''') - U(x,x') - U(x,x''') - U(x'',x') - U(x'',x'''))(V(y,y'') + V(y',y'''))\\ &+ (2U(x,x''') + 2U(x'',x') - U(x,x'') - U(x,x') - U(x''',x'') - U(x''',x'))(V(y,y''') + V(y',y''))\Big\}.\end{aligned}$$

1.1.1 Proof of Proposition 2.1

Proof. For the ease of notation, we drop the subscript n, that is, $(X_i, Y_i) =^D (X, Y)$, where the distribution of (X, Y) is allowed to depend on n. Under the null of mutual independence between X and Y, $dCov^2(X,Y) = 0$. It can be easily seen that $h_1(z) = 0$. And h_2 and h_3 can be simplified as,

$$h_2(z,z') = \frac{1}{6}U(x,x')V(y,y'),$$

and

$$\begin{aligned}h_3(z,z',z'') =& \frac{1}{12}\Big\{(2U(x,x') - U(x',x'') - U(x,x''))V(y,y') \\ & + (2U(x,x'') - U(x,x') - U(x',x''))V(y,y'') \\ & + (2U(x',x'') - U(x,x') - U(x,x''))V(y',y'')\Big\}.\end{aligned}$$

We deduce that

$$\mathrm{var}(h_2(Z,Z')) = \frac{1}{36}\mathbb{E}U(X,X')^2V(Y,Y')^2 := \nu^2,$$

and

$$\begin{aligned}\mathrm{var}(h_3(Z,Z',Z'')) =& \frac{3}{144}\mathrm{var}\{(2U(X,X') - U(X',X'') - U(X,X''))V(Y,Y')\} \\ =& \frac{1}{24}\Big[2\mathbb{E}U(X,X')^2V(Y,Y')^2 + \mathbb{E}U(X,X'')^2V(Y,Y')^2\Big] \\ =& o(n\nu^2),\end{aligned}$$

and also

$$\begin{aligned}\mathrm{var}(h_4(Z,Z',Z'',Z''')) =& \frac{6}{144}\mathbb{E}V(Y,Y')^2[U(X,X'') + U(X',X''') + U(X',X'') \\ & + U(X,X''') - 2U(X,X') - 2U(X'',X''')]^2 \\ =& \frac{1}{6}\{\mathbb{E}V(Y,Y')^2U(X,X'')^2 + \mathbb{E}U(X,X')^2\mathbb{E}V(Y,Y')^2 \\ & + \mathbb{E}U(X,X')^2V(Y,Y')^2\} \\ =& o(n^2\nu^2).\end{aligned}$$

The sample distance covariance can be decomposed as in (4) under the null. The readers are referred to Serfling (1980) for more details.

Under the local alternative, we assume that

$$\mathrm{var}(K(X,Y)) = o(n^{-1}\nu^2), \quad \mathrm{var}(K(X,Y')) = o(\nu^2).$$

This condition implies that

$$\mathrm{var}(h_1(Z)) = o(n^{-1}\nu^2), \quad \mathrm{var}(h_2(Z,Z')) = \nu^2(1+o(1)).$$

Moreover, we have

$$\begin{aligned}\mathrm{var}(h_3(Z,Z',Z'')) \le& C\Big\{\nu^2 + \mathbb{E}U(X,X'')^2V(Y,Y')^2 + \mathbb{E}U(X,X'')U(X',X'')V(Y,Y')^2\Big\} \\ \le& C\Big\{\nu^2 + \mathbb{E}U(X,X'')^2V(Y,Y')^2\Big\},\end{aligned}$$

and

$$\mathrm{var}(h_4(Z,Z',Z'',Z''')) \leq C'\left\{\nu^2 + \mathbb{E}U(X,X'')^2V(Y,Y')^2 + \mathbb{E}U(X,X')^2\mathbb{E}V(Y,Y')^2\right\},$$

where C and C' are some constants which are independent of n and p. Therefore, the same decomposition can be derived under assumptions (1)-(3). $\square$

1.2 Proofs of the main results

1.2.1 Proof of Lemma 2.1

Proof. Denote $\mathbf{1} \in \mathbb{R}^n$ as the vector of all ones, $(n)_k = n!/(n-k)!$, I_k^n is the collections of k-tuples of indices from $\{1,2,\dots,n\}$ such that each index occurs only once. By Lemma 1 of Park et al. (2015), it can be shown that

$$\begin{aligned} dCov_n^2(X,Y) &= \frac{1}{n(n-3)}\left(\mathrm{tr}(AB) + \frac{\mathbf{1}^TA\mathbf{1}\mathbf{1}^TB\mathbf{1}}{(n-1)(n-2)} - \frac{2\mathbf{1}^TAB\mathbf{1}}{(n-2)}\right) \\ &= (n)_4^{-1}\sum_{(i,j,k,l)\in I_4^n}(A_{ij}B_{ij} + A_{ij}B_{kl} - 2A_{ij}B_{ik}) \\ &= \frac{1}{\binom{n}{4}}\sum_{i<j<k<l} h(Z_i,Z_j,Z_k,Z_l) \end{aligned}$$

where

$$\begin{aligned} h(Z_i,Z_j,Z_k,Z_l) &= \frac{1}{4!}\sum_{(s,t,u,v)}^{(i,j,k,l)}(A_{st}B_{st} + A_{st}B_{uv} - 2A_{st}B_{su}) \\ &= \frac{1}{6}\sum_{s<t,u<v}^{(i,j,k,l)}(A_{st}B_{st} + A_{st}B_{uv}) - \frac{1}{12}\sum_{(s,t,u)}^{(i,j,k,l)}A_{st}B_{su} \end{aligned}$$

with $Z_i = (X_i,Y_i)$, and the last summation is over all permutations of the 4-tuples of indices (i,j,k,l). It is straightforward to verify that

$$\mathbb{E}\left[\sum_{(i,j)\in I_2^n}A_{ij}B_{ij}\right] = \mathbb{E}[\mathrm{tr}(AB)] = (n)_2\cdot\mathbb{E}|X-X'||Y-Y'|,$$

$$\mathbb{E}\left[\sum_{(i,j,q,r)\in I_4^n}A_{ij}B_{qr}\right] = \mathbb{E}[\mathbf{1}^TA\mathbf{1}\mathbf{1}^TB\mathbf{1} - 4\mathbf{1}^TAB\mathbf{1} + 2\mathrm{tr}(AB)] = (n)_4\cdot\mathbb{E}|X-X'|\mathbb{E}|Y-Y'|,$$

$$\mathbb{E}\left[\sum_{(i,j,r)\in I_3^n}A_{ij}B_{ir}\right] = \mathbb{E}[\mathbf{1}^TAB\mathbf{1} - \mathrm{tr}(AB)] = (n)_3\cdot\mathbb{E}|X-X'||Y-Y''|.$$

Therefore, $dCov_n^2(X,Y)$ is unbiased and it is a fourth-order U-statistic. $\square$

1.2.2 Proof of Theorem 3.1

Define the following quantities,

$$\mathcal{V}_1 = \mathbb{E}[H(W,W')^2 H(W,W'')^2],$$
$$\mathcal{V}_2 = \mathbb{E}[H(W,W')H(W,W'')H(W''',W')H(W''',W'')],$$
$$\mathcal{V}_3 = \mathbb{E}[H(W,W')^4].$$

We first present the following three propositions.

Proposition 1.1. *Define $M_r := \sum_{j=2}^{r}\sum_{i=1}^{j-1} H(W_i,W_j)$. Then M_r is a martingale relative to the natural filtration with respect to $\{W_i\}_{i=1}^{r}$.*

Proof. Define the natural filtration $\mathcal{F}_j = \sigma(W_1, W_2, ..., W_j)$. Notice that under the null

$$\mathbb{E}[H(W_i,W_j)] = \mathbb{E}[H(W_i,W_j)|W_i] = \mathbb{E}[H(W_i,W_j)|W_j] = 0.$$

It follows that $M_r \in \mathcal{F}_r$ and $\mathbb{E}(M_r) = 0$. For any $s \geq r$,

$$\begin{aligned}
\mathbb{E}(M_s|\mathcal{F}_r) =& \sum_{j=2}^{r}\sum_{i=1}^{j-1} H(W_i,W_j) + \mathbb{E}\left[\sum_{j=r+1}^{s}\sum_{i=1}^{j-1} H(W_i,W_j)\middle|\mathcal{F}_r\right] \\
=& M_r + \sum_{j=r+1}^{s}\sum_{i=1}^{r} \mathbb{E}\left[H(W_i,W_j)|\mathcal{F}_r\right] + \sum_{j=r+2}^{s}\sum_{i=r+1}^{j-1} \mathbb{E}\left[H(W_i,W_j)\right] \\
=& M_r + \sum_{j=r+1}^{s}\sum_{i=1}^{r} \mathbb{E}\left[\sum_{1\leq l<m\leq p} U_l(W_i^{(l)},W_j^{(l)})U_m(W_i^{(m)},W_j^{(m)})|W_i\right] \\
=& M_r.
\end{aligned}$$

Therefore, M_r is a zero mean martingale sequence. □

Proposition 1.2. *Define $\mathcal{W}_j = \sum_{i=1}^{j-1} H(W_i,W_j)$ and the natural filtration $\mathcal{F}_j$ with respect to W_j. Then under the assumption that*

$$\frac{\mathcal{V}_1}{nS^4} \to 0, \quad \frac{\mathcal{V}_2}{S^4} \to 0, \tag{1}$$

we have

$$B_n^{-2}\sum_{j=2}^{n} \mathbb{E}(\mathcal{W}_j^2|\mathcal{F}_{j-1}) \to^p 1, \tag{2}$$

where $B_n^2 = n(n-1)\mathcal{S}^2/2$.

