Supplementary Material of Testing the Conditional Mean Independence for Functional Data

BY CHUNG EUN LEE, XIANYANG ZHANG, AND XIAOFENG SHAO
University of Tennessee, Knoxville, Texas A&M University, and University of Illinois at Urbana-Champaign

The supplementary material includes all the technical proofs and additional simulations.

1. APPENDIX

1.1. *Technical Proofs*

Proof of Proposition 1. If $E(Y \mid X) = \mu_Y$ almost surely, it is clear that $\text{FMDD}(Y \mid X) = 0$. We only need to show the other direction i.e., $\text{FMDD}(Y \mid X) = 0$ implies $E(Y \mid X) = \mu_Y$ almost surely. Without loss of generality, we can assume that $\mu_Y = 0$ (otherwise we work with $Y - \mu_Y$). By Theorem 3.16 and Proposition 3.1 in Lyons (2013), there exists an embedding $\phi : \mathcal{L}_2(\mathcal{I})$ (or $\mathbb{R}^q$) $\to \mathcal{H}$ such that $|x - x'| = |\phi(x) - \phi(x')|^2$ and $\beta_\phi(v) = \int \phi(x) dv(x)$ is injective on the set of measures v on $\mathcal{L}_2(\mathcal{I})$ (or $\mathbb{R}^q$) such that $|v|$ has a finite first moment, i.e., $\int |x - o| d|v|(x) < \infty$, for some $o \in \mathcal{L}_x$. Using this result and the definition of FMDD, we have

$$\begin{aligned}\text{FMDD}(Y \mid X) &= 2 \int < y, y' >< \phi(x) - \beta_\phi(\mu), \phi(x') - \beta_\phi(\mu) > d\theta(x, y)\theta(x', y') \\ &= \left| E\{Y \otimes \phi(X)\}^2 \right| \geq 0,\end{aligned}$$

where μ denotes the distribution of X and θ is the joint distribution of (X, Y). Hence, $FMDD(Y \mid X) = 0$ implies that

$$E\{Y \otimes \phi(X)\} = \int y\phi(x) d\theta(x, y) = 0 \quad \text{almost surely.}$$

For any Borel set $B \subseteq \mathcal{L}_2(\mathcal{I})$ (or $\mathbb{R}^q$) and $k \in \mathcal{L}_2(\mathcal{I})$ (or $\mathbb{R}^p$), define the sign measure,

$$v_k(B) = \int < y, k > 1_B(x) d\theta(x, y) = E\{< Y, k > 1_B(X)\},$$

where $|v_k|$ has a finite first moment under the assumptions that $E(|X| + |Y|) < \infty$ and $E(|X - \mu_X||Y - \mu_Y|) < \infty$. Then we have

$$\beta_\phi(v_k) = \int < y, k > \phi(x) d\theta(x, y) =< \int y\phi(x) d\theta(x, y), k >= 0.$$

The injectivity of β_ϕ gives $v_k(B) = E\{< Y, k > 1_B(X)\} = 0$. Thus, by the definition of conditional mean independence, we have

$$E(< Y, k >\mid X) = 0, \tag{1}$$

for any $k \in \mathcal{L}_2(\mathcal{I})$ (or $\mathbb{R}^p$). Therefore, (1) implies that $E(Y \mid X) = \mu_Y$, which completes the proof of Proposition 1. $\square$

Proof of Proposition 2. Following the arguments in Section 1.1 of the supplement of Zhang et al. (2018), we can show that $\text{FMDD}_n(Y \mid X)$ is an unbiased estimator of $\text{FMDD}(Y \mid X)$, and it is a fourth-order U-statistic which has the form of

$$\mathrm{FMDD}_n(Y \mid X) = \frac{1}{\binom{n}{4}} \sum_{i<j<q<r} h(Z_i, Z_j, Z_q, Z_r),$$

$$h(Z_i, Z_j, Z_q, Z_r) = \frac{1}{4!} \sum_{(s,t,u,v)}^{(i,j,q,r)} (a_{st}b_{uv} + a_{st}b_{st} - a_{st}b_{su} - a_{st}b_{tv}),$$

