Biometrika (2015), **xx**, x, *pp*. 1–14

Printed in Great Britain

Supplementary Material for "Diagnostic studies in sufficient dimension reduction"

BY XIN CHEN
Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546
stacx@nus.edu.sg

R. DENNIS COOK
School of Statistics, University of Minnesota, Minneapolis, Minnesota 55455, U.S.A.
dennis@stat.umn.edu

AND CHANGLIANG ZOU
Institute of Statistics, Nankai University, Tianjin 300071, China
nk.chlzou@gmail.com

In this file, we present the proof of Theorems 1-3.

APPENDIX A: PROOF OF THEOREM 1

We first give a sketch of the proof of Theorem 1. $V_n(A;B)$ is a V-statistic. Denote the corresponding U-type statistic of $V_n(A;B)$ as $U_n(A;B)$. We derive the asymptotic distribution of $U_n(A;B)$ in Lemma 2. The proof of Lemma 2 requires Lemma 1 which establishes the joint distribution of the components in $U_n(A;B)$. Lemma 3 quantifies the asymptotic bias of $V_n(A;B)$ with respect to $U_n(A;B)$ from which we can prove Theorem 1.

Denote $U_n(A;B) = \tilde{S}_{n,1}(A;B) + \tilde{S}_{n,2}(A;B) - 2\tilde{S}_{n,3}(A;B)$, where

$$
\begin{aligned}
\tilde{S}_{n,1}(A;B) =& \frac{1}{P_n^2 \hat{f}_0^2(B)} \sum_{k\neq l} K_h(A^{\mathsf{T}}X_k - B)K_h(A^{\mathsf{T}}X_l - B)\|X_k - X_l\||Y_k - Y_l|,\\
\tilde{S}_{n,2}(A;B) =& \frac{1}{P_n^2 \hat{f}_0^2(B)} \sum_{k\neq l} K_h(A^{\mathsf{T}}X_k - B)K_h(A^{\mathsf{T}}X_l - B)\|X_k - X_l\|\\
& \times \frac{1}{P_n^2 \hat{f}_0^2(B)} \sum_{k\neq l} K_h(A^{\mathsf{T}}X_k - B)K_h(A^{\mathsf{T}}X_l - B)|Y_k - Y_l| \equiv \tilde{S}_{n,2x}(A;B)\tilde{S}_{n,2y}(A;B),\\
\tilde{S}_{n,3}(A;B) =& \frac{1}{P_n^3 \hat{f}_0^3(B)} \sum_{k\neq l\neq m} \frac{1}{6} \sum_{(i_1,i_2,i_3)\in(k,l,m)} K_h(A^{\mathsf{T}}X_{i_1} - B)K_h(A^{\mathsf{T}}X_{i_2} - B)K_h(A^{\mathsf{T}}X_{i_3} - B)\\
& \|X_{i_1} - X_{i_2}\||Y_{i_1} - Y_{i_3}|,
\end{aligned}
$$

where $P_n^d = n!/(n-d)!$.

Furthermore, let $\hat{S}_{n,i}(A;B)$ be the same form of $\tilde{S}_{n,i}(A,B)$, replacing $\hat{f}_0(B)$ with $f_0(B)$ and accordingly

$$U_n(A;B) = \frac{1}{\eta_n^2(B)}\left\{\hat{S}_{n,1}(A;B) + \frac{\hat{S}_{n,2x}(A;B)\hat{S}_{n,2y}(A;B)}{\eta_n^2(B)} - 2\frac{\hat{S}_{n,3}(A;B)}{\eta_n(B)}\right\},$$

where we denote $\eta_n(B) = \hat{f}_0(B)/f_0(B)$. It is well known that $\eta_n(B) \to 1$ in probability under the conditions.

The following lemma establishes the joint distribution of $\hat{S}_{n,i}(A;B)$ and $\eta_n(B)$, which allows us to further derive the distribution of $U_n(B)$.

LEMMA 1. *Under Conditions 1–5,*

$$\left\{\frac{nh^d f_0(B)}{D}\right\}^{1/2}\left[\left\{\hat{S}_{n,1}(A;B), \hat{S}_{n,2x}(A;B), \hat{S}_{n,2y}(A;B), \hat{S}_{n,3}(A;B), \eta_n(B)\right\} - \mu\right] \to N_5(0,\Sigma) \text{ in distribution},$$

where $\mu = \{E(\|P_1-P_2\||Q_1-Q_2|), \mu_x, \mu_y, E(\|P_1-P_2\||Q_1-Q_3|), 1\}^{\mathsf{T}}$, $\Sigma = (v_{ij})_{5\times 5}$ *with*

$$\begin{aligned}
v_{11} =& 4E\left\{E^2(\|X-P\||Y-Q| \mid Z)\right\},\ v_{21} = 4E\left\{E(\|X-P\||Y-Q| \mid Z)E(\|X-P\| \mid Z)\right\},\\
v_{31} =& 4E\left\{E(\|X-P\||Y-Q| \mid Z)E(\|Y-Q\| \mid Z)\right\},\\
v_{41} =& 2E\left\{E(\|X-P\||Y-Q| \mid Z)E(\|X-P_1\||Y-Q_2| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P\||Y-Q| \mid Z)E(\|X-P_1\||Q_1-Q_2| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P\||Y-Q| \mid Z)E(\|P_1-P_2\||Y-Q_1| \mid Z)\right\},\\
v_{51} =& 2E(\|P_1-P_2\||Q_1-Q_2|),\\
v_{22} =& 4E\left\{E^2(\|X-P\| \mid Z)\right\},\ v_{32} = 4E\left\{E(\|X-P\| \mid Z)E(|Y-Q\| \mid Z)\right\},\\
v_{42} =& 2E\left\{E(\|X-P_1\||Y-Q_2| \mid Z)E(\|X-P\| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P\| \mid Z)E(\|X-P_1\||Q_1-Q_2| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P\| \mid Z)E(\|P_1-P_2\||Y-Q_1| \mid Z)\right\},\\
v_{52} =& 2E(\|P_1-P_2\|),\ v_{33} = 4E\left\{E^2(|Y-Q| \mid Z)\right\},\\
v_{43} =& 2E\left\{E(\|X-P_1\||Y-Q_2| \mid Z)E(|Y-Q\| \mid Z)\right\}\\
&+ 2E\left\{E(|Y-Q\| \mid Z)E(\|X-P_1\||Q_1-Q_2| \mid Z)\right\}\\
&+ 2E\left\{E(|Y-Q| \mid Z)E(\|P_1-P_2\||Y-Q_1| \mid Z)\right\},\\
v_{53} =& 2E(|Q_1-Q_2|),\\
v_{44} =& E\left\{E^2(\|X-P_1\||Y-Q_2| \mid Z)\right\} + E\left\{E^2(\|X-P_1\||Q_1-Q_2| \mid Z)\right\}\\
&+ E\left\{E^2(\|P_1-P_2\||Y-Q_1| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P_1\||Y-Q_2| \mid Z)E(\|X-P_1\||Q_1-Q_2| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P_1\||Y-Q_2| \mid Z)E(\|P_1-P_2\||Y-Q_1| \mid Z)\right\}\\
&+ 2E\left\{E(\|X-P_1\||Q_1-Q_2| \mid Z)E(\|P_1-P_2\||Y-Q_1| \mid Z)\right\},\\
v_{45} =& E\left\{E(\|X-P_1\||Y-Q_2| \mid Z)\right\} + E\left\{E(\|X-P_1\||Q_1-Q_2| \mid Z)\right\}\\
&+ E\left\{E(\|P_1-P_2\||Y-Q_1| \mid Z)\right\},\ v_{55} = 1.
\end{aligned}$$