Proof of Proposition 1.2. Notice that

$$\begin{aligned}\sum_{j=2}^{n}\mathbb{E}[\mathcal{W}_j^2] =& \sum_{j=2}^{n}\mathbb{E}\Big[\sum_{i,i'=1}^{j-1}\sum_{1\le l<m\le p} U_l(W_i^{(l)},W_j^{(l)})U_m(W_i^{(m)},W_j^{(m)})\\
&\cdot \sum_{1\le l'<m'\le p} U_{l'}(W_{i'}^{(l')},W_j^{(l')})U_{m'}(W_{i'}^{(m')},W_j^{(m')})\Big]\\
=&\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathbb{E}H(W_i,W_j)^2\\
=&\frac{n(n-1)}{2}S^2 = B_n^2.\end{aligned}$$

Define $L_j(W_i,W_k)=\mathbb{E}[H(W_i,W_j)H(W_k,W_j)|\mathcal{F}_{j-1}]$ for $i,k<j$, and note that

$$\mathbb{E}[\mathcal{W}_j^2|\mathcal{F}_{j-1}] = \mathbb{E}[\sum_{i=1}^{j-1}\sum_{k=1}^{j-1}H(W_i,W_j)H(W_k,W_j)|\mathcal{F}_{j-1}] = \sum_{i=1}^{j-1}\sum_{k=1}^{j-1}L_j(W_i,W_k).$$

If $i\le k$ and $i'\le k'$ then

$$\begin{aligned}&\mathbb{E}[L_j(W_i,W_k)L_{j'}(W_{i'},W_{k'})]\\
=&\mathbb{E}H(W,W')^2H(W,W'')^2 && \text{if } i=k=i'=k',\\
=&\mathbb{E}[H(W,W')H(W,W'')H(W''',W')H(W''',W'')] && \text{if } i=i'\neq k=k',\ \text{or } i=k'\neq k=i',\\
=&[\mathbb{E}H(W,W')^2]^2 && \text{if } i=k\neq i'=k',\\
=&0 && \text{otherwise,}\end{aligned}$$

and also

$$\begin{aligned}&\mathbb{E}[L_j(W_i,W_k)]\mathbb{E}[L_{j'}(W_{i'},W_{k'})]\\
=&\mathbb{E}H(W_i,W_j)H(W_k,W_j)\mathbb{E}H(W_{i'},W_{j'})H(W_{k'},W_{j'})\\
=&[\mathbb{E}H(W,W')^2]^2 \qquad \text{if } i=k,\ i'=k',\\
=&0 \qquad \text{otherwise.}\end{aligned}$$

Therefore,

$$\begin{aligned}\operatorname{var}\left(\sum_{j=2}^{n}\mathbb{E}[\mathcal{W}_j^2|\mathcal{F}_{j-1}]\right) =& \sum_{j,j'=2}^{n}\sum_{i,k=1}^{j-1}\sum_{i',k'=1}^{j'-1}\operatorname{cov}(L_j(W_i,W_k),L_{j'}(W_{i'},W_{k'}))\\
=&\sum_{j=j'}[(j-1)\mathcal{V}_1+2(j-1)(j-2)\mathcal{V}_2-(j-1)S^4]\\
&+\ 2\sum_{2\le j<j'\le n}[(j-1)\mathcal{V}_1+2(j-1)(j-2)\mathcal{V}_2-(j-1)S^4].\end{aligned}$$

Under the assumption (1), we have

$$\frac{4}{n^2(n-1)^2S^4}\ \mathrm{var}\left(\sum_{j=2}^{n}\mathbb{E}[\mathcal{W}_j^2|\mathcal{F}_{j-1}]\right)\to 0.$$

Therefore (2) holds. □

Proposition 1.3. *Define $\mathcal{W}_j=\sum_{i=1}^{j-1}H(W_i,W_j)$ and the natural filtration $\mathcal{F}_j$ with respect to W_j. Under the assumption*

$$\frac{\mathcal{V}_1}{nS^4}\to 0,\quad \frac{\mathcal{V}_3}{n^2S^4}\to 0, \tag{3}$$

we have

$$\sum_{j=2}^{n}B_n^{-2}\mathbb{E}\left(\mathcal{W}_j^2\ \mathbf{I}(|\mathcal{W}_j|>\epsilon B_n)|\mathcal{F}_{j-1}\right)\to^p 0, \tag{4}$$

where $B_n=n(n-1)\mathcal{S}^2/2$.

Proof of Proposition 1.3. Notice that

$$\sum_{j=2}^{n}B_n^{-2}\mathbb{E}\left(\mathcal{W}_j^2\ \mathbf{1}(|\mathcal{W}_j|>\epsilon B_n)|\mathcal{F}_{j-1}\right)\le\sum_{j=2}^{n}B_n^{-2}(\epsilon B_n)^{-s}\mathbb{E}\left(|\mathcal{W}_j|^{2+s}|\mathcal{F}_{j-1}\right)$$

for some $s>0$. It suffices to show that for $s=2$

$$\sum_{j=2}^{n}B_n^{-4}\mathbb{E}\left(\mathcal{W}_j^4|\mathcal{F}_{j-1}\right)\to^p 0.$$

To this end, we show that

$$\sum_{j=2}^{n}B_n^{-4}\mathbb{E}\left(\mathcal{W}_j^4\right)\to^p 0. \tag{5}$$

Some algebra yields that

$$\begin{aligned}\sum_{j=2}^{n}\mathbb{E}[\mathcal{W}_j^4]=&\sum_{j=2}^{n}\sum_{i_1,i_2,i_3,i_4=1}^{j-1}\mathbb{E}H(W_{i_1},W_j)H(W_{i_2},W_j)H(W_{i_3},W_j)H(W_i,W_j)\\=&\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathbb{E}[H(W_i,W_j)^4]+3\sum_{j=2}^{n}\sum_{i\neq i'}^{j-1}\mathbb{E}[H(W_i,W_j)^2H(W_{i'},W_j)^2]\\=&\frac{n(n-1)}{2}\mathcal{V}_3+O(n^3\mathcal{V}_1).\end{aligned}$$

Therefore, under (3), (5) holds. □

We present the following lemma which is useful in the proof of Theorem 3.1.

Lemma 1.1. *Let* $a(x) = \max\{|\mathbb{E}[|X - X'|] - 2\mathbb{E}[|x - X'|]|, \mathbb{E}[|X - X'|]\}$. *Then we have* $|U(x, x')| \leq \max\{a(x), a(x')\}$.

Proof of Lemma 1.1. By the triangle inequality, we have $|\mathbb{E}[|X - x'|] - |x - x'|| \leq \mathbb{E}[|x - X'|]$ for $x, x' \in \mathbb{R}$. Thus $|U(x, x')| \leq \max\{|\mathbb{E}[|X - X'|] - 2\mathbb{E}[|x - X'|]|, \mathbb{E}[|X - X'|]\} = a(x)$. Switching x and x', we get $|U(x, x')| \leq a(x')$. The conclusion thus follows. $\square$

Proof of Theorem 3.1. We show that Assumption A1 implies both (1) and (3) under the null, i.e., $\frac{\mathcal{V}_1}{nS^4} \to 0$, $\frac{\mathcal{V}_2}{S^4} \to 0$ and $\frac{\mathcal{V}_3}{n^2 S^4} \to 0$. We write $a \lesssim b$ if a is less or equal to b up to a multiplicative constant. By Lemma 1.1 and the fact that $\mathbb{E}[a(X)] \lesssim E[|X - \mathbb{E}[X]|]$, we have

$$\begin{aligned}
\frac{\sum_{l=1}^p dCov^4(W^{(l)})}{[\sum_{l=1}^p dCov^2(W^{(l)})]^2} &= \frac{\sum_{l=1}^p \{\mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2]\}^2}{[\sum_{l=1}^p dCov^2(W^{(l)})]^2} \\
&\leq \frac{\sum_{l=1}^p \{\mathbb{E}[a(W^{(l)})]\}^4}{[\sum_{l=1}^p dCov^2(W^{(l)})]^2} \\
&\lesssim \frac{\sum_{l=1}^p \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^4}{[\sum_{l=1}^p dCov^2(W^{(l)})]^2}.
\end{aligned}$$

By Assumption A1, $\sum_{l=1}^p dCov^4(W^{(l)}) = o([\sum_{l=1}^p dCov^2(W^{(l)})]^2)$. Therefore, we have

$$\begin{aligned}
2S^2 &= \sum_{l \neq m} dCov^2(W^{(l)}) dCov^2(W^{(m)}) \\
&= \left\{\sum_{l=1}^p dCov^2(W^{(l)})\right\}^2 - \sum_{l=1}^p dCov^4(W^{(l)}) \\
&= \left\{\sum_{l=1}^p dCov^2(W^{(l)})\right\}^2 \cdot \{1 + o(1)\}.
\end{aligned}$$