where $\sum_{(s,t,u,v)}^{(i,j,q,r)}$ denotes the summation over all permutations of the 4-tuple of indices (i, j, q, r) and $Z_i = (X_i, Y_i)$. Under the assumption that $E(|X| + |Y|) < \infty$ and $E(|X - \mu_X||Y - \mu_Y|) < \infty$, we have

$$\begin{aligned} E\{|h(Z_i, Z_j, Z_q, Z_r)|\} \leq & \frac{1}{4!} \sum_{(s,t,u,v)}^{(i,j,q,r)} E|a_{st}b_{uv} + a_{st}b_{st} - a_{st}b_{su} - a_{st}b_{tv}| \\ \leq & E(|X - X'|)E(|Y - \mu_Y|)^2 + E(|X - X'||Y - \mu_Y||Y' - \mu_Y|) \\ & + 2E(|X - X'||Y - \mu_Y|)E(|Y - \mu_Y|) < \infty. \end{aligned}$$

Proposition 2 follows from the law of large numbers for U-statistics [see e.g. Hoeffding (1961) and Lee (1990)]. □

Proof of Theorem 1. For $c = 1, 2, 3, 4$, define

$$h_c(z_1, \cdots, z_c) = E\{h(z_1, \cdots, z_c, Z_{c+1}, \cdots, Z_4)\},$$

where $z_i = (x_i, y_i)$ for $1 \leq i \leq 4$. Denote by $Z' = (X', Y')$ and $Z'' = (X'', Y'')$ two independent copies of $Z = (X, Y)$. When $FMDD(Y \mid X) = 0$, following the calculations in Section 1.2 of the supplement of Zhang et al. (2018), we have $h_1(z) = 0$ and $h_2(z, z') = U(x, x')V(y, y')/6$ for $z = (x, y)$ and $z' = (x', y')$. Under the assumption $E(|X|^2 + |Y|^2) < \infty$ and $E(|X - \mu_X|^2|Y - \mu_Y|^2) < \infty$, we have $E\{h(Z_i, Z_j, Z_q, Z_r)^2\} < \infty$. Applying Theorem 5.5.2 in Serfling (1980), we obtain $n\mathrm{FMDD}_n(Y \mid X) \to^D \sum_{k=1}^{\infty} \lambda_k(G_k^2 - 1)$. □

Proof of Theorem 2. Under the local alternative $H_{1,n} : Y = \mu_Y + \frac{g(X)}{n^a} + \epsilon$, we have

$$\begin{aligned} |Y_i - Y_j|^2 = & \frac{1}{n^{2a}}|g(X_i) - g(X_j)|^2 + |\epsilon_i - \epsilon_j|^2 + \frac{2}{n^a} < g(X_i) - g(X_j), \epsilon_i - \epsilon_j > \\ = & \frac{1}{n^{2a}}|g(X_i) - g(X_j)|^2 + |\epsilon_i - \epsilon_j|^2 \\ & + \frac{1}{n^a}\left[|g(X_i) + \epsilon_i - \{g(X_j) + \epsilon_j\}|^2 - |g(X_i) - g(X_j)|^2 - |\epsilon_i - \epsilon_j|^2\right] \end{aligned}$$

Using the above result, $\mathrm{FMDD}_n\left\{\frac{g(X)}{n^a} + \epsilon \middle| X\right\}$ can be decomposed into three terms

$$\begin{aligned} \mathrm{FMDD}_n\left\{\frac{g(X)}{n^a} + \epsilon \middle| X\right\} = & \frac{1}{n^{2a}n(n-3)} \sum_{i \neq j} \widetilde{A}_{ij}\widetilde{B}_{ij}^g + \frac{1}{n(n-3)} \sum_{i \neq j} \widetilde{A}_{ij}\widetilde{B}_{ij}^\epsilon \\ & + \frac{1}{n^a n(n-3)} \sum_{i \neq j} \widetilde{A}_{ij}\left(\widetilde{B}_{ij}^{g+\epsilon} - \widetilde{B}_{ij}^g - \widetilde{B}_{ij}^\epsilon\right), \end{aligned} \tag{2}$$