Proof. Firstly, $\{\hat{S}_{n,1}(A;B), \hat{S}_{n,2x}(A;B), \hat{S}_{n,2y}(A;B), \hat{S}_{n,3}(A;B), \eta_n(B)\}^{\mathsf{T}}$ can be viewed as a five-dimensional U-statistic of order three. Thus, by the standard central limit theory of multivariate U-

statistic, it remains to evaluate its asymptotic expectation and covariance. Denote $Z_k = (X_k, Y_k)$. Note

$$\begin{aligned}
E\{\hat{S}_{n,1}(A;B)\} =& f_0^{-2}(B) E\{K_h(A^\intercal X_k - B) K_h(A^\intercal X_l - B) \|X_k - X_l\| |Y_k - Y_l|\} \\
=& f_0^{-2}(B) E[E\{K_h(A^\intercal X_k - B) K_h(A^\intercal X_l - B) \|X_k - X_l\| |Y_k - Y_l| \mid Z_k, A^\intercal X_l = t\}] \\
=& f_0^{-2}(B) E\{K_h(A^\intercal X_k - B)\} \int K_h(t - B) \|X_k - u\| |Y_k - v| f_{A^\intercal X}^{(X,Y)}(u, v \mid t) f_0(t) dt du dv \\
\equiv& f_0^{-2}(B) E\{K_h(A^\intercal X_k - B)\} \Delta_1.
\end{aligned}$$

By Condition 5 and Taylor expansion,

$$\begin{aligned}
\Delta_1 =& \frac{1}{h^d} \int_{u,v} \|X_k - u\| |Y_k - v| \int_w K(w) f_1(u, v, B + hw) d(hw) du dv \\
\leq& \int_{u,v} \|X_k - u\| |Y_k - v| \left\{ f_1(u, v, B) + C_0 h^q \right\} du dv \\
\leq& f_0(B) E(\|X_k - P\| |Y_k - Q| \mid Z_k) + C_0 h^q \int_{u,v} \|X_k - u\| |Y_k - v| du dv,
\end{aligned}$$

where $(P, Q) \sim f_{A^\intercal X}^{(X,Y)}(\cdot, \cdot \mid B)$. Thus,

$$\begin{aligned}
& E\{\hat{S}_{n,1}(A;B)\} - E(\|P_1 - P_2\| |Q_1 - Q_2|) \\
\leq& f_0^{-1}(B) E\{K_h(A^\intercal X_k - B)\} E(\|X_k - P\| |Y_k - Q| \mid Z_k) - E(\|P_1 - P_2\| |Q_1 - Q_2|) \\
& + C_0^2 f_0^{-1}(B) h^q \int E\{\|u - P\| |v - Q| \mid (u, v)\} du dv \\
=& O(h^q),
\end{aligned}$$

where we use similar techniques to those in calculating Δ_1. Similarly, we can obtain

$$\begin{aligned}
E\{\hat{S}_{n,2x}(A;B)\} =& \mu_x + O(h^q), \\
E\{\hat{S}_{n,2y}(A;B)\} =& \mu_y + O(h^q), \\
E\{\hat{S}_{n,3}(A;B)\} =& E(\|P_1 - P_2\| |Q_1 - Q_3|) + O(h^q),
\end{aligned}$$

from which we finish the calculation of the asymptotic expectations.

Next, we handle the asymptotic variances. Take $\hat{S}_{n,1}(A;B)$ for example and the calculation for the other terms is similar. By using the standard result of U-statistics, we know

$$\operatorname{var}\{\hat{S}_{n,1}(A,B)\} = \frac{4}{n} \operatorname{var}[E\{H_n(A;B) \mid Z_k\}] + O\{(nh^d)^{-2}\},$$

where $H_n(A;B)$ denotes the kernel function of U-statistic $\hat{S}_{n,1}(A;B)$. It can be verified that

$$E\{H_n(A;B) \mid Z_k\} = K_h(A^\intercal X_k - B) \left\{ f_0^{-1}(B) E(\|X_k - P\| |Y_k - Q| \mid Z_k) + O(h^q) \right\}.$$

Consequently,

$$\begin{aligned}
\operatorname{var}\{\hat{S}_{n,1}(A;B)\} =& \frac{4}{n f_0^2(B)} E\left\{ K_h^2(A^\intercal X_k - B) E^2(\|X_k - P\| |Y_k - Q| \mid Z_k) \right\} \{1 + o(1)\} \\
=& \frac{4D}{nh^d f_0(B)} E\left\{ E^2(\|X - P\| |Y - Q| \mid Z) \right\} \{1 + o(1)\}. \qquad \text{(A.1)}
\end{aligned}$$

Similarly, we can compute the asymptotic covariance terms in the following way

$$
\begin{aligned}
&\operatorname{cov}\{\hat{S}_{n,1}(A;B), \hat{S}_{n,2x}(A;B)\} \\
&\quad = \frac{4}{n}\operatorname{cov}[E\{H_{n,1}(A,B) \mid Z_k\}, E\{H_{n,2x}(A,B) \mid Z_k\}] + O\{(nh^d)^{-2}\} \\
&\quad = \frac{4}{n}E[E\{H_{n,1}(A,B) \mid Z_k\}E\{H_{n,2x}(A,B) \mid Z_k\}] + o\{(nh^d)^{-1}\} \\
&\quad = \frac{4}{nf_0^2(B)}E\left\{K_h^2(A^{\mathsf{T}}X_k - B)E(\|X_k - P\||Y_k - Q| \mid Z_k)E(\|X_k - P\| \mid Z_k)\right\} + o\{(nh^d)^{-1}\} \\
&\quad = \frac{4}{nh^d f_0(B)}E\left\{K_h^2(A^{\mathsf{T}}X_k - B)E(\|X_k - P\||Y_k - Q| \mid Z_k)E(\|X_k - P\| \mid Z_k)\right\} + o\{(nh^d)^{-1}\} \\
&\quad = \frac{4D}{nh^d f_0(B)}E\left\{E(\|X - P\||Y - Q| \mid Z)E(\|X - P\| \mid Z)\right\} + o\{(nh^d)^{-1}\},
\end{aligned}
$$

from which we can complete the proof of this lemma. ■

LEMMA 2. *Under H_0 and Conditions 1–4,*

$$\frac{nh^d f_0(B)}{D}U_n(A;B) \to \mu_x\mu_y\mathcal{N}_4^2 + 2\mathcal{N}_4(\mathcal{N}_3 - \mu_x\mathcal{N}_2 - \mu_y\mathcal{N}_1) + \mathcal{N}_1\mathcal{N}_2 \ \textit{in distribution},$$

where $(\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3, \mathcal{N}_4)$ are normally distributed with mean zero and covariance matrix V given by

$$V = \begin{pmatrix} 4\nu_x & 4\mu_x\mu_y & 4\nu_x\mu_y + 2\mu_x^2\mu_y & 2\mu_x \\ 4\mu_x\mu_y & 4\nu_y & 4\nu_y\mu_x + 2\mu_y^2\mu_x & 2\mu_y \\ 4\nu_x\mu_y + 2\mu_x^2\mu_y & 4\nu_y\mu_x + 2\mu_y^2\mu_x & \lambda & 3\mu_x\mu_y \\ 2\mu_x & 2\mu_y & 3\mu_x\mu_y & 1 \end{pmatrix}.$$