Again using Lemma 1.1 and the fact that $\mathbb{E}[a(X)^2] \lesssim \mathrm{var}(X)$, we have

$$\begin{aligned}
\mathcal{V}_1 =& \mathbb{E}[H(W, W')^2 H(W, W'')^2] \\
=& \sum_{l<m} \sum_{l'<m'} \mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2 U_m(W^{(m)}, W'^{(m)})^2 U_{l'}(W^{(l')}, W''^{(l')})^2 U_{m'}(W^{(m')}, W''^{(m')})^2] \\
\lesssim& \left\{\sum_l \mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2]\right\}^4 + \left\{\sum_l \mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2 U_l(W^{(l)}, W''^{(l)})^2]\right\}^2 \\
&+ \left\{\sum_l \mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2 U_l(W^{(l)}, W''^{(l)})^2]\right\} \cdot \left\{\sum_l \mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2]\right\}^2 \\
\lesssim& \left\{\sum_l dCov^2(W^{(l)})\right\}^4 + \left\{\sum_l \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^2 \mathrm{var}(W^{(l)})\right\}^2.
\end{aligned}$$

Together with Assumption A1, we can show that

$$\frac{\mathcal{V}_1}{nS^4} \lesssim \left\{\sum_l dCov^2(W^{(l)})\right\}^4 \cdot \frac{1}{nS^4} + \left\{\sum_l \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^2 \mathrm{var}(W^{(l)})\right\}^2 \cdot \frac{1}{nS^4} \to 0,$$

where we have used the Cauchy-Schwarz inequality to show $\left\{\sum_l \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^2 \mathrm{var}(W^{(l)})\right\}^2 \leq \sum_l \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^4 \sum_l \mathrm{var}(W^{(l)})^2$. Similarly, we have

$$\begin{aligned}
\mathcal{V}_2 =& \mathbb{E}[H(W,W')H(W,W'')H(W''',W')H(W''',W'')] \\
=& \sum_{1\leq l<m\leq p} \mathbb{E}[U_l(W^{(l)},W'^{(l)})U_l(W^{(l)},W''^{(l)})U_l(W'''^{(l)},W'^{(l)})U_l(W'''^{(l)},W''^{(l)}) \\
& \cdot U_m(W^{(m)},W'^{(m)})U_m(W^{(m)},W''^{(m)})U_m(W'''^{(m)},W'^{(m)})U_m(W'''^{(m)},W''^{(m)})] \\
\leq& \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})U_l(W^{(l)},W''^{(l)})U_l(W'''^{(l)},W'^{(l)})U_l(W'''^{(l)},W''^{(l)})]\right\}^2 \\
\lesssim& \left\{\sum_l \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^4\right\}^2,
\end{aligned}$$

which implies that

$$\frac{\mathcal{V}_2}{S^4} \leq \left\{\frac{1}{S^2}\sum_l \{\mathbb{E}[|W^{(l)} - \mu^{(l)}|]\}^4\right\}^2 \to 0.$$

Lastly, we have

$$\begin{aligned}
\mathcal{V}_3 = \mathbb{E}[H(W,W')^4] \lesssim& \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})^4]\right\}^2 + \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})^2]\right\}^4 \\
&+ \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})^4]\right\} \cdot \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})^2]\right\}^2 \\
&+ \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})^2]\right\} \cdot \left\{\sum_l \mathbb{E}[U_l(W^{(l)},W'^{(l)})^3]\right\}^2 \\
\lesssim& \left\{\sum_l \mathrm{var}(W^{(l)})^2\right\}^2 + \left\{\sum_l dCov^2(W^{(l)})\right\}^4,
\end{aligned}$$

Hence,

$$\frac{\mathcal{V}_3}{n^2 S^4} \lesssim \left\{\frac{1}{nS^2}\sum_l \mathrm{var}(W^{(l)})^2\right\}^2 + \left\{\sum_l dCov^2(W^{(l)})\right\}^4 \cdot \frac{1}{n^2 S^4} \to 0.$$

In view of Corollary 3.1 of Hall & Heyde (1980), the conclusion follows from Proposition 1.2 and 1.3. □

Theorem 3.3 and Theorem 4.1 can be proved using similar arguments in Proposition 1.2 and Proposition 1.3, we omit the details.

1.2.3 Proof of Theorem 3.2

Proof. Under the null of mutual independence, we have

$$\begin{aligned}\mathbb{E}\hat{S}^2 &= \sum_{1\le l<m\le p} \mathbb{E}[dCov_n^2(W^{(l)})dCov_n^2(W^{(m)})] \\ &= \sum_{1\le l<m\le p} dCov^2(W^{(l)})dCov^2(W^{(m)}) = S^2.\end{aligned}$$

Thus it suffice to show that

$$\mathbb{E}\left(\left|\frac{\hat{S}^2}{S^2}-1\right|^2\right) = \frac{\mathrm{var}(\hat{S}^2)}{S^4} \to 0.$$

Notice that

$$\begin{aligned}\mathrm{var}(\hat{S}^2) &= \sum_{1\le l<m\le p}\sum_{1\le l'<m'\le p} \mathrm{cov}(dCov_n^2(W^{(l)})dCov_n^2(W^{(m)}), dCov_n^2(W^{(l')})dCov_n^2(W^{(m')})) \\ &= \sum_{l<m} \mathrm{var}(dCov_n^2(W^{(l)})dCov_n^2(W^{(m)})) \\ &+ 2\sum_{l<m<m'} \mathrm{cov}(dCov_n^2(W^{(l)})dCov_n^2(W^{(m)}), dCov_n^2(W^{(l)})dCov_n^2(W^{(m')})) \\ &= \sum_{l<m} \mathrm{var}(dCov_n^2(W^{(l)}))\mathrm{var}(dCov_n^2(W^{(m)})) \\ &+ \sum_{l\neq m} \mathrm{var}(dCov_n^2(W^{(l)}))dCov^4(W^{(m)}) \\ &+ 2\sum_{l<m<m'} \mathrm{var}(dCov_n^2(W^{(l)}))dCov^2(W^{(m)})dCov^2(W^{(m')}) \\ &:= J_1 + J_2 + J_3 \quad \text{(say)}.\end{aligned}$$

Since $dCov_n^2(W^{(l)})$ is a fourth order U-statistics, by the Hoeffding decomposition, the dominant term of its variance is

$$\binom{n}{4}^{-1}\binom{4}{1}\binom{n-4}{3}\mathrm{var}(h_1(W^{(l)}))$$

with

$$\begin{aligned}\mathrm{var}(h_1(W^{(l)})) &= \frac{1}{4}\mathrm{var}(\mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2|W^{(l)}]) \\ &= \frac{1}{4}\{\mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2U_l(W^{(l)}, W''^{(l)})^2] - dCov^4(W^{(l)})\}.\end{aligned}$$

Under Assumption A1 and by Lemma 1.1, we can derive that

$$\frac{\sum_{l=1}^p \mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2U_l(W^{(l)}, W''^{(l)})^2]}{nS^2} \lesssim \frac{\sum_l\{\mathbb{E}[|W^{(l)}-\mu^{(l)}|]\}^2\mathrm{var}(W^{(l)})}{nS^2} \to 0,$$

and

$$\frac{\sum_{l=1}^{p} dCov^4(W^{(l)})}{S^2} \to 0,$$

as we have shown in the proof of Theorem 3.1. The higher order terms of the variance of $dCov_n^2(W^{(l)})$ can be handled in a similar fashion. Hence we have

$$\frac{\sum_{l=1}^{p} \mathrm{var}(dCov_n^2(W^{(l)}))}{S^2} = O\left(\frac{\sum_{l=1}^{p}\{\mathbb{E}[U_l(W^{(l)}, W'^{(l)})^2 U_l(W^{(l)}, W''^{(l)})^2] - dCov^4(W^{(l)})\}}{nS^2}\right) \to 0.$$

Therefore, we obtain that

$$\frac{J_1}{S^4} \leq \left[\frac{\sum_{l=1}^{p} \mathrm{var}(dCov_n^2(W^{(l)}))}{S^2}\right]^2 \to 0,$$

and

$$\frac{J_2}{S^4} \leq \left(\frac{\sum_{l=1}^{p} \mathrm{var}(dCov_n^2(W^{(l)}))}{S^2}\right) \cdot \left(\frac{\sum_{l=1}^{p} dCov^4(W^{(l)})}{S^2}\right) \to 0,$$

and also

$$\frac{J_3}{S^4} \leq \frac{2S^2 \sum_{l=1}^{p} \mathrm{var}(dCov_n^2(W^{(l)}))}{S^4} \to 0.$$

Thus $\hat{S}^2$ is ratio consistent under the null and Assumption A1. □

1.2.4 Proof of Theorem 3.4

Proof. When $W^{(l)}$ is standard Gaussian, we can directly calculate that $dCov^2(W^{(l)}) = f(1) = \frac{4}{\pi}(1 + \frac{\pi}{3} - \sqrt{3})$. Therefore,

$$S^2 = \sum_{1 \leq l < m \leq p} dCov^2(W^{(l)}) dCov^2(W^{(m)}) = \frac{p(p-1)}{2}[f(1)]^2.$$