where $\widetilde{B}^{\epsilon}_{ij} = e_{ij} - e_{i\cdot} - e_{\cdot j} + e_{\cdot\cdot}$, with $e_{ij} = \frac{1}{2}|\epsilon_i - \epsilon_j|^2$, and $e_{ij},\ e_{i\cdot},\ e_{\cdot j},\ e_{\cdot\cdot}$ are defined similarly as $b_{ij},\ b_{i\cdot},\ b_{\cdot j},\ b_{\cdot\cdot}$. Moreover, $\widetilde{B}^{g+\epsilon}_{ij}$ and $\widetilde{B}^{g}_{ij}$ are defined similarly by replacing (ϵ_i, ϵ_j) in $\widetilde{B}^{\epsilon}_{ij}$ with $\{g(X_i) + \epsilon_i, g(X_j) + \epsilon_j\}$ and $\{g(X_i), g(X_j)\}$ respectively.

Because $\frac{1}{n(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}^{\epsilon}_{ij}$ is a degenerate U-statistic, whereas $\frac{1}{n(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}^{g}_{ij}$ and $\frac{1}{n(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\left(\widetilde{B}^{g+\epsilon}_{ij} - \widetilde{B}^{g}_{ij} - \widetilde{B}^{\epsilon}_{ij}\right)$ are nondegenerate U-statistics (to be shown below), (2) implies that

$$n\mathrm{FMDD}_n\left\{\frac{g(X)}{n^a} + \epsilon \,\middle|\, X\right\} = n^{1-2a}FMDD\{g(X) \mid X\} + O_p(n^{1/2-2a}) + O_p(1) + O_p(n^{1/2-a}) \tag{3}$$

We shall consider three scenarios: *1. if* $0 < a < 1/2$, *2. if* $a = 1/2$, *3. if* $a > 1/2$.
1. if $0 < a < 1/2$:
Based on (3), we can easily show that $n\mathrm{FMDD}_n\left\{\frac{g(X)}{n^a} + \epsilon \,\middle|\, X\right\} \to^p \infty$ which implies that our test has consistency under this scenario.
3. if $a > 1/2$:
Similarly, using (2) and Theorem 1, we have

$$n\mathrm{FMDD}_n\left\{\frac{g(X)}{n^a} + \epsilon \,\middle|\, X\right\} = \frac{1}{(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}^{\epsilon}_{ij} + o_p(1) \to^D \sum_{k=1}^{\infty}\lambda_k(G_k^2 - 1),$$

which is the same as the limiting null distribution.
2. if $a = 1/2$:
When $a = 1/2$, $\mathrm{FMDD}_n\left\{\frac{g(X)}{n^{1/2}} + \epsilon \,\middle|\, X\right\}$ can be written as a sum of linear combination of two U-statistics, $U_n^{\epsilon},\ U_n^{g,\epsilon}$ and a sequence of random variables c_n where $nc_n \to^{a.s.} \mathrm{FMDD}\{g(X) \mid X\} = c$, i.e.,

$$\begin{aligned}\mathrm{FMDD}_n\left\{\frac{g(X)}{n^{1/2}} + \epsilon \,\middle|\, X\right\} &= \frac{1}{n^2(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}^{g}_{ij} + \frac{1}{n(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\left\{\widetilde{B}^{\epsilon}_{ij} + \frac{1}{n^{1/2}}(\widetilde{B}^{g+\epsilon}_{ij} - \widetilde{B}^{g}_{ij} - \widetilde{B}^{\epsilon}_{ij})\right\} \\ &= c_n + U_n^{\epsilon} + \frac{1}{n^{1/2}}U_n^{g,\epsilon}. \end{aligned} \tag{4}$$