Proof. Under the null hypothesis,

$$\mu = (\mu_x\mu_y, \mu_x, \mu_y, \mu_x\mu_y, 1)^{\mathsf{T}}\{1 + O(h^q)\},$$

and Σ becomes

$$\begin{pmatrix} 4\nu_x\nu_y & 4\nu_x\mu_y & 4\nu_y\mu_x & 2\nu_x\nu_y + 2\nu_x\mu_y^2 + 2\nu_y\mu_x^2 & 2\mu_x\mu_y \\ 4\nu_x\mu_y & 4\nu_x & 4\mu_x\mu_y & 4\nu_x\mu_y + 2\mu_x^2\mu_y & 2\mu_x \\ 4\nu_y\mu_x & 4\mu_x\mu_y & 4\nu_y & 4\nu_y\mu_x + 2\mu_y^2\mu_x & 2\mu_y \\ 2\nu_x\nu_y + 2\nu_x\mu_y^2 + 2\nu_y\mu_x^2 & 4\nu_x\mu_y + 2\mu_x^2\mu_y & 4\nu_y\mu_x + 2\mu_y^2\mu_x & v_{44} & 3\mu_x\mu_y \\ 2\mu_x\mu_y & 2\mu_x & 2\mu_y & 3\mu_x\mu_y & 1 \end{pmatrix}$$

where v_{44} reduces to $\nu_x\nu_y + \nu_x\mu_y^2 + \nu_y\mu_x^2 + 2\nu_x\mu_y^2 + 2\nu_y\mu_x^2 + 2\mu_x^2\mu_y^2$. It can be easily seen that

$$(1, \mu_y, \mu_x, -2, 0)^{\mathsf{T}}\Sigma(1, \mu_y, \mu_x, -2, 0) = 0,$$

where $(1, \mu_y, \mu_x, -2, 0)$ is the gradient of the function $g(u_1, u_2, u_3, u_4, u_5) = (u_1 + u_2u_3/u_5^2 - 2u_4/u_5)$ evaluated at $(\mu_x\mu_y, \mu_x, \mu_y, \mu_x\mu_y, 1)$. Thus, the first-order multivariate Delta-theorem is not valid. We consider to apply a second-order multivariate Delta-theorem.

Let $\gamma_n \equiv \{\hat{S}_{n,1}(A;B), \hat{S}_{n,2x}(A;B), \hat{S}_{n,2y}(A;B), \hat{S}_{n,3}(A;B), \eta_n(B)\}^\top$. First, obtain the Hessian matrix of $g(u_1, u_2, u_3, u_4, u_5)$ evaluated at $(\mu_x\mu_y, \mu_x, \mu_y, \mu_x\mu_y, 1)$

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -2\mu_y \\ 0 & 1 & 0 & 0 & -2\mu_x \\ 0 & 0 & 0 & 0 & 2 \\ 0 & -2\mu_y & -2\mu_x & 2 & 2\mu_x\mu_y \end{pmatrix}.$$

As a consequence,

$$nh^d U_n(A;B) = \frac{nh^d}{2}(\gamma_n - \mu)^\top H(\gamma_n - \mu) + o_p(1),$$

where we use the fact that $\eta_n(B) \to 1$ in probability, the condition $nh^{d/2+2q} \to 0$ and the Slutsky theorem. By Lemma 1, we can obtain the assertion by some calculation immediately. ■

The following lemma characterizes the difference between $U_n(A;B)$ and $V_n(A;B)$.

LEMMA 3. *Under Conditions 1–4,*

$$nh^d\{V_n(A;B) - U_n(A;B)\} \to -\frac{2D}{f_0(B)}E(\|P_1 - P_2\||Q_1 - Q_2|), \quad \textit{in probability}.$$

Proof. Simple algebra leads to

$$\begin{aligned} S_{n,1} &= (1 - n^{-1})\tilde{S}_{n,1}(A;B), \\ S_{n,2} &= (1 - n^{-1})^2\tilde{S}_{n,2x}(A;B)\tilde{S}_{n,2y}(A;B), \\ S_{n,3} &= \{1 - 3n^{-1} + O(n^{-2})\}\tilde{S}_{n,3}(A;B) + \frac{1}{nh^d\hat{f}_0(B)}\acute{S}_{n,1}(A;B), \end{aligned}$$

where $\acute{S}_{n,1}(A;B) = h^d\{n\hat{f}_0(B)\}^{-2}\sum_{k\neq l}K_h^2(A^\top X_k - B)K_h(A^\top X_l - B)\|X_k - X_l\||Y_k - Y_l|$. Similar to the proof of Lemma 1, it can be verified that $\acute{S}_{n,1}(A;B) \to DE(\|P_1 - P_2\||Q_1 - Q_2|)$ in probability. By using the fact that $\tilde{S}_{n,i}(A;B) \to S_i(A;B)$ $(i = 1, 2, 3)$ in probability, which can be seen from Lemma 5, we can obtain the assertion immediately. ■

Proof of Theorem 1 Combining Lemmas 2–3, the assertion follows immediately. ■

APPENDIX B: PROOF OF THEOREMS 2–3

To prove Theorems 2–3, we need the following lemma from Hall (1984). Define $G_n(x,y) = E\{H_n(X_1,x)H_n(X_1,y)\}$.

LEMMA 4. *Assume H_n is symmetric, $E\{H_n(X_1, X_2 \mid X_1)\} = 0$ almost surely and $E\{H_n^2(X_1, X_2)\} < \infty$ for each n. If*

$$[E\{G_n^2(X_1,X_2)\} + n^{-1}E\{H_n^4(X_1,X_2)\}]/[E\{H_n^2(X_1,X_2)\}]^2 \to 0, \tag{A.2}$$

as $n \to \infty$, then $U_n \equiv \sum_{1<i<j<n} H_n(X_i, X_j)$ is asymptotically normal with zero mean and variance given by $C_n^2E\{H_n^2(X_1,X_2)\}$.