Our test $\phi_{n,\alpha} = 1$ if $D_n > z_\alpha$, where z_α is $100(1-\alpha)\%$ quantile of standard normal. Hence we have

$$\begin{aligned}
1 - \mathbb{E}[\phi_{n,\alpha}] &= P\left(\frac{\sum_{1 \leq l < m \leq p} \sqrt{\binom{n}{2}} dCov_n^2(W^{(l)}, W^{(m)})}{S} \leq z_\alpha\right) \\
&= P\left(\sum_{1 \leq l < m \leq p} dCov_n^2(W^{(l)}, W^{(m)}) - |\Theta|^2 \leq z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}} - |\Theta|^2\right) \\
&\leq P\left(|\sum_{1 \leq l < m \leq p} dCov_n^2(W^{(l)}, W^{(m)}) - |\Theta|^2| \geq \left|z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}} - |\Theta|^2\right|\right) \\
&\leq \frac{\mathrm{var}[\sum_{1 \leq l < m \leq p} dCov_n^2(W^{(l)}, W^{(m)})]}{\left(z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}} - |\Theta|^2\right)^2},
\end{aligned}$$

where $z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}} - |\Theta|^2$ is negative for large enough $\tilde{c}$ and the last inequality uses the fact that $\mathbb{E}[\sum_{1\le l<m\le p} dCov_n^2(W^{(l)}, W^{(m)})] = |\Theta|^2$ and Chebyshev's inequality. Now let $Z_i^{(lm)} = (W_i^{(l)}, W_i^{(m)})$, by lemma 2.1 we have

$$\sum_{1\le l<m\le p} dCov_n^2(W^{(l)}, W^{(m)}) := \frac{1}{\binom{n}{4}} \sum_{1\le i_1<i_2<i_3<i_4\le n} h^s(W_{i_1}, W_{i_2}, W_{i_3}, W_{i_4}),$$

where

$$h^s(W_{i_1}, W_{i_2}, W_{i_3}, W_{i_4}) = \sum_{1\le l<m\le p} h(Z_{i_1}^{(lm)}, Z_{i_2}^{(lm)}, Z_{i_3}^{(lm)}, Z_{i_4}^{(lm)}).$$

Therefore $\sum_{1\le l<m\le p} dCov_n^2(W^{(l)}, W^{(m)})$ is a fourth order U-statistic with kernel h^s and its variance is given by

$$\operatorname{var}\left[\sum_{1\le l<m\le p} dCov_n^2(W^{(l)}, W^{(m)})\right] = \binom{n}{4}^{-1} \sum_{c=1}^{4} \binom{4}{c}\binom{n-4}{4-c} \operatorname{var}(h_c^s) \le C \sum_{c=1}^{4} \operatorname{var}(h_c^s) n^{-c},$$

for some constant $C > 0$. Here $h_c^s = \sum_{1\le l<m\le p} h_c(z_1^{(lm)}, \ldots, z_c^{(lm)})$ for $c = 1,2,3,4$ with $h_c(z_1^{(lm)}, \ldots, z_c^{(lm)}) = Eh(z_1^{(lm)}, \ldots, z_c^{(lm)}, Z_{c+1}^{(lm)}, ;Z_4^{(lm)})$ defined in Section 1.1.

Use the results from Lemma 1.2 or similar arguments from the proof, we can work out the variance of the fourth order U-statistic. Specifically, for some constant c', we have

$$\begin{aligned}
4\operatorname{var}(h_1^s) =& \sum_{\substack{1\le l<m\le p\\ 1\le l'<m'\le p}} \mathbb{E}[h_1(Z_{i_1}^{(lm)}, Z_{i_2}^{(lm)}, Z_{i_3}^{(lm)}, Z_{i_4}^{(lm)}) h_1(Z_{i_1}^{(l'm')}, Z_{i_5}^{(l'm')}, Z_{i_6}^{(l'm')}, Z_{i_7}^{(l'm')})] - |\Theta|^4 \\
=& \sum_{\substack{1\le l<m\le p\\ 1\le l'<m'\le p}} \mathbb{E}[U(W^{(l)}, W'^{(l)}) U(W^{(m)}, W'^{(m)}) U(W^{(l')}, W''^{(l')}) U(W^{(m')}, W''^{(m')})] - |\Theta|^4 \\
\le& c'\Big\{p^4 \mathbb{E}[U(W^{(1)}, W'^{(1)}) U(W^{(2)}, W'^{(2)}) U(W^{(3)}, W''^{(3)}) U(W^{(4)}, W''^{(4)})] \\
&+ p^3 \mathbb{E}[U(W^{(1)}, W'^{(1)}) U(W^{(1)}, W''^{(1)}) U(W^{(2)}, W'^{(2)}) U(W^{(3)}, W''^{(3)})] \\
&+ p^2 \mathbb{E}[U(W^{(1)}, W'^{(1)}) U(W^{(1)}, W''^{(1)}) U(W^{(2)}, W'^{(2)}) U(W^{(2)}, W''^{(2)})]\Big\} \\
=& O(|\Theta|^4) + O(|\Theta|^3) + O(|\Theta|^2).
\end{aligned}$$

Therefore, we have

$$\frac{Cn^{-1}\operatorname{var}(h_1^s)}{\left(z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}} - |\Theta|^2\right)^2} \le \frac{Cn^{-1}\{O(|\Theta|^4) + O(|\Theta|^3) + O(|\Theta|^2)\}}{z_\alpha^2 f(1)^2 \frac{p(p-1)}{n(n-1)} + |\Theta|^4 - 2z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}}|\Theta|^2}.$$

Hence, the right hand side can be made less than $\frac{1-\beta}{4}$ when $p/n \to \lambda \in (0,\infty)$, and also

$|\Theta|^2 > \tilde{c}$ for some large enough constant $\tilde{c} = \tilde{c}(\alpha, \beta, \lambda)$. Similarly, we have

$$
\begin{aligned}
\operatorname{var}(h_2^s) \leq & c'\Big\{ \sum_{\substack{1\leq l<m\leq p\\ 1\leq l'<m'\leq p}} \mathbb{E}[U(W^{(l)}, W'^{(l)})U(W^{(m)}, W'^{(m)})U(W^{(l')}, W'^{(l')})U(W^{(m')}, W'^{(m')})] \\
& + \sum_{\substack{1\leq l<m\leq p\\ 1\leq l'<m'\leq p}} \mathbb{E}[U(W^{(l)}, W'^{(l)})U(W^{(m)}, W'^{(m)})U(W^{(l')}, W''^{(l')})U(W^{(m')}, W''^{(m')})] \\
& + \sum_{\substack{1\leq l<m\leq p\\ 1\leq l'<m'\leq p}} \mathbb{E}[U(W^{(l)}, W'^{(l)})U(W^{(m)}, W''^{(m)})U(W'''^{(l')}, W'^{(l')})U(W'''^{(m')}, W''^{(m')})]\Big\}.
\end{aligned}
$$

In particular

$$
\begin{aligned}
& \sum_{\substack{1\leq l<m\leq p\\ 1\leq l'<m'\leq p}} \mathbb{E}[U(W^{(l)}, W'^{(l)})U(W^{(m)}, W'^{(m)})U(W^{(l')}, W'^{(l')})U(W^{(m')}, W'^{(m')})] \\
\leq & c'\Big\{p^4\mathbb{E}[U(W^{(1)}, W'^{(1)})U(W^{(2)}, W'^{(2)})U(W^{(3)}, W'^{(3)})U(W^{(4)}, W'^{(4)})] \\
& + p^3\mathbb{E}[U(W^{(1)}, W'^{(1)})^2U(W^{(2)}, W'^{(2)})U(W^{(3)}, W'^{(3)})] \\
& + p^2\mathbb{E}[U(W^{(1)}, W'^{(1)})^2U(W^{(2)}, W'^{(2)})^2]\Big\} \\
= & O(|\Theta|^4) + O(p|\Theta|^2) + O(p^2),
\end{aligned}
$$

and also

$$
\begin{aligned}
& \sum_{\substack{1\leq l<m\leq p\\ 1\leq l'<m'\leq p}} \mathbb{E}[U(W^{(l)}, W'^{(l)})U(W^{(m)}, W''^{(m)})U(W'''^{(l')}, W'^{(l')})U(W'''^{(m')}, W''^{(m')})] \\
\leq & c'\Big\{p^4\mathbb{E}[U(W^{(1)}, W'^{(1)})U(W^{(2)}, W''^{(2)})U(W'''^{(3)}, W'^{(3)})U(W'''^{(4)}, W''^{(4)})] \\
& + p^3\mathbb{E}[U(W^{(1)}, W'^{(1)})U(W'''^{(1)}, W''^{(1)})U(W^{(2)}, W''^{(2)})U(W'''^{(3)}, W'^{(3)})] \\
& + p^3\mathbb{E}[U(W^{(1)}, W'^{(1)})U(W'''^{(1)}, W'^{(1)})U(W^{(2)}, W''^{(2)})U(W'''^{(3)}, W''^{(3)})] \\
& + p^2\mathbb{E}[U(W^{(1)}, W'^{(1)})U(W'''^{(1)}, W'^{(1)})U(W^{(2)}, W''^{(2)})U(W'''^{(2)}, W''^{(2)})]\Big\} \\
= & O(|\Theta|^4) + O(|\Theta|^3) + O(|\Theta|^2).
\end{aligned}
$$

Therefore,

$$
\frac{Cn^{-2}\operatorname{var}(h_2^s)}{\left(z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}} - |\Theta|^2\right)^2} \leq \frac{Cn^{-2}\{O(|\Theta|^4) + O(|\Theta|^3) + O(|\Theta|^2) + O(p|\Theta|^2) + O(p^2)\}}{z_\alpha^2 f(1)^2\frac{p(p-1)}{n(n-1)} + |\Theta|^4 - 2z_\alpha f(1)\sqrt{\frac{p(p-1)}{n(n-1)}}|\Theta|^2}.
$$

The right hand side can also be made less than $\frac{1-\beta}{4}$ when $p/n \to \lambda \in (0, \infty)$ and $\tilde{c}$ is large.