Specifically, $U_n^{\epsilon},\ U_n^{g,\epsilon}$ are mean zero U-statistics of degree 4, i.e.,

$$\begin{aligned} U_n^{\epsilon} &= \frac{1}{n(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}^{\epsilon}_{ij} = \frac{1}{\binom{n}{4}}\sum_{i<j<q<r} H_1(\mathcal{Z}_i, \mathcal{Z}_j, \mathcal{Z}_q, \mathcal{Z}_r), \\ U_n^{g,\epsilon} &= \frac{1}{n(n-3)}\sum_{i\neq j}\widetilde{A}_{ij}(\widetilde{B}^{g+\epsilon}_{ij} - \widetilde{B}^{g}_{ij} - \widetilde{B}^{\epsilon}_{ij}) = \frac{1}{\binom{n}{4}}\sum_{i<j<q<r} h_1(\mathcal{Z}_i, \mathcal{Z}_j, \mathcal{Z}_q, \mathcal{Z}_r), \end{aligned}$$

where $\mathcal{Z}_i = (X_i, \epsilon_i)$,

$$H_1(\mathcal{Z}_i, \mathcal{Z}_j, \mathcal{Z}_q, \mathcal{Z}_r) = \frac{1}{4!}\sum_{(s,t,u,v)}^{(i,j,q,r)} a_{st}(e_{uv} + e_{st} - e_{su} - e_{tv}),$$

$$h_1(\mathcal{Z}_i, \mathcal{Z}_j, \mathcal{Z}_q, \mathcal{Z}_r) = \frac{1}{4!}\sum_{(s,t,u,v)}^{(i,j,q,r)} a_{st}(m_{uv} + m_{st} - m_{su} - m_{tv}),$$

and $m_{ij} =< g(X_i) - g(X_j), \epsilon_i - \epsilon_j >$. Define

$$\widetilde{h}_{1,i}(z_1, \cdots, z_i) = E\{h_1(z_1, \cdots, z_i, \mathcal{Z}_{i+1}, \cdots, \mathcal{Z}_4)\},$$

$$\widetilde{H}_{1,i}(z_1, \cdots, z_i) = E\{H_1(z_1, \cdots, z_i, \mathcal{Z}_{i+1}, \cdots, \mathcal{Z}_4)\},$$

for $1 \leq i \leq 4$. By the results in the supplement of Zhang et al. (2018) and some calculations, we show that $\widetilde{h}_{1,1}(z_1) = \frac{1}{2}E\{U(x_1, X)V(\epsilon_1, g(X))\},\ \widetilde{H}_{1,1}(z_1) = 0,$

$$\begin{aligned}
\widetilde{h}_{1,2}(z_1, z_2) = \frac{1}{6}\Big[& U(x_1, x_2)\{V(\epsilon_1, g(x_2)) + V(g(x_1), \epsilon_2)\} \\
& + E\{U(x_1, X)V(\epsilon_1, g(X))\} + E\{U(x_2, X)V(\epsilon_2, g(X))\} \\
& + E\left[\{U(x_1, X) - U(x_2, X)\}\{V(\epsilon_1, g(X)) - V(\epsilon_2, g(X)\}\right]\Big], \\
\widetilde{H}_{1,2}(z_1, z_2) = \frac{1}{6} & U(x_1, x_2)V(\epsilon_1, \epsilon_2),
\end{aligned}$$