The following lemma establishes the uniform convergence of $\hat{S}_{n,i}(A;B)$ $(i = 1, 2, 3)$ which also plays important role in the proof of Theorems 2–3.

LEMMA 5. *Suppose Conditions 1–4 hold. Then,*

$$\sup_{B\in\mathbb{R}^d} |\hat{f}_0(B) - f_0(B)| = O\{h^q + (nh^d)^{-1/2}\log n\}, \text{ almost surely},$$
$$\sup_{B\in\mathbb{R}^d} |\hat{S}_{n,i}(A;B) - S_i(A;B)| = O\{h^q + (nh^d)^{-1/2}\log n\}\ (i=1,2,3), \text{almost surely}.$$

Proof. The first part is a well-known result concerning the uniform convergence rate of kernel density estimators; see Silverman (1978) or Gine and Guillou (2002). We present the proof only for the case of $\hat{S}_{n,1}(A,B)$ since the proofs for the other two cases are very similar.

By similar arguments used in (A.1) and Condition 5, we can verify that

$$\begin{aligned}
&\frac{1}{h^d}\sup_{B\in\mathbb{R}^d}\left[E\left\{K_h(A^{\mathsf T}X_k - B)K_h(A^{\mathsf T}X_l - B)\|X_k - X_l\||Y_k - Y_l|\right\}^2\right]\\
=&D\sup_{B\in\mathbb{R}^d}\left[f_0(B)E\left\{E^2(\|X-P\||Y-Q|\mid Z)\right\}\right]\{1+O(h^q)\}.
\end{aligned}$$

Therefore, by invoking similar arguments to those of theorem 37 in Pollard (1984), page 34,

$$\sup_{B\in\mathbb{R}^d} |\hat{S}_{n,1}(A;B) - E\{\hat{S}_{n,1}(A;B)\}| = O\{(nh^d)^{-1/2}\log n\}.$$

On the other hand, expanding $E\{\hat{S}_{n,1}(A,B)\}$ in a Taylor series with Lagrange form of the remainder term and using conditions, we have that

$$\sup_{B\in\mathbb{R}^d} |E\{\hat{S}_{n,1}(A;B)\} - S_1(A;B)| = O(h^q),$$

which completes the proof. ■

By this lemma, Proposition 2 holds immediately.

Proof of Theorem 2 $U_n(A;B)$ can be rewritten as

$$\begin{aligned}
U_n(A;B) &= \frac{1}{\eta_n^4(B)}\left\{\eta_n^2(B)\hat{S}_{n,1}(A;B) + \hat{S}_{n,2x}(A;B)\hat{S}_{n,2y}(A;B) - 2\eta_n(B)\hat{S}_{n,3}(A;B)\right\}\\
&\equiv \frac{1}{\eta_n^4(B)}\tilde{U}_n(A;B).
\end{aligned}$$

First of all, we observe that

$$\begin{aligned}
\eta_n^2(B)\hat{S}_{n,1}(A;B) =& \frac{1}{p_n^4}\sum_{j\neq k\neq l\neq m} L_1 \prod_{r=j,k,l,m} K_h(A^\top X_r - B) \\
&+ \frac{1}{nh^d f_0(B)}\left\{\frac{h^d}{nf_0(B)}\sum_i K_h^2(A^\top X_i - B)\hat{S}_{n,1} + 4\acute{S}_{n,31}(A,B)\right\} \\
&+ O_p\{n^{-1} + (nh^d)^{-2}\}, \\
\hat{S}_{n,2x}(A;B)\hat{S}_{n,2y}(A;B) =& \frac{1}{p_n^4}\sum_{j\neq k\neq l\neq m} L_2 \prod_{r=j,k,l,m} K_h(A^\top X_r - B) + \frac{4\acute{S}_{n,32}(A,B)}{nh^d f_0(B)} \\
&+ O_p\{n^{-1} + (nh^d)^{-2}\}, \\
\eta_n(B)\hat{S}_{n,3}(A;B) =& \frac{1}{p_n^4}\sum_{j\neq k\neq l\neq m} L_3 \prod_{r=j,k,l,m} K_h(A^\top X_r - B) \\
&+ \frac{1}{n^4 f_0^4(B)}\sum_{k\neq l\neq m}\sum_{r\in(k,l,m)} K_h(A^\top X_r - B)\|X_k - X_l\||Y_k - Y_m| + O_p(n^{-1})
\end{aligned}$$

uniformly in B, where

$$\begin{aligned}
\acute{S}_{n,31}(A;B) &= \frac{h^d}{n^3 f_0^3(B)}\sum_{k\neq l\neq m} K_h^2(A^\top X_k - B)K_h(A^\top X_l - B)K_h(A^\top X_m - B)\|X_k - X_l\||Y_k - Y_l|, \\
\acute{S}_{n,32}(A;B) &= \frac{h^d}{n^3 f_0^3(B)}\sum_{k\neq l\neq m} K_h^2(A^\top X_k - B)K_h(A^\top X_l - B)K_h(A^\top X_m - B)\|X_k - X_l\||Y_k - Y_m|,
\end{aligned}$$

and

$$\begin{aligned}
L_1(Z_j, Z_k, Z_l, Z_m) &= \frac{1}{6}\sum_{(i_1,i_2)\in(j,k,l,m)} \|X_{i_1} - X_{i_2}\||Y_{i_1} - Y_{i_2}|, \\
L_2(Z_j, Z_k, Z_l, Z_m) &= \frac{1}{6}\sum_{(i_1,i_2,i_3,i_4)\in(j,k,l,m)} \|X_{i_1} - X_{i_2}\||Y_{i_3} - Y_{i_4}|, \\
L_3(Z_j, Z_k, Z_l, Z_m) &= \frac{1}{24}\sum_{(i_1,i_2,i_3)\in(j,k,l,m)} \|X_{i_1} - X_{i_2}\||Y_{i_1} - Y_{i_3}|.
\end{aligned}$$

By similar arguments in the proof of Lemma 1, it can be verified that under H_0, all the $h^d\{nf_0(B)\}^{-1}\sum_i K_h^2(A^\top X_i - B)\hat{S}_{n,1}$, $\acute{S}_{n,31}(A,B)$ and $\acute{S}_{n,32}(A,B)$ uniformly converge to $D\mu_x\mu_y$. Additionally, we can have

$$\frac{h^d}{n^3 f_0^3(B)}\sum_{k\neq l\neq m}\sum_{r\in(k,l,m)} K_h(A^\top X_r - B)\|X_k - X_l\||Y_k - Y_m| \to 3D\mu_x\mu_y.$$

uniformly in B.