Using similar arguments, we can show that $\operatorname{var}(h_3^s), \operatorname{var}(h_4^s) = O(\operatorname{var}(h_2^s))$ and accordingly we obtain that $1 - \mathbb{E}[\phi_{n,\alpha}] \leq 1 - \beta$ as $p/n \to \lambda$ and the theorem is proved.

□

Lemma 1.2. *For multivariate Gaussian (W_1, W_2, W_3, W_4) with pairwise correlation ρ, we have*

$$\mathbb{E}[U(W_1, W_1')U(W_2, W_2')U(W_3, W_3')U(W_4, W_4')] \le C'|\rho|^4, \tag{6}$$
$$\mathbb{E}[U(W_1, W_1')U(W_2, W_2')U(W_3, W_3'')U(W_4, W_4'')] \le C'|\rho|^4,$$
$$\mathbb{E}[U(W_1, W_1')U(W_1, W_1'')U(W_2, W_2')U(W_3, W_3'')] \le C'|\rho|^3,$$
$$\mathbb{E}[U(W_1, W_1')U(W_1, W_1'')U(W_2, W_2')U(W_2, W_2'')] \le C'|\rho|^2,$$
$$\mathbb{E}[U(W_1, W_1')U(W_2, W_2'')U(W_3''', W_3')U(W_4''', W_4'')] \le C'|\rho|^4,$$
$$\mathbb{E}[U(W_1, W_1')U(W_1''', W_1')U(W_2, W_2'')U(W_3''', W_3'')] \le C'|\rho|^3,$$
$$\mathbb{E}[U(W_1, W_1')U(W_1''', W_1')U(W_2, W_2'')U(W_2''', W_2'')] \le C'|\rho|^2,$$
$$\mathbb{E}[U(W_1, W_1')^2U(W_2, W_2')U(W_3, W_3')] \le C'|\rho|^2,$$

for some positive constant C' which is different from line to line.

Proof. We provide the details for (6). The other inequalities can be obtained in a similar way. Using Lemma 1 in Szekeley et al. (2007), we can show that

$$U(W_1, W_1') = \int_{\mathbb{R}} (f(t_1) - e^{it_1W_1})(\overline{f(t_1)} - e^{-it_1W_1})\frac{dt_1}{\pi t_1^2},$$

where $\overline{f(t)} = f(t) = e^{-t^2/2}$. Therefore,

$$\begin{aligned}
&\mathbb{E}[U(W_1, W_1')U(W_2, W_2')U(W_3, W_3')U(W_4, W_4')]\\
=&\mathbb{E}\Big\{\int_{\mathbb{R}^4} \pi^{-4}(e^{-t_1^2/2} - e^{it_1W_1})(e^{-t_2^2/2} - e^{it_2W_2})(e^{-t_3^2/2} - e^{it_3W_3})(e^{-t_4^2/2} - e^{it_4W_4})\\
&\times (e^{-t_1^2/2} - e^{it_1W_1'})(e^{-t_2^2/2} - e^{it_2W_2'})(e^{-t_3^2/2} - e^{it_3W_3'})(e^{-t_4^2/2} - e^{it_4W_4'})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\Big\}\\
=&\int_{\mathbb{R}^4} \pi^{-4}\left|\mathbb{E}(e^{-t_1^2/2} - e^{it_1W_1})(e^{-t_2^2/2} - e^{it_2W_2})(e^{-t_3^2/2} - e^{it_3W_3})(e^{-t_4^2/2} - e^{it_4W_4})\right|^2\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}.
\end{aligned}$$

It is straightforward to verify that

$$\begin{aligned}
&\mathbb{E}(e^{-t_1^2/2} - e^{it_1W_1})(e^{-t_2^2/2} - e^{it_2W_2})(e^{-t_3^2/2} - e^{it_3W_3})(e^{-t_4^2/2} - e^{it_4W_4})\\
=&e^{-\frac{t_1^2+t_2^2+t_3^2+t_4^2}{2}}(e^{-\rho t_1t_2} + e^{-\rho t_1t_3} + e^{-\rho t_1t_4} + e^{-\rho t_2t_3} + e^{-\rho t_2t_4} + e^{-\rho t_3t_4}\\
&- e^{-\rho t_1t_2-\rho t_1t_3-\rho t_2t_3} - e^{-\rho t_1t_2-\rho t_1t_4-\rho t_2t_4} - e^{-\rho t_1t_3-\rho t_1t_4-\rho t_3t_4}\\
&- e^{-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4} + e^{-\rho t_1t_2-\rho t_1t_3-\rho t_1t_4-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4} - 3)\\
=&e^{-\frac{t_1^2+t_2^2+t_3^2+t_4^2}{2}}(3\rho^2t_1t_2t_3t_4 + \text{Remainder terms}),
\end{aligned}$$

where the last step uses the Taylor expansion $e^x = 1 + x + x^2/2 + \sum_3^\infty x^k/k!$. Therefore we have

$$
\begin{aligned}
&\mathbb{E}[U(W_1,W_1')U(W_2,W_2')U(W_3,W_3')U(W_4,W_4')] \\
&= \int_{\mathbb{R}^4} \pi^{-4} \left| e^{-\frac{t_1^2+t_2^2+t_3^2+t_4^2}{2}} (3\rho^2 t_1t_2t_3t_4 + \text{Remainder terms}) \right|^2 \frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2} \\
&= \int_{\mathbb{R}^4} 9\pi^{-4}\rho^4 e^{-(t_1^2+t_2^2+t_3^2+t_4^2)} dt_1dt_2dt_3dt_4 \qquad (7)\\
&\quad + \int_{\mathbb{R}^4} 6\pi^{-4}\rho^2 t_1t_2t_3t_4 e^{-(t_1^2+t_2^2+t_3^2+t_4^2)} (\text{Remainder terms}) \frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2} \qquad (8)\\
&\quad + \int_{\mathbb{R}^4} \pi^{-4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)} (\text{Remainder terms})^2 \frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}. \qquad (9)
\end{aligned}
$$

We first consider term (8). Denote $a_1 = t_1t_2$, $a_2 = t_1t_3$, $a_3 = t_1t_4$, $a_4 = t_2t_3$, $a_5 = t_2t_4$ and $a_6 = t_3t_4$. By the Vitali convergence theorem, we can show that

$$
\begin{aligned}
&\int_{\mathbb{R}^4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)} (\text{Remainder terms}) \frac{dt_1}{t_1}\frac{dt_2}{t_2}\frac{dt_3}{t_3}\frac{dt_4}{t_4} \\
&= \sum_{k=3}^{\infty} \frac{(-\rho)^k}{k!} \int_{\mathbb{R}^4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)} \Big\{ (a_1)^k + (a_2)^k + (a_3)^k + (a_4)^k + (a_5)^k + (a_6)^k \\
&\quad - (a_1+a_2+a_4)^k - (a_1+a_3+a_5)^k - (a_2+a_3+a_6)^k - (a_4+a_5+a_6)^k \\
&\quad + (a_1+a_2+a_3+a_4+a_5+a_6)^k \Big\} \frac{dt_1}{t_1}\frac{dt_2}{t_2}\frac{dt_3}{t_3}\frac{dt_4}{t_4} \\
&:= \sum_{k=3}^{\infty} \frac{(-\rho)^k}{k!} I_k.
\end{aligned}
$$

Using the multinomial expansion, we have

$$
\begin{aligned}
a(k) :=& \Big\{ (a_1)^k + (a_2)^k + (a_3)^k + (a_4)^k + (a_5)^k + (a_6)^k - (a_1+a_2+a_4)^k - (a_1+a_3+a_5)^k \\
&- (a_2+a_3+a_6)^k - (a_4+a_5+a_6)^k + (a_1+a_2+a_3+a_4+a_5+a_6)^k \Big\} \\
=& \sum_{*} \frac{k!}{k_1!k_2!k_3!k_4!k_5!k_6!} t_1^{k_1+k_2+k_3} t_2^{k_1+k_4+k_5} t_3^{k_2+k_4+k_6} t_4^{k_3+k_5+k_6},
\end{aligned}
$$

where $\sum_*$ denotes the summation over all $(k_1,k_2,k_3,k_4,k_5,k_6)$ such that $\sum_{i=1}^6 k_i = k$, $k_1 + k_2+k_3 \geq 1$, $k_1+k_4+k_5 \geq 1$, $k_2+k_4+k_6 \geq 1$ and $k_3+k_5+k_6 \geq 1$. Since $\int_{\mathbb{R}} e^{-t^2} t^{2k+1} dt = 0$ and $0 < \int_{\mathbb{R}} e^{-t^2} t^{2k} dt < \infty$, we have $I_k > 0$ for $k \geq 3$. We first consider the case $-1/3 \leq \rho < 0$. By Hölder's inequality, we have

$$
\sum_{k=3}^{\infty} \frac{(-\rho)^k}{k!} I_k = \sum_{k=3}^{\infty} \frac{|\rho|^k}{k!} I_k \leq \mathbb{E}[U(W^{(1)}, W'^{(1)})^4] < \infty,
$$

which implies that

$$\sum_{k=3}^{\infty}\frac{(-\rho)^k}{k!}I_k=\sum_{k=3}^{\infty}\frac{|\rho|^k}{k!}I_k\le\frac{|\rho|^3}{(1/3)^3}\sum_{k=3}^{\infty}\frac{(1/3)^k}{k!}I_k\le C|\rho|^3.$$