$$\begin{aligned}
\widetilde{h}_{1,3}(z_1, z_2, z_3) = \frac{1}{12}\Big[& \{2U(x_1, x_2) - U(x_2, x_3) - U(x_1, x_3)\}\{V(\epsilon_1, g(x_2)) + V(g(x_1), \epsilon_2)\} \\
& + \{2U(x_1, x_3) - U(x_1, x_2) - U(x_2, x_3)\}\{V(\epsilon_1, g(x_3)) + V(g(x_1), \epsilon_3)\} \\
& + \{2U(x_2, x_3) - U(x_1, x_2) - U(x_1, x_3)\}\{V(\epsilon_2, g(x_3)) + V(g(x_2), \epsilon_3)\} \\
& + E\left[\{2U(x_1, X) - U(x_2, X) - U(x_3, X)\}V(\epsilon_1, g(X))\right] \\
& + E[\{2U(x_2, X) - U(x_1, X) - U(x_3, X)\}V(\epsilon_2, g(X))] \\
& + E[\{2U(x_3, X) - U(x_1, X) - U(x_2, X)\}V(\epsilon_3, g(X))]\Big], \\
\widetilde{H}_{1,3}(z_1, z_2, z_3) = \frac{1}{12}\Big[& \{2U(x_1, x_2) - U(x_2, x_3) - U(x_1, x_3)\}V(\epsilon_1, \epsilon_2) \\
& + \{2U(x_1, x_3) - U(x_1, x_2) - U(x_2, x_3)\}V(\epsilon_1, \epsilon_3) \\
& + \{2U(x_2, x_3) - U(x_1, x_2) - U(x_1, x_3)\}V(\epsilon_2, \epsilon_3)\Big].
\end{aligned}$$

Note that $\widetilde{h}_{1,4}(z_1, z_2, z_3, z_4) = h_1(z_1, z_2, z_3, z_4)$ and $\widetilde{H}_{1,4}(z_1, z_2, z_3, z_4) = H_1(z_1, z_2, z_3, z_4)$. Under the assumptions that $E(|X|^2 + |g(X)|^2 + |\epsilon|^2) < \infty$ and $E[|X - \mu_X|^2\{|g(X)|^2 + |\epsilon|^2\}] < \infty$, it is guaranteed that $\text{var}\{h_1(\mathcal{Z}, \mathcal{Z}', \mathcal{Z}'', \mathcal{Z}''')\} < \infty$ and $\text{var}\{H_1(\mathcal{Z}, \mathcal{Z}', \mathcal{Z}'', \mathcal{Z}''')\} < \infty$. Moreover, by the results in Section 5.2.1 (page 182) and Lemma 5.1.5A in Serfling (1980), we have $0 < \text{var}\{\widetilde{h}_{1,1}(\mathcal{Z})\} < \infty$ and obtain

$$n^{1/2}U_n^{g,\epsilon} = \frac{4}{n^{1/2}}\sum_{i=1}^{n}\widetilde{h}_{1,1}(\mathcal{Z}_i) + \mathcal{R}_{1,n}, \tag{5}$$

where $\mathcal{R}_{1,n}$ is asymptotically negligible. Similarly we have $0 < \text{var}\{\widetilde{H}_{1,2}(\mathcal{Z}, \mathcal{Z}')\} < \infty$ and obtain

$$nU_n^{\epsilon} = \frac{6}{(n-1)}\sum_{i \neq j}\widetilde{H}_{1,2}(\mathcal{Z}_i, \mathcal{Z}_j) + \mathcal{R}_{2,n}, \tag{6}$$

where $\mathcal{R}_{2,n}$ is asymptotically negligible. Based on (5) and (6), we deduce that

$$\begin{aligned} n^{1/2}U_n^{g,\epsilon} + nU_n^{\epsilon} &= \frac{4}{n^{1/2}}\sum_{i=1}^{n}\widetilde{h}_{1,1}(\mathcal{Z}_i) + \frac{6}{(n-1)}\sum_{i\neq j}\widetilde{H}_{1,2}(\mathcal{Z}_i,\mathcal{Z}_j) + \mathcal{R}_n \\ &= n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2} + \mathcal{R}_n, \end{aligned} \tag{7}$$

where $\mathcal{R}_n$ is asymptotically negligible and

$$\mathcal{U}_{n1} = \frac{4}{n}\sum_{i=1}^{n}\widetilde{h}_{1,1}(\mathcal{Z}_i),\ \mathcal{U}_{n2} = \frac{6}{n(n-1)}\sum_{i\neq j}\widetilde{H}_{1,2}(\mathcal{Z}_i,\mathcal{Z}_j).$$