As a consequence, $\tilde{U}_n(A;B)$ can be further decomposed as

$$\tilde{U}_n(A;B) = \tilde{T}_n(A;B) + \frac{3}{nh^d f_0(B)}D\mu_x\mu_y + O_p\{h^q + (nh^d)^{-1/2}\log n\},$$

and furthermore by Lemma 3

$$V_n(A;B) = \tilde{T}_n(A;B) + \frac{1}{nh^d f_0(B)} D\mu_x\mu_y + O_p\{h^q + (nh^d)^{-1/2}\log n\},$$

uniformly in B. Let $\zeta_0(B) = f_0(B)/S_2(A;B)$. It remains mainly to derive the asymptotic distribution of $\tilde{T}_n(A) = \int \tilde{T}_n(A;B)\zeta_0(B)dB$ which can be re-written as a U-statistic of order four with the kernel function $\gamma_h(Z_1,\ldots,Z_4)$

$$\begin{aligned}\tilde{T}_n(A) &= \frac{1}{p_n^4}\sum_{j\neq k\neq l\neq m}(L_1+L_2-2L_3)\int \prod_{r=j,k,l,m} K_h(A^{\mathsf{T}}X_r - B)\zeta_0(B)dB \\ &\equiv \frac{1}{p_n^4}\sum_{j\neq k\neq l\neq m}\gamma_h(Z_j,Z_k,Z_l,Z_m).\end{aligned}$$

Next, we will show that $\tilde{T}_n(A)$ is asymptotically normal by using Lemma 4. First of all, $\gamma_h(Z_1,\ldots,Z_4)$ is symmetric, $\theta_n \equiv E\{\gamma_h(Z_1,\ldots,Z_r)\} = O\{(nh^d)^{-1}\}$ (by Lemma 2 and Delta theorem) and $E\{\nu_n(Z_1,\ldots,Z_r \mid Z_1)\} = O_p(h^q)$. Say, $\tilde{T}_n(A)$ is a degenerate U-statistic. The limit distribution of a degenerate U-statistics when its kernel function is fixed is a linear combination of independent, centered χ_1^2 distributions, and cannot be derived using classical martingale methods. However, in certain cases in which the kernel function of the U-statistic depends on n, a normal distribution can result (c.f., Hall 1984).

Let us define the projection of $\tilde{T}_n(A)$ to be $W_n(A)$ given by

$$W_n(A) = \frac{6}{\binom{n}{2}}\sum_{1\leq j<k\leq n}\{\gamma_{h2}(Z_j,Z_k) - \theta_n\},$$

where $\gamma_{h2}(Z_j,Z_k) = E\{\gamma_h(Z_j,Z_k,Z_l,Z_m) \mid Z_j,Z_k\}$. The remaining proof consists of two parts: (i) show that $W_n(A)$ is asymptotically normal by Lemma 4; (ii) verify that

$$\frac{E\left\{W_n(A)+\theta_n-\tilde{T}_n(A)\right\}^2}{\mathrm{var}\{W_n(A)\}} \to 0. \tag{A.3}$$

Define

$$\tilde{W}_n(A) = \frac{1}{\binom{n}{2}}\sum_{1\leq j<k\leq n}\gamma_{h2}(Z_j,Z_k) - \frac{2}{n}\sum_{j=1}^{n}\gamma_{h1}(Z_j) + \theta_n,$$

where $\gamma_{h1}(Z_j) = E\{\gamma_h(Z_j,Z_k,Z_l,Z_m) \mid Z_j\}$. $\tilde{W}_n(A)$ is a U-statistic, based on the kernel

$$\tilde{H}(Z_j,Z_k) = \gamma_{h2}(Z_j,Z_k) - \{\gamma_{h1}(Z_j)+\gamma_{h1}(Z_k)\} + \theta_n.$$

Obviously, $E\{\tilde{H}(Z_j,Z_k)\} = 0$ and $E\{\tilde{H}(Z_j,Z_k) \mid Z_j\} = 0$. To show the asymptotic normality of $\tilde{W}_n(A)$, it suffices to check the other conditions in Lemma 4.

Next, we will verify that $E\{\tilde{H}_n^2(Z_1, Z_2)\} \sim h^{-d}$. Note that

$$
\begin{aligned}
&\gamma_{h2}(Z_j, Z_k) && \text{(A.4)}\\
&= \frac{1}{6}\int \Big\{E(\|P_1 - P_2\||Q_1 - Q_2|) + \|X_j - X_k\||Y_j - Y_k| + E(\|X_j - P_1\||Y_j - Q_1|)\\
&\quad + E(\|X_k - P_1\||Y_k - Q_1|) + \|X_j - X_k\|E(|Q_1 - Q_2|) + E(\|P_1 - P_2\|)|Y_j - Y_k|\\
&\quad + E(\|X_j - P_1\|)E(|Y_k - Q_2|) + E(\|X_k - P_1)\|E(|Y_j - Q_2|) - \|X_j - X_k\|E(|Y_j - Q_1|)\\
&\quad - E(\|X_j - P_1\|)|Y_j - Y_k| - \|X_j - X_k\|E(|Y_k - Q_1|) - E(\|X_k - P_1\|)|Y_j - Y_k|\\
&\quad - E(\|P_1 - X_j\|)E(|Q_1 - Q_2|) - E(\|P_1 - X_k\|)E(|Q_1 - Q_2|) - E(\|P_1 - P_2\|E|Q_1 - Y_j|)\\
&\quad - E(\|P_1 - P_2\|)E(|Q_1 - Y_k|)\Big\}.
\end{aligned}
$$

$E\{\gamma_{h2}^2(Z_1, Z_2)\}$ may be expanded into several terms, each of which is of $O(h^{-d})$. We treat only the first such term. Observe that

$$
\begin{aligned}
&\{E(\|P_1 - P_2\||Q_1 - Q_2|)\}^2 \int \cdots \int \Big\{\int K_h(A^{\top}x^{(1)} - B)K_h(A^{\top}x^{(2)} - B)\zeta_0(B)dB\Big\}^2\\
&\times \prod_{r=1}^{2} f_{X,Y}(x^{(r)}, y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}\\
&= h^{-2d} \int\int \Big[\int K(v)K\{v + A^{\top}(x^{(2)} - x^{(1)})/h\}\zeta_0(A^{\top}x^{(1)} - vh)dv\Big]^2\\
&\quad \times \prod_{r=1}^{2} f_{X,Y}(x^{(r)}, y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}dw^{(2)}\\
&\sim h^{-2d} \int\int_{A^{\top}(x^{(2)} - x^{(1)}) = w^{(2)}h} \Big\{K(v)K(v + w^{(2)})\zeta_0(A^{\top}x^{(1)})dv\Big\}^2\\
&\quad \times \prod_{r=1}^{2} f_{X,Y}(x^{(r)}, y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}dw^{(2)} = O(h^{-d}).
\end{aligned}
$$

Now, we deal with $E\{\gamma_{h2}^4(X_1, X_2)\}$. Similar to the arguments above, the integrand can be expanded into several terms, and each of these shown to be of order h^{-3d}. We shall illustrate the purpose in the cause of the first of these terms.