For $0\le\rho\le 1$, $\sum_{k=3}^{\infty}\frac{(-\rho)^k}{k!}I_k=\rho^3\sum_{k=3}^{\infty}\frac{(-1)^k}{k!}I_k\rho^{k-3}$. First notice that the above power series is convergent at $\rho=1$, that is,

$$\sum_{k=3}^{\infty}\frac{(-1)^k}{k!}I_k\rho^{k-3}=\sum_{k=3}^{\infty}\frac{(-1)^k}{k!}I_k\le\mathbb{E}[U(W^{(1)},W'^{(1)})^4]<\infty.$$

By the Abel theorem, the power series is continuous as a function of ρ for $\rho\in[0,1]$ and therefore bounded. Equivalently, we can use the Abel's uniform convergence test to show the power series is uniformly convergent for all $\rho\in[0,1]$. Hence, $\sum_{k=3}^{\infty}\frac{(-\rho)^k}{k!}I_k\le C|\rho|^3$ for some constant C that is independent of ρ and accordingly term (8) $\le C|\rho|^5$. Similarly , we can show that

$$\begin{aligned}&\int_{\mathbb{R}^4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(\text{Remainder terms})^2\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\rho^6\int_{\mathbb{R}^4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(\rho^{-3}\times\text{Remainder terms})^2\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\rho^6\int_{\mathbb{R}^4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}\sum_{k=6}^{\infty}(-\rho)^{k-6}J_k\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
\le&C|\rho|^6,\end{aligned}$$

where $J_k=\sum_{k_1,k_2:k_1+k_2=1}\frac{1}{k_1!k_2!}a(k_1)a(k_2)$. Therefore

$$\mathbb{E}[U(W_1,W_1')U(W_2,W_2')U(W_3,W_3')U(W_4,W_4')]\le C|\rho|^4.$$

Using similarly arguments, we can show that

$$\begin{aligned}&\mathbb{E}[U(W_1,W_1')U(W_2,W_2')U(W_3,W_3'')U(W_4,W_4'')]\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-\rho t_1t_2}+e^{-\rho t_1t_3}+e^{-\rho t_1t_4}+e^{-\rho t_2t_3}+e^{-\rho t_2t_4}+e^{-\rho t_3t_4}\\
&-e^{-\rho t_1t_2-\rho t_1t_3-\rho t_2t_3}-e^{-\rho t_1t_2-\rho t_1t_4-\rho t_2t_4}-e^{-\rho t_1t_3-\rho t_1t_4-\rho t_3t_4}-e^{-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4}\\
&+e^{-\rho t_1t_2-\rho t_1t_3-\rho t_1t_4-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4}-3)(e^{-\rho t_1t_2}-1)(e^{-\rho t_3t_4}-1)\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(3\rho^4t_1^2t_2^2t_3^2t_4^2+\text{Remainder terms})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
\le&C'|\rho|^4,\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}[U(W_1,W_1')U(W_1,W_1'')U(W_2,W_2')U(W_3,W_3'')]\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-t_1t_2}+e^{-\rho t_1t_3}+e^{-\rho t_1t_4}+e^{-\rho t_2t_3}+e^{-\rho t_2t_4}+e^{-\rho t_3t_4}\\
&-e^{-t_1t_2-\rho t_1t_3-\rho t_2t_3}-e^{-t_1t_2-\rho t_1t_4-\rho t_2t_4}-e^{-\rho t_1t_3-\rho t_1t_4-\rho t_3t_4}-e^{-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4}\\
&+e^{-t_1t_2-\rho t_1t_3-\rho t_1t_4-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4}-3)(e^{-\rho t_1t_3}-1)(e^{-\rho t_2t_4}-1)\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(2\rho^3t_1^2t_2^2t_3^2t_4^2+\text{Remainder terms})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
\le&C'|\rho|^3,
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[U(W_1,W_1')U(W_1,W_1'')U(W_2,W_2')U(W_2,W_2'')]\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-t_1t_2}+e^{-\rho t_1t_3}+e^{-\rho t_1t_4}+e^{-\rho t_2t_3}+e^{-\rho t_2t_4}+e^{-t_3t_4}\\
&-e^{-t_1t_2-\rho t_1t_3-\rho t_2t_3}-e^{-t_1t_2-\rho t_1t_4-\rho t_2t_4}-e^{-\rho t_1t_3-\rho t_1t_4-t_3t_4}-e^{-\rho t_2t_3-\rho t_2t_4-t_3t_4}\\
&+e^{-t_1t_2-\rho t_1t_3-\rho t_1t_4-\rho t_2t_3-\rho t_2t_4-t_3t_4}-3)(e^{-\rho t_1t_3}-1)(e^{-\rho t_2t_4}-1)\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(2\rho^2t_1^2t_2^2t_3^2t_4^2+\text{Remainder terms})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
\le&C'|\rho|^2,
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[U(W_1,W_1')U(W_2,W_2'')U(W_3''',W_3)U(W_4''',W_4'')]\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-\rho t_1t_2}-1)(e^{-\rho t_1t_3}-1)(e^{-\rho t_2t_4}-1)(e^{-\rho t_3t_4}-1)\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(\rho^4t_1^2t_2^2t_3^2t_4^2+\text{Remainder terms})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
\le&C'|\rho|^4,
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[U(W_1,W_1')U(W_1''',W_1')U(W_2,W_2'')U(W_3''',W_3'')]\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-t_1t_2}-1)(e^{-\rho t_1t_3}-1)(e^{-\rho t_2t_4}-1)(e^{-\rho t_3t_4}-1)\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
=&\int_{\mathbb{R}^4}\pi^{-4}e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(\rho^3t_1^2t_2^2t_3^2t_4^2+\text{Remainder terms})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2}\\
\le&C'|\rho|^3,
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[U(W_1, W_1')U(W_1''', W_1')U(W_2, W_2'')U(W_2''', W_2'')] \\
=&\int_{\mathbb{R}^4} \pi^{-4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-t_1t_2}-1)(e^{-\rho t_1 t_3}-1)(e^{-\rho t_2 t_4}-1)(e^{-t_3t_4}-1)\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2} \\
=&\int_{\mathbb{R}^4} \pi^{-4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(\rho^2 t_1^2t_2^2t_3^2t_4^2 + \text{Remainder terms})\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2} \\
\leq& C'|\rho|^2,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}[U(W_1, W_1')^2 U(W_2, W_2')U(W_3, W_3')] \\
=&\int_{\mathbb{R}^4} \pi^{-4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(e^{-t_1t_2}+e^{-\rho t_1t_3}+e^{-\rho t_1t_4}+e^{-\rho t_2t_3}+e^{-\rho t_2t_4}+e^{-\rho t_3t_4} \\
&-e^{-t_1t_2-\rho t_1t_3-\rho t_2t_3}-e^{-t_1t_2-\rho t_1t_4-\rho t_2t_4}-e^{-\rho t_1t_3-\rho t_1t_4-\rho t_3t_4}-e^{-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4} \\
&+e^{-t_1t_2-\rho t_1t_3-\rho t_1t_4-\rho t_2t_3-\rho t_2t_4-\rho t_3t_4}-3)^2\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2} \\
=&\int_{\mathbb{R}^4} \pi^{-4} e^{-(t_1^2+t_2^2+t_3^2+t_4^2)}(2\rho t_1t_2t_3t_4 + \text{Remainder terms})^2\frac{dt_1}{t_1^2}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3^2}\frac{dt_4}{t_4^2} \\
\leq& C'|\rho|^2.
\end{aligned}
$$

□

1.2.5 Proof of Proposition 5.2

Proof. By (2.11) of Székely et al. (2007), we have

$$
\int (e^{\imath\langle t_i, W^{(i)}\rangle}-1)(e^{-\imath\langle t_i, W^{(i)'}\rangle}-1)(c_{a,d_i}|t_i|^{1+a})^{-1}dt = \mathcal{K}(W^{(i)}, W^{(i)'}).
$$

Therefore, direct calculation shows that

$$
\begin{aligned}
&\int \left| \mathbb{E}\prod_{i=1}^p (e^{\imath\langle t_i, W^{(i)}\rangle}-1) - \prod_{i=1}^p \mathbb{E}(e^{\imath\langle t_i, W^{(i)}\rangle}-1)\right|^2 d\tilde{t} \\
=&\int \mathbb{E}\prod_{i=1}^p (e^{\imath\langle t_i, W_1^{(i)}\rangle}-1)(e^{-\imath\langle t_i, W_2^{(i)}\rangle}-1)d\tilde{t} + \int \mathbb{E}\prod_{i=1}^p (e^{\imath\langle t_i, W_{2i-1}^{(i)}\rangle}-1)(e^{-\imath\langle t_i, W_{2i}^{(i)}\rangle}-1)d\tilde{t} \\
&-2\int \mathbb{E}\prod_{i=1}^p (e^{\imath\langle t_i, W_1^{(i)}\rangle}-1)(e^{-\imath\langle t_i, W_{i+1}^{(i)}\rangle}-1)d\tilde{t} \\
=&\mathbb{E}\prod_{i=1}^p \mathcal{K}_i(W_1^{(i)}, W_2^{(i)}) + \prod_{i=1}^p \mathbb{E}\mathcal{K}_i(W_{2i-1}^{(i)}, W_{2i}^{(i)}) - 2\mathbb{E}\prod_{i=1}^p \mathcal{K}_i(W_1^{(i)}, W_{i+1}^{(i)}) \\
=&MdCov^2(W;a).
\end{aligned}
$$

□

2 Additional Simulation Results

2.1 Testing for mutual independence

In this section, we provide additional simulation examples to compare the power from our proposed test and LD_{t^*}. LD_{t^*} is studied in Leung & Drton (2017), which is based on the the sign covariance introduced by Bergsma & Dassios (2014). The sign covariance also targets at non-linear dependence as distance covariance. We consider several non-Gaussian data generating processes as follows. The power (rejection probabilities) reported below are based on 5000 Monte Carlo simulations at the nominal level $\alpha = 0.05$.