Next we shall find the limiting distribution of $n^{1/2}U_n^{g,\epsilon} + nU_n^{\epsilon}$. Applying Dunford and Schwartz (1963) to $6\widetilde{H}_{1,2}(\mathcal{Z},\mathcal{Z}')$, we obtain

$$6\widetilde{H}_{1,2}(\mathcal{Z},\mathcal{Z}') = \sum_{k=1}^{\infty}\lambda_k\psi_k(\mathcal{Z})\psi_k(\mathcal{Z}'),$$

where $\{\lambda_k,\psi_k(\cdot)\}$ is a sequence of eigenvalues and eigenfunctions of $6\widetilde{H}_{1,2}$ and the eigenfunctions are orthogonal in the sense that $E\{\psi_i(\mathcal{Z})\psi_j(\mathcal{Z})\} = \delta_{ij}$. Note that

$$\begin{aligned} E\{6^2\widetilde{H}_{1,2}(\mathcal{Z},\mathcal{Z}')\widetilde{H}_{1,2}(\mathcal{Z},\mathcal{Z}'')\} &= E\{U(X,X')U(X,X'')V(\epsilon,\epsilon')V(\epsilon,\epsilon'')\} = 0 \\ &= E\left\{\sum_k\sum_l \lambda_k\lambda_l\psi_k(\mathcal{Z})\psi_k(\mathcal{Z}')\psi_l(\mathcal{Z})\psi_l(\mathcal{Z}'')\right\} \\ &= \sum_k \lambda_k^2 E\{\psi_k(\mathcal{Z}')\}E\{\psi_k(\mathcal{Z}'')\} \\ &= \sum_k \lambda_k^2 E\{\psi_k(\mathcal{Z}')\}^2, \end{aligned}$$

which implies that

$$E\{\psi_k(\mathcal{Z})\} = 0,\ \forall k. \tag{8}$$

For convenience, we let $\mathcal{U}_{n2}$ be $\frac{6}{n^2}\sum_{i\neq j}\widetilde{H}_{1,2}(\mathcal{Z}_i,\mathcal{Z}_j)$ which will not affect the limiting distribution of $n^{1/2}U_n^{g,\epsilon} + nU_n^{\epsilon}$. Then the leading term of $n^{1/2}U_n^{g,\epsilon} + nU_n^{\epsilon}$, which is $n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2}$ can be rewritten as

$$n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2} = \frac{4}{n^{1/2}}\sum_{i=1}^{n}\widetilde{h}_{1,1}(\mathcal{Z}_i) + \sum_{k=1}^{\infty}\lambda_k\left[\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\psi_k(\mathcal{Z}_i)\right\}^2 - \frac{1}{n}\sum_{i=1}^{n}\psi_k(\mathcal{Z}_i)^2\right].$$

Then we apply multivariate CLT to $\frac{1}{n^{1/2}}\sum_{i=1}^{n}\psi_k(\mathcal{Z}_i)$ and $\frac{1}{n^{1/2}}\sum_i\widetilde{h}_{1,1}(\mathcal{Z}_i)$. Due to (8) and $E\{\widetilde{h}_{1,1}(\mathcal{Z})\} = 0$, for a fixed positive integer K, we have

$$\begin{pmatrix} \frac{1}{n^{1/2}}\sum_{i=1}^{n}\psi_1(\mathcal{Z}_i) \\ \vdots \\ \frac{1}{n^{1/2}}\sum_{i=1}^{n}\psi_K(\mathcal{Z}_i) \\ \frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{h}_{1,1}(\mathcal{Z}_i) \end{pmatrix} \to^D Z^* \sim N(\mathbf{0},\Sigma), \tag{9}$$

where

$$\Sigma = \begin{pmatrix} 1 & \cdots & 0 & E\{\psi_1(\mathcal{Z})\widetilde{h}_{1,1}(\mathcal{Z})\} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & E\{\psi_K(\mathcal{Z})\widetilde{h}_{1,1}(\mathcal{Z})\} \\ E\{\psi_1(\mathcal{Z})\widetilde{h}_{1,1}(\mathcal{Z})\} & \cdots & E\{\psi_K(\mathcal{Z})\widetilde{h}_{1,1}(\mathcal{Z})\} & \mathrm{var}\{\widetilde{h}_{1,1}(\mathcal{Z})\} \end{pmatrix}. \tag{10}$$