$$
\begin{aligned}
&\{E(\|P_1-P_2\||Q_1-Q_2|)\}^4 \int\cdots\int \Big\{\int K_h(A^{\mathsf T}x^{(1)}-B)K_h(A^{\mathsf T}x^{(2)}-B)\zeta_0(B)dB\Big\}^4 \\
&\quad\times\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)} \\
&=h^{-4d}\int\int\Big[\int K(v)K\{v+A^{\mathsf T}(x^{(2)}-x^{(1)})/h\}\zeta_0(A^{\mathsf T}x^{(1)}-vh)dv\Big]^4 \\
&\quad\times\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}dw^{(2)} \\
&\sim h^{-4d}h^{d}\int\int_{A^{\mathsf T}(x^{(2)}-x^{(1)})=w^{(2)}}\Big\{K(v)K(v+w^{(2)})\zeta_0(A^{\mathsf T}x^{(1)})dv\Big\}^2 \\
&\quad\times\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}dw^{(2)}=O(h^{-3d}).
\end{aligned}
$$

Similarly, we can show that $E\{\gamma_{h1}^2(Z_j)\}=O(h^{2q})$ and $E\{\gamma_{h1}^4(Z_j)\}=O(h^{4q})$. This implies that

$$\{\gamma_{h2}^4(Z_1,Z_2)\}/[E\{\gamma_{h2}^2(Z_1,Z_2)\}]^2\to 0$$

by the condition $nh^{2d}\to\infty$. Finally, let us handle the $E\{\tilde G_n^2(Z_1,Z_2)\}$, where

$$\tilde G_n(x,y)=E\{\tilde H_n(Z_1,x)\tilde H_n(Z_1,y)\}.$$

Again, we focus on $E\{\gamma_{h2}(Z_1,x)\gamma_{h2}(Z_1,y)\}$ because the other terms involved in $E\{\tilde G_n^2(Z_1,Z_2)\}$ are of smaller order.

Note that

$$
\begin{aligned}
&E_{Z_1,Z_2}[E_{Z_1^{(1)}}\{\gamma_{h2}(Z_1^{(1)},Z_1)\gamma_{h2}(Z_1^{(1)},Z_2)\}]^2 \\
&\quad=\int\int\Big\{\int f_{X,Y}(x_1^{(1)},y_1^{(1)})\gamma_{h2}(Z_1^{(1)},Z_1)\gamma_{h2}(Z_1^{(1)},Z_2)dx_1^{(1)}dy_1^{(1)}\Big\}^2 \\
&\quad\quad\times\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}.
\end{aligned}
$$

This integral may be expanded into several terms, each of which is of order h^{-d}. We treat only the first such term

$$
\begin{aligned}
&\{E(\|P_1-P_2\||Q_1-Q_2|)\}^4\int\cdots\int\Big\{\int_{x_1^{(1)},y_1^{(1)}}\int_B K_h(A^\top x_1^{(1)}-B)K_h(A^\top x^{(1)}-B)\zeta_0(B)dB \\
&\quad\times\int_B K_h(A^\top x_1^{(1)}-B)K_h(A^\top x^{(2)}-B)\zeta_0(B)dB f_{X,Y}(x_1^{(r)},y_1^{(r)})dx_1^{(1)}dy_1^{(1)}\Big\}^2 \\
&\quad\times\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)} \\
&=h^{-8d}\int\cdots\int\Big[\int_{x_1^{(1)},y_1^{(1)}}\int_v K(v)K\{v+A^\top(x_1^{(1)}-x^{(1)})/h\}\zeta_0(A^\top x^{(1)}-vh)dv \\
&\quad\int_v K\{v+A^\top(x_1^{(1)}-x^{(1)})/h\}K\{v+A^\top(x^{(2)}-x^{(1)})/h\}\zeta_0(A^\top x^{(1)}-vh)dv \\
&\quad\times f_{X,Y}(x_1^{(r)},y_1^{(r)})dx_1^{(1)}dy_1^{(1)}\Big]^2\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)} \\
&\sim h^{-4d}h^{2d}h^{d}\iint_{A^\top(x^{(2)}-x^{(1)})=w^{(2)}}\Big\{\int_{A^\top(x_1^{(1)}-x^{(1)})=w_1^{(1)}}\int_v K(v)K(v+w_1^{(1)})\zeta_0(A^\top x^{(1)})dv \\
&\quad\int_v K(v+w_1^{(1)})K(v+w^{(2)})\zeta_0(A^\top x^{(1)})dv f_{X,Y}(x_1^{(r)},y_1^{(r)})dx_1^{(1)}dy_1^{(1)}w_1^{(1)}\Big\}^2 \\
&\quad\times\prod_{r=1}^{2} f_{X,Y}(x^{(r)},y^{(r)})dx^{(1)}dx^{(2)}dy^{(1)}dy^{(2)}dw^{(2)}=O(h^{-d}).
\end{aligned}
$$

Thus, the condition (A.2) holds for $\tilde{H}_n(Z_1,Z_2)$. By Lemma 4, $nh^{d/2}\tilde{W}_n(A)$ is asymptotically normal with zero mean and variance

$$
V'_{T0}=2\lim_{n\to\infty}h^dE\{\gamma_{h2}^2(X_1,X_2)\}. \tag{A.5}
$$

Furthermore, by noticing that $E\{2n^{-1}\sum_{j=1}^n\gamma_{h1}(Z_j)\}^2=O(n^{-1}h^{2q})$ and $\theta_n=O(h^{2q})$, $6\tilde{W}_n(A)$ and $W_n(A)$ are asymptotically equivalent from which we finish the first part of the proof.

Now, it remains to check (A.3). $W_n(A)+\theta_n-\tilde{T}_n(A)$ is itself a U-statistic based on the kernel

$$
\breve{H}(Z_1,Z_2,Z_3,Z_4)=\gamma_h(Z_1,Z_1,Z_3,Z_4)-\sum_{1\le i_1<i_2\le 4}\gamma_{h2}(Z_{i_1},Z_{i_2})+5\theta_n.
$$

By using similar arguments above, it can be checked that

$$
\begin{aligned}
\operatorname{var}\{\breve{H}_1(Z_1)\}&=O(h^{2q}),\\
\operatorname{var}\{\breve{H}_2(Z_1,Z_2)\}&=O(h^{2q}),\\
\operatorname{var}\{\breve{H}_3(Z_1,Z_2,Z_3)\}&=O(h^{-2d}),
\end{aligned}
$$

where $\breve{H}_r(z_1,\ldots,z_r)$ denotes the conditional expectation of $\breve{H}$ given $X_1,\ldots,X_r$. By using the standard result on the variance of U-statistic (Serfling 1980), we have

$$
\operatorname{var}\{W_n(A)+\theta_n-\tilde{T}_n(A)\}=n^{-4}O(n^3h^{2q}+n^2h^{2q}+nh^{-2d}).
$$

Recall that $\mathrm{var}\{W_n(A)\} = O(n^{-2}h^{-d})$, (A.3) follows immediately by using the condition $nh^{2d} \to \infty$ and $nh^{2q+d/2} \to 0$.