Example 2.1. *The data $W = (W_1, ..., W_p) \in \mathbb{R}^p$, where $W_i = Z_i^3$ for $i = 1, ..., p$ and $Z = (Z_1, ..., Z_p)$ are generated from multivariate t-distribution with degrees of freedom 5 and the following three covariance matrices $\mathbf{\Sigma} = (\sigma_{ij}(\rho))_{i,j=1}^p$ for $\rho = 0.1$.*

- **AR(1) structure:** *$\sigma_{ii} = 1$ and $\sigma_{ij} = \rho^{|i-j|}$ for all $i, j \in \{1, ..., d\}$;*
- **Band structure:** *$\sigma_{ii} = 1$ for $i = 1, ..., d$; $\sigma_{ij} = \rho$ if $0 < |i-j| < 3$ and $\sigma_{ij} = 0$ if $|i-j| \geq 3$;*
- **Block structure:** *Define $\mathbf{\Sigma}_{block} = (\sigma^*_{ij})$ with $\sigma_{ii} = 1$ and $\sigma_{ij} = \rho$ if $i \neq j$ for all $i, j \in \{1, ..., 5\}$. The covariance matrix is given by the following Kronecker product $\mathbf{\Sigma} = \boldsymbol{I}_{\lfloor p/5 \rfloor} \otimes \mathbf{\Sigma}_{block}$.*

Table 1: Power of Example 2.1

		AR(1)		Band		Block	
n	p	dCov	LD_{t^*}	dCov	LD_{t^*}	dCov	LD_{t^*}
30	25	0.699	0.307	0.723	0.365	0.718	0.377
30	50	0.918	0.609	0.924	0.620	0.920	0.642
30	100	0.992	0.920	0.995	0.932	0.995	0.942
30	200	1.000	1.000	1.000	0.999	1.000	1.000
30	400	1.000	1.000	1.000	1.000	1.000	1.000
60	25	0.959	0.357	0.971	0.482	0.974	0.492
60	50	0.999	0.665	0.999	0.746	1.000	0.752
60	100	1.000	0.961	1.000	0.974	1.000	0.980
60	200	1.000	1.000	1.000	1.000	1.000	1.000
60	400	1.000	1.000	1.000	1.000	1.000	1.000

Example 2.2. *Similar to the example used in Leung & Drton (2017), consider the data* $W = (W_1, ..., W_p) \in \mathbb{R}^p$*, generated from multivariate power exponential distribution with kurtosis parameter equals 20 and a compound symmetry covariance matrix* $\Sigma = (\sigma_{ij})_{i,j=1}^p$ *for* $\sigma_{ii} = 1$ *and* $\sigma_{ij} = 0.03$.

Table 2: Power of Example 2.2

	$n = 30$		$n = 60$	
p	dCov	LD_{t^*}	dCov	LD_{t^*}
25	0.108	0.092	0.148	0.116
50	0.167	0.136	0.311	0.209
100	0.293	0.199	0.619	0.421
200	0.560	0.424	0.915	0.788
400	0.859	0.740	0.995	0.979
800	0.974	0.942	1.000	1.000

From both examples we observed that our proposed test shows higher power than LD_{t^*} when the dimension and sample size are low. Notice that as dimension decreases, the overall dependence also decreases. Therefore, our test outperforms LD_{t^*} under weak signal situations and performs equally well as LD_{t^*} for strong signal cases.

2.2 Testing for banded dependence structure

In this subsection, we conduct additional simulations to evaluate the performance of the proposed test for the banded dependence structure. Adaptations of the CJ and HL_ρ tests to testing the banded dependence structure are also carried out to compare with the proposed test. The simulation setting is the same as in Section 6.1.

Example 2.3. *Consider the following banded dependence structures*

i) *The data is generated from multivariate normal distribution with banded covariance matrix* $\mathbf{\Sigma} = (\sigma_{ij})_{i,j=1}^p$*, where* $\sigma_{ii} = 1$ *for* $i = 1, ..., d$*,* $\sigma_{ij} = 0.3$ *if* $0 < |i - j| < 5$ *and* $\sigma_{ij} = 0$ *if* $|i - j| \geq 5$*;*

ii) *The data is generated as* $W = Z^3$*, where* Z *is generated from i);*

iii) *The data is generated as* $W = Z^{1/3}$*, where* Z *is generated from i).*

Table 3 shows the result from Example 2.3. The true bandwidth is 4 in this example, we choose $h = 5$ and $h = 10$ in the tests. It can be seen from the table that dCov-based banded dependence structure test has slight size inflation when $n = 60$, which subsides as sample size grows. In contrast, HL_τ test is a little bit conservative in some scenarios. CJ test is more conservative in cases i & iii and shows strong size distortion when the distribution is too far from Gaussian in case ii. It appears that there is no big difference between using $h = 5$ and $h = 10$ for all of the three tests. Likewise, we provide the histogram of the test statistics from 5000 Monte Carlo simulation and also the kernel density estimate using the Gaussian kernel with the comparison of standard normal density as the red dashed line in Figure 2.1 for the three cases in this example where $n = 100$, $p = 800$ and $h = 10$. It is shown that the normal approximation is quite close to the null distribution of the proposed test statistic in all the three cases. The plots for $h = 5$ are almost identical to those for $h = 10$ and therefore omitted.

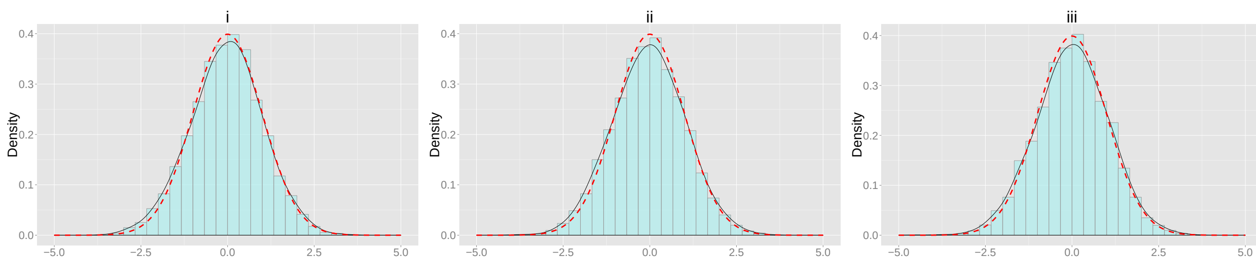

Figure 2.1: The histogram and kernel density estimate for the null distribution of the dCov-based test statistic for Example 2.3. The red dashed line is the density of the standard normal.

Example 2.4. *Consider the following cases*

i) *The data is generated from multivariate normal distribution with banded covariance matrix* $\mathbf{\Sigma} = (\sigma_{ij})_{i,j=1}^p$, *where* $\sigma_{ii} = 1$ *for* $i = 1, ..., d$, $\sigma_{ij} = 0.1$ *if* $0 < |i - j| \leq 20$ *and* $\sigma_{ij} = 0$ *if* $|i - j| > 20$;

ii) *The data is generated as* $W = Z^{1/3}$, *where* Z *is generated from i).*

iii) *The data is generated the same way as Example 6.3 in Section 6.1.*

Table 4 collects the results from Example 2.4. We choose $h = 5$ and $h = 10$ in all the tests whereas the true bandwidths in cases i and ii are both 20; in case iii, there is no banded dependence structure. We observe that the power for the proposed test is consistently higher than other methods and the power increases as sample size and dimension increase, whereas HL_τ test suffers from significant power reduction in all cases. Moreover, CJ test is the worst among these three tests with power less than the nominal level in most of the scenarios. This example demonstrates that our proposed banded dependence test has very good power performance.