We shall use the truncation method to show

$$n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2} \to^D G + \sum_{k=1}^{\infty} \lambda_k (G_k^2 - 1),$$

where (G_k) are independent standard normal random variables and G is normal random variable with zero mean and variance equal to $16\mathrm{var}\{\widetilde{h}_{1,1}(\mathcal{Z})\}$ which is possibly correlated with (G_k); see (10). Define $n\mathcal{U}_{n2}^{(K)} = \frac{1}{n}\sum_{i\neq j}\sum_{k=1}^{K} \lambda_k \psi_k(\mathcal{Z}_i)\psi_k(\mathcal{Z}_j)$ and notice that

$$E\{(n\mathcal{U}_{n2} - n\mathcal{U}_{n2}^{(K)})^2\} = \frac{1}{n^2}\sum_{i\neq j}\sum_{k=K+1}^{\infty} \lambda_k^2 \to 0, \tag{11}$$

as $K \to +\infty$ due to the fact that $E\{6^2\widetilde{H}_{1,2}(\mathcal{Z},\mathcal{Z}')^2\} = \sum_{k=1}^{\infty}\lambda_k^2 < \infty$. Then by (11) and the Markov inequality, we can show that for any $\delta > 0$,

$$\lim_{K\to+\infty}\limsup_{n\to\infty} \mathrm{pr}\left\{|n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2} - (n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2}^{(K)})| \geq \delta\right\} = 0. \tag{12}$$

Moreover, due to (9), it is obvious that for any fixed K,

$$\begin{aligned} n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2}^{(K)} &= \frac{4}{n^{1/2}}\sum_{i=1}^{n}\widetilde{h}_1(\mathcal{Z}_i) + \sum_{k=1}^{K}\lambda_k\left[\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\psi_k(\mathcal{Z}_i)\right\}^2 - \frac{1}{n}\sum_{i=1}^{n}\psi_k(\mathcal{Z}_i)^2\right] \\ &\to^D G + \sum_{k=1}^{K}\lambda_k(G_k^2 - 1). \end{aligned} \tag{13}$$

Since (12) and (13) are satisfied, we have the following result by using Theorem 2 in Dehling et al. (2009).

$$n^{1/2}\mathcal{U}_{n1} + n\mathcal{U}_{n2} \to^D G + \sum_{k=1}^{\infty}\lambda_k(G_k^2 - 1). \tag{14}$$

□

Therefore, due to (4), (7), and (14), we finally conclude that

$$n\left[\mathrm{FMDD}_n\left\{\frac{g(X)}{n^{1/2}} + \epsilon\middle|X\right\} - \frac{1}{n}\mathrm{FMDD}\{g(X) \mid X\}\right] \to^D G + \sum_{k=1}^{\infty}\lambda_k(G_k^2 - 1).$$

Proof of Theorem 3. By the Hoeffding decomposition, we have

$$\mathrm{FMDD}_n(Y \mid X) - \mathrm{FMDD}(Y \mid X) = \frac{2}{n}\sum_{i=1}^{n}\{K(Z_i) - \mathrm{FMDD}(Y \mid X)\} + R_n,$$

where R_n is asymptotically negligible. Hence, using Theorem 5.5.1 in Serfling (1980), we obtain

$$n^{1/2}\{\mathrm{FMDD}_n(Y \mid X) - \mathrm{FMDD}(Y \mid X)\} \to^D N(0, 4\sigma_1^2),$$

where $\sigma_1^2 = \text{var}(K(Z))$. $\square$

LEMMA 1. *Let $\{X_i\}_{i\geq 1}$ be a sequence of identically distributed random elements defined on the same probability space $(\Omega, \mathcal{B}, P)$ with $E(|X_1|) < \infty$. Let $Y_n = n^{-1}\max_{1\leq i\leq n}|X_i|$. Then $Y_n \to^{a.s.} 0$.*