Finally, by uniformly convergence of $S_{n,2}(A;B)$ and $\hat{f}(B)$

$$\begin{aligned} nh^{d/2}T_n(A) - h^{-d/2}D &= nh^{d/2}\left\{\int_B V_n(A;B)\zeta_0(B)\frac{\hat{f}(B)}{\zeta_0(B)S_{n,2}(A;B)}dB - (nh^d)^{-1}D\right\} \\ &= nh^{d/2}\int \tilde{T}_n(A;B)\zeta_0(B)dB + h^{-d/2}O_p\{h^q + (nh^d)^{-1/2}\log n\} \\ &= nh^{d/2}\tilde{T}_n(A) + o_p(1), \end{aligned}$$

from which we completes the proof of this theorem. ■

Proof of Theorem 3

(i) By Lemma 5, it is straightforward to see $T_n(A) \to \int V(A,B)f_0(B)/S_2(A;B)dB$ in probability. Thus, by the conditions the assertion follows immediately.

(ii) The proof of this part is analogous to that of Theorem 2. Here we only highlight the differences between them. Define $G_{u,v}(t), G_u(t), G_v(t)$ and $G_f(t)$ as the gradients of the functions $f_{X,Y|\mathcal{P}_\mathcal{S}X}(u,v \mid t)$, $f_{X|\mathcal{P}_\mathcal{S}X}(u \mid t)$, $f_{Y|\mathcal{P}_\mathcal{S}X}(v \mid t)$ and $f_0(t)$ with respect to t, respectively. Similarly, denote the corresponding Hessian matrices by $H_{u,v}(t)$, $H_u(t)$, $H_v(t)$, and $H_f(t)$. By Condition 5, we can expand $S_1\{\mathcal{P}_\mathcal{S}^\mathsf{T} + (nh^{d/2})^{-1/2}\Delta; B\}$ at $\mathcal{P}_\mathcal{S}$ in a Taylor series with Lagrange form of the remainder term and obtain

$$\begin{aligned} S_1(A,B) = &\int\int \|x_1 - x_2\||y_1 - y_2|\Big\{f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_1,y_1 \mid B) + G_{u,v}^\mathsf{T}(x_1,y_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})x_1 \\ &+ \frac{1}{2}x_1^\mathsf{T}(A - \mathcal{P}_\mathcal{S}^\mathsf{T})H_{u,v}(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1\Big\}\Big\{f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_2,y_2 \mid B) + G_{u,v}^\mathsf{T}(x_2,y_2)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_2 \\ &+ \frac{1}{2}x_2^\mathsf{T}(A - \mathcal{P}_\mathcal{S}^\mathsf{T})H_{u,v}(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_2\Big\}dx_1dx_2dy_1dy_2 + o\{(nh^{d/2})\}^{-1} \\ =&\mu_x\mu_y + \Lambda_1 + o\{(nh^{d/2})^{-1}\} \\ &+ 2\int\int \|x_1 - x_2\||y_1 - y_2|G_{u,v}^\mathsf{T}(x_1,y_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_2,y_2 \mid B)dx_1dx_2dy_1dy_2, \end{aligned}$$

where Λ_1 denotes those terms associated with the order $(nh^{d/2})^{-1}$.

Similarly,

$$\begin{aligned} S_2(A,B) =&\mu_x\mu_y + \Lambda_2 + 2\mu_y\int\int \|x_1 - x_2\|G_u^\mathsf{T}(x_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1f_{X|\mathcal{P}_\mathcal{S}X}(x_2 \mid B)dx_1dx_2 \\ &+ 2\mu_x\int\int |y_1 - y_2|G_{u,v}^\mathsf{T}(x_1,y_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1f_{Y|\mathcal{P}_\mathcal{S}X}(y_2 \mid B)dx_1dy_1dy_2 + o\{(nh^{d/2})^{-1}\} \\ S_3(A,B) =&\mu_x\mu_y + \Lambda_3 + o\{(nh^{d/2})^{-1}\} \\ &+ \mu_y\int\int \|x_1 - x_2\|G_u^\mathsf{T}(x_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1f_{X|\mathcal{P}_\mathcal{S}X}(x_2 \mid B)dx_1dx_2 \\ &+ \mu_x\int\int\int |y_1 - y_2|G_{u,v}^\mathsf{T}(x_1,y_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1f_{Y|\mathcal{P}_\mathcal{S}X}(y_2 \mid B)dx_1dy_1dy_2 \\ &+ \int\int \|x_1 - x_2\||y_1 - y_2|G_{u,v}^\mathsf{T}(x_1,y_1)(A - \mathcal{P}_\mathcal{S}^\mathsf{T})^\mathsf{T}x_1f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_2,y_2 \mid B)dx_1dx_2dy_1dy_2. \end{aligned}$$

Recall the definition of L_1, L_2 and L_3 in the proof of Theorem 2. By the arguments in Lemma 2, we get

$$
\begin{aligned}
E(\gamma_n) =& \int \{S_1(A,B) + S_2(A,B) - 2S_3(A,B)\} f_0(B) dB + o\{(nh^d)^{-1}\} \\
=&(nh^{d/2})^{-1}\delta + o\{(nh^{d/2})^{-1}\},
\end{aligned}
$$

where $\delta = nh^{d/2}\int(\Lambda_1 + \Lambda_2 - 2\Lambda_3) f_0(B) dB$ with

$$
\begin{aligned}
\Lambda_1 =& \int\cdots\int \|x_1 - x_2\||y_1 - y_2|\Big\{ f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_1, y_1 \mid B) x_2^\top (A - \mathcal{P}_\mathcal{S}^\top) H_{u,v} (A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 \\
&+ G_{u,v}^\top(x_1, y_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 G_{u,v}^\top(x_2, y_2)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 \Big\} dx_1 dx_2 dy_1 dy_2 \\
\Lambda_2 =& \mu_y \int\int \|x_1 - x_2\| \Big\{ f_{X|\mathcal{P}_\mathcal{S}X}(x_1 \mid B) x_2^\top (A - \mathcal{P}_\mathcal{S}^\top) H_u (A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 \\
&+ G_u^\top(x_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 G_u^\top(x_2)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 \Big\} dx_1 dx_2 dB \\
&+ \mu_x \Big\{ \int\int |y_1 - y_2| f_{y|\mathcal{P}_\mathcal{S}X}(y_1 \mid B) x_2^\top (A - \mathcal{P}_\mathcal{S}^\top) H_{u,v}(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 dx_2 dy_1 dy_2 \\
&+ G_{u,v}^\top(x_1, y_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 G_{u,v}^\top(x_2, y_2)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 dx_1 dx_2 dy_1 dy_2 \Big\} \\
&+ 4 \int\int \|x_1 - x_2\| G_u^\top(x_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 f_{X|\mathcal{P}_\mathcal{S}X}(x_2 \mid B) dx_1 dx_2 \\
&\times \int\int\int |y_1 - y_2| G_{u,v}^\top(x_1, y_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 f_{Y|\mathcal{P}_\mathcal{S}X}(y_2 \mid B) dx_1 dy_1 dy_2
\end{aligned}
$$