Table 3: Size for the banded dependence tests from Example 2.3

			(i)			(ii)			(iii)		
	n	p	dCov	CJ	HL_τ	dCov	CJ	HL_τ	dCov	CJ	HL_τ
$h=5$	60	50	0.063	0.012	0.041	0.065	0.938	0.038	0.062	0.027	0.046
	60	100	0.067	0.008	0.048	0.070	1.000	0.044	0.064	0.020	0.044
	60	200	0.061	0.004	0.044	0.058	1.000	0.042	0.069	0.016	0.041
	60	400	0.060	0.002	0.043	0.054	1.000	0.037	0.060	0.013	0.042
	60	800	0.064	0.001	0.034	0.066	1.000	0.036	0.066	0.008	0.038
	100	50	0.056	0.021	0.040	0.059	0.943	0.040	0.060	0.029	0.042
	100	100	0.056	0.018	0.044	0.053	1.000	0.048	0.058	0.029	0.044
	100	200	0.055	0.015	0.042	0.052	1.000	0.042	0.053	0.026	0.046
	100	400	0.057	0.013	0.043	0.057	1.000	0.043	0.055	0.021	0.043
	100	800	0.059	0.005	0.041	0.058	1.000	0.039	0.056	0.017	0.040
$h=10$	60	50	0.069	0.009	0.034	0.063	0.893	0.031	0.060	0.022	0.036
	60	100	0.067	0.007	0.044	0.066	0.999	0.039	0.062	0.017	0.040
	60	200	0.062	0.004	0.041	0.062	1.000	0.040	0.073	0.015	0.039
	60	400	0.061	0.002	0.042	0.055	1.000	0.036	0.061	0.013	0.041
	60	800	0.064	0.001	0.033	0.064	1.000	0.036	0.065	0.008	0.038
	100	50	0.053	0.016	0.033	0.060	0.901	0.032	0.058	0.024	0.034
	100	100	0.058	0.016	0.040	0.058	1.000	0.043	0.056	0.025	0.040
	100	200	0.057	0.014	0.042	0.052	1.000	0.040	0.054	0.025	0.043
	100	400	0.057	0.012	0.042	0.060	1.000	0.042	0.054	0.021	0.042
	100	800	0.061	0.005	0.040	0.059	1.000	0.038	0.057	0.017	0.040

Table 4: Power for the banded dependence tests from Example 2.4

			(i)			(ii)			(iii)		
	n	p	dCov	CJ	HL_τ	dCov	CJ	HL_τ	dCov	CJ	HL_τ
$h=5$	60	50	0.983	0.070	0.185	0.938	0.091	0.192	1.000	0.019	0.311
	60	100	0.998	0.038	0.149	0.980	0.062	0.151	1.000	0.014	0.359
	60	200	0.999	0.020	0.116	0.993	0.037	0.115	1.000	0.010	0.407
	60	400	1.000	0.005	0.072	0.996	0.020	0.073	1.000	0.008	0.449
	60	800	1.000	0.002	0.064	0.998	0.012	0.051	1.000	0.003	0.511
	100	50	1.000	0.249	0.345	0.999	0.209	0.357	1.000	0.023	0.310
	100	100	1.000	0.170	0.302	1.000	0.163	0.285	1.000	0.023	0.368
	100	200	1.000	0.101	0.215	1.000	0.107	0.210	1.000	0.023	0.423
	100	400	1.000	0.050	0.150	1.000	0.067	0.154	1.000	0.015	0.490
	100	800	1.000	0.026	0.114	1.000	0.038	0.106	1.000	0.010	0.537
$h=10$	60	50	0.916	0.050	0.130	0.804	0.062	0.129	1.000	0.015	0.305
	60	100	0.957	0.028	0.114	0.877	0.045	0.116	1.000	0.013	0.356
	60	200	0.973	0.016	0.095	0.902	0.028	0.094	1.000	0.010	0.406
	60	400	0.980	0.004	0.063	0.915	0.017	0.061	1.000	0.007	0.449
	60	800	0.983	0.002	0.057	0.921	0.011	0.044	1.000	0.003	0.511
	100	50	0.998	0.171	0.244	0.988	0.147	0.250	1.000	0.018	0.304
	100	100	1.000	0.122	0.223	0.997	0.123	0.212	1.000	0.020	0.365
	100	200	1.000	0.073	0.159	1.000	0.080	0.157	1.000	0.021	0.420
	100	400	1.000	0.036	0.119	1.000	0.051	0.116	1.000	0.014	0.490
	100	800	1.000	0.020	0.095	1.000	0.033	0.086	1.000	0.009	0.536

2.3 Computational complexity

In the high-dimensional setting, computation cost is a worthy consideration. In this section, we compare the computational complexity for different methods theoretically and also provide a runtime analysis. The discussion only focuses on the $\mathcal{L}_2$ methods mentioned in the paper and also dHSIC, because the $\mathcal{L}_\infty$ methods have the same order of computational complexity as their $\mathcal{L}_2$ counterparts.

SC's test uses the Pearson correlation, which is very straightforward to implement in $O(n)$ operations; LD_τ and LD_ρ use the rank correlation coefficients Kendall's τ and Spearman's ρ, which are U-statistics of degrees 2 and 3, respectively. Naive implementation involves $O(n^2)$ and $O(n^3)$ operations. However, Spearman's ρ statistics can be easily calculated in $O(n \log n)$ operations based on its alternative definition. Christensen (2005) showed that Kendall's τ can also be computed in $O(n \log n)$. Bergsma-Dassios' sign covariance t^* is a U-statistics of degree 4. Direct computing has a $O(n^4)$ complexity, but Weihs et al. (2016) showed that it can be computed in $O(n^2 \log n)$ operations; Heller & Heller (2016) further improved it to $O(n^2)$. Distance covariance can also be computed in $O(n^2)$ operations; Huo & Székely (2016) proposed a fast computing algorithm which only requires $O(n \log n)$ operations. For the corresponding $\mathcal{L}_2/\mathcal{L}_\infty$ statistics for testing pairwise independence, we need to evaluate the underlying sample dependence measures $\binom{p}{2}$ times. Since all the existing $\mathcal{L}_2$ statistics are asymptotically pivotal, no further calibration is needed for these tests. However, in our proposed test statistic, we do need to estimate the variance part. Hence the runtime for our test is slightly longer than other methods with the same order of computational complexity.

On the other hand, dHSIC itself can be computed in $O(p^2 n^2)$ operations. Pfister et al. (2016) proposed the dHSIC independence test based on permutation test, bootstrap and Gamma approximation. If the first two approaches are used, the overall complexity becomes $O(Bp^2 n^2)$, where B is the number of permutation/bootstrap. This is quite demanding as compared with other methods discussed above. Table 5 shows the summary of computational complexity for different methods. Figure 2.2 presents the runtime (at log scale) results, which is consistent with the theory.

References

Bergsma, W. & Dassios, A. (2014), 'A consistent test of independence based on a sign covariance related to kendalls tau', *Bernoulli* **20**(2), 1006–1028.

Table 5: Computational Complexity for All Tests

	Naive Algo	Fast Algo
dCov	$O(p^2n^2)$	$O(p^2n \log n)$
SC	$O(p^2n)$	
LD_τ	$O(p^2n^2)$	$O(p^2n \log n)$
LD_ρ	$O(p^2n^3)$	$O(p^2n \log n)$
LD_{t^*}	$O(p^2n^4)$	$O(p^2n^2 \log n)$/ $O(p^2n^2)$
dHSIC	$O(p^2n^2)$	

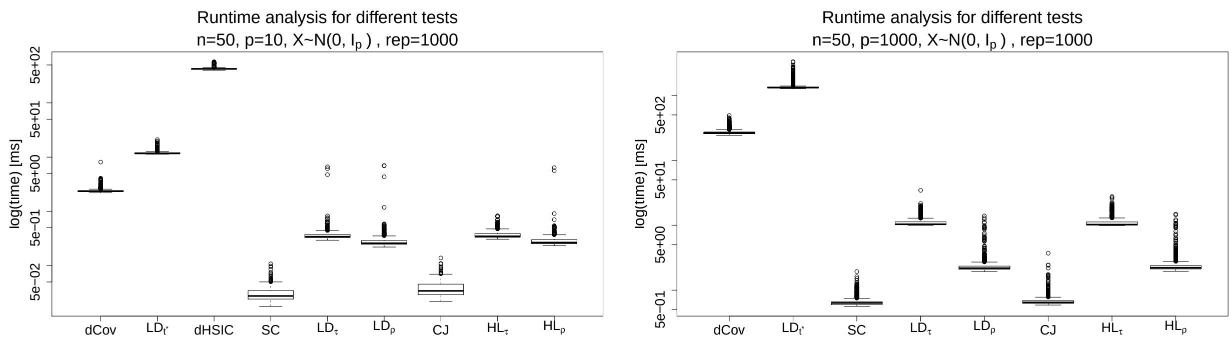

Figure 2.2: Runtime analysis. dCov is implemented using the naive algorithm for simplicity, dHSIC is implemented using permutation with $B = 200$, other tests are implemented using the fast algorithms.

Christensen, D. (2005), 'Fast algorithms for the calculation of kendalls τ', *Computational Statistics* **20**(1), 51–62.

Hall, P. & Heyde, C. C. (1980), *Martingale limit theory and its application*, Academic press.

Heller, Y. & Heller, R. (2016), 'Computing the bergsma dassios sign-covariance', *arXiv preprint arXiv:1605.08732* .

Huo, X. & Székely, G. J. (2016), 'Fast computing for distance covariance', *Technometrics* **58**(4), 435–447.

Leung, D. & Drton, M. (2017), 'Testing mutual independence in high dimensions with sums of squares of rank correlations', *The Annals of Statistics, to appear, arXiv:1501.01732* .

Park, T., Shao, X. & Yao, S. (2015), 'Partial martingale difference correlation', *Electronic Journal of Statistics* **9**, 1492–1517.

Pfister, N., Bühlmann, P., Schölkopf, B. & Peters, J. (2016), 'Kernel-based tests for joint independence', *Journal of the Royal Statistical Society, Series B, to appear, arXiv:1603.00285* .

Serfling, R. J. (1980), *Approximation theorems of mathematical statistics*, Vol. 162, John Wiley & Sons.

Székely, G. J., Rizzo, M. L. & Bakirov, N. K. (2007), 'Measuring and testing dependence by correlation of distances', *The Annals of Statistics* **35**(6), 2769–2794.

Weihs, L., Drton, M. & Leung, D. (2016), 'Efficient computation of the bergsma–dassios sign covariance', *Computational Statistics* **31**(1), 315–328.