$$
\begin{aligned}
\Lambda_3 =& \frac{1}{2}\int\cdots\int \|x_1 - x_2\||y_1 - y_2| f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_1, y_1 \mid B) x_2^\top (A - \mathcal{P}_\mathcal{S}^\top) H_{u,v}(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 dx_1 dx_2 dy_1 dy_2 \\
&+ \frac{1}{2}\mu_y \int\int \|x_1 - x_2\| f_{X|\mathcal{P}_\mathcal{S}X}(x_1 \mid B) x_2^\top (A - \mathcal{P}_\mathcal{S}^\top) H_u (A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 dx_1 dx_2 \\
&+ \frac{1}{2}\mu_x \int\int |y_1 - y_2| f_{y|\mathcal{P}_\mathcal{S}X}(y_1 \mid B) x_2^\top (A - \mathcal{P}_\mathcal{S}^\top) H_{u,v}(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 dx_2 dy_1 dy_2 \\
&+ \int\cdots\int \|x_1 - x_2\||y_1 - y_3| \Big\{ \\
&G_{u,v}^\top(x_1, y_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 G_u^\top(x_2)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 f_{Y|\mathcal{P}_\mathcal{S}X}(y_3 \mid B) dx_1 dx_2 dy_1 dy_3 \\
&+ G_{u,v}^\top(x_1, y_1)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_1 G_{u,v}^\top(x_3, y_3)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_3 f_{X|\mathcal{P}_\mathcal{S}X}(x_2 \mid B) dx_1 dx_2 dx_3 dy_1 dy_3 \\
&+ G_u^\top(x_2)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_2 G_{u,v}^\top(x_3, y_3)(A - \mathcal{P}_\mathcal{S}^\top)^\top x_3 f_{X,Y|\mathcal{P}_\mathcal{S}X}(x_1, y_1 \mid B) dx_1 dx_2 dx_3 dy_1 dy_3 \Big\}.
\end{aligned}
$$

Note that $G_{u,v}(x, y) = G_u(x) f_{Y|\mathcal{P}_\mathcal{S}X}(y) + G_v(y) f_{X|\mathcal{P}_\mathcal{S}X}(x)$. Accordingly, tedious algebras yield

$$
\delta = \int \left\{ \int G_v^\top(B) \Delta^\top x f_{X|\mathcal{P}_\mathcal{S}X}(x \mid B) dx \right\}^2 dB.
$$

The remaining proofs follow exactly same as those of Theorem 2 and thus omitted. The variance V_{T1} is given by $V_{T1} = 72 \lim_{n\to\infty} h^d E\{\gamma_{h2}^2(X_1, X_2)\}$, where $\gamma_{h2}^2(\cdot,\cdot)$ is given by (A.4) but the expectations now are all taken under the alternative hypothesis. ∎

APPENDIX C: CROSS VALIDATION BOXPLOTS

Fig. 1 gives the cross validation boxplots based on the simulations leading to Fig. 1 in the paper. On each replication the data were split randomly into two sets of $n/2$ observations, with one set being used to determine $\hat{A}$ and the other being used for the permutation test.

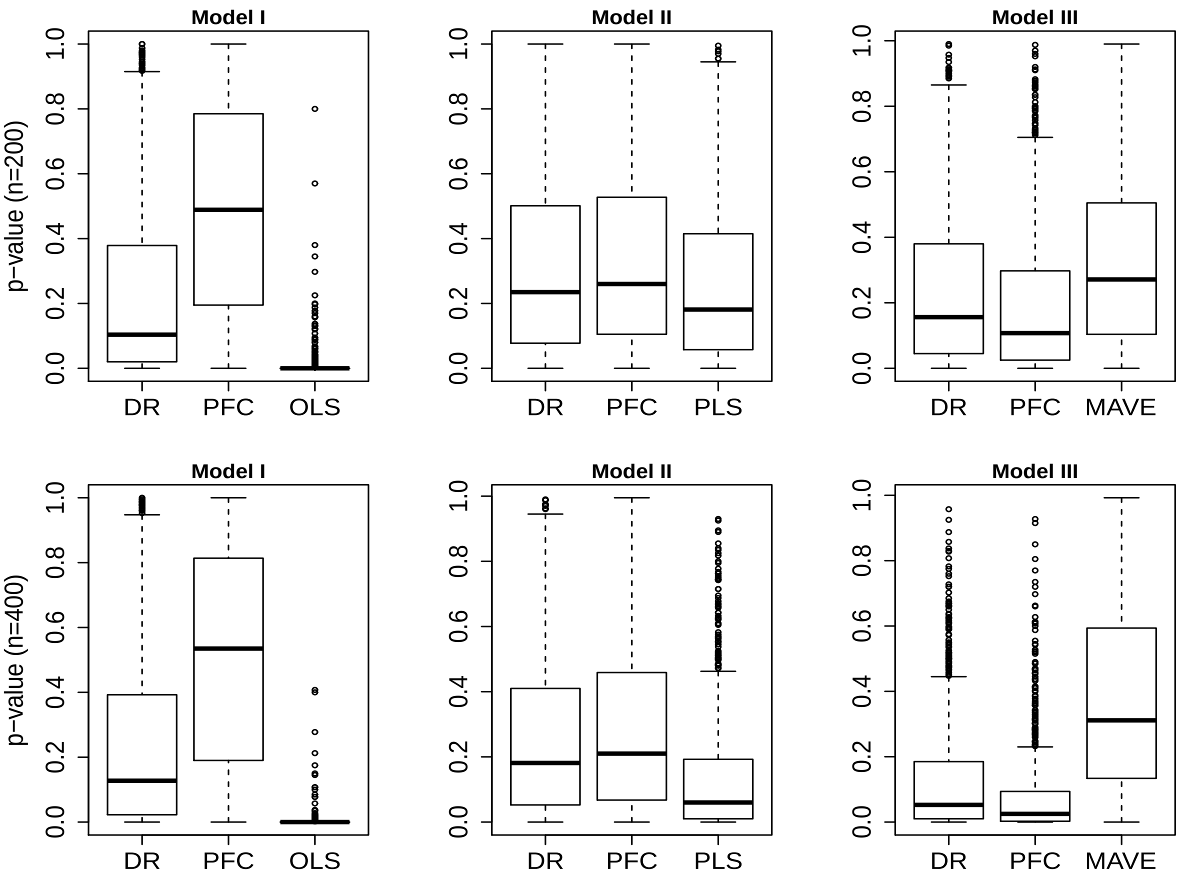

Fig. 1. Cross validation boxplots of p-values from the proposed test using directional regression (DR), ordinary least squares (OLS), principal fitted components (PFC), partial least squares (PLS) and minimum average variance estimation (MAVE)

REFERENCES

SILVERMAN, B. W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.* **6** 177–189.

GINE, E. & GUILLOU, A. (2002). Rates of strong uniform consistency for multivariate kernel density estimators. *Ann. I. H. Poincare - PR* **38** 907–921.

HALL, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric estimators. *J. Multivar. Anal.* **14**, 1–6.

POLLARD, D. (1984). *Convergence of Stochastic Processes*, Springer, New York.

SERFLING, R. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.