Proof of Lemma 1. For any $\epsilon > 0$, we have

$$\sum_{n=1}^{+\infty}\text{pr}(|X_n| > \epsilon n) = \sum_{n=1}^{+\infty}\text{pr}(|X_1| > \epsilon n) < \infty,$$

as $E(|X_1|) < \infty$ [see Lemma 7.5.1 of Resnick (2005)]. By the Borel-Cantelli Lemma, we have $\text{pr}\{\liminf_n(|X_n| \leq \epsilon n)\} = 1$. Let $A = \cap_{m=1}^{\infty}\liminf_n(|X_n| \leq n/m)$. Then $\text{pr}(A) = 1$. For any $w \in A$, there exists $n_0 = n_0(w; m)$ such that for $n \geq n_0(w; m)$, $|X_n| \leq n/m$. Thus we have

$$\begin{aligned} Y_n(w) \leq & n^{-1}\max_{1\leq i\leq n_0-1}|X_i(w)| + n^{-1}\max_{n_0\leq i\leq n}|X_i(w)| \\ \leq & n^{-1}\max_{1\leq i\leq n_0-1}|X_i(w)| + 1/m \to 1/m. \end{aligned}$$

Since m can be arbitrarily large, $\lim_{n\to+\infty} Y_n(w) = 0$, which implies that $Y_n \to^{a.s.} 0$. $\square$

LEMMA 2. *If $E\{\mathcal{H}(Z, Z')^4\} < \infty$ and $\nu_k \neq 0$, then $E\{\phi_k(Z)^4\} < \infty$, where ν_k is an eigenvalue which corresponds to the kth eigenfunction of $\mathcal{H}$, $\phi_k(\cdot)$.*

Proof of Lemma 2. Note that $\nu_k\phi_k(Z) = E\{\mathcal{H}(Z, Z')\phi_k(Z') \mid Z\}$. By the Cauchy-Schwarz inequality and the fact that $E\{\phi_k(Z')^2 \mid Z\} = E\{\phi_k(Z')^2\} = 1$, we have

$$\begin{aligned} \nu_k^4 E\{\phi_k(Z)^4\} =& E[E\{\mathcal{H}(Z, Z')\phi_k(Z') \mid Z\}^4] \\ \leq& E[E\{\mathcal{H}(Z, Z')^2 \mid Z\}^2 E\{\phi_k(Z')^2 \mid Z\}^2] \\ \leq& E\{\mathcal{H}(Z, Z')^4\} < \infty. \end{aligned}$$

Thus $E\{\phi_k(Z)^4\} < \infty$ as $\nu_k \neq 0$. $\square$

Proof of Theorem 4. Let $\mathcal{L}_2(\mu)$ be the space consisting of all square integrable functions with respect to the measure induced by Z (say μ). Let $\mathcal{H}(\cdot, \cdot)$ be a symmetric bivariate function with $E\{\mathcal{H}(Z, Z')^2\} < \infty$, where Z' is an independent copy of Z. Define the linear operator $(Hf)(s) = \int \mathcal{H}(s, t)f(t)\mu(dt)$ for $f \in \mathcal{L}_2(\mu)$. According to Dunford and Schwartz (1963, p108, Exercise 56), $\mathcal{H}(z, z')$ admits the series decomposition,

$$\mathcal{H}(z, z') = \sum_{k=1}^{\infty}\nu_k\phi_k(z)\phi_k(z'),$$

where (ν_k) and (ϕ_k) are the eigenvalues and eigenfunctions of H (with respect to μ) respectively, i.e., $H\phi_k = \nu_k\phi_k$ and $E\{\phi_i(Z)\phi_j(Z)\} = \delta_{ij}$.

Define $\mathcal{H}^{(K)}(Z, Z') = \sum_{k=1}^{K}\nu_k\phi_k(Z)\phi_k(Z')$. As $E\{\mathcal{H}(Z, Z')^2\} = \sum_{k=1}^{\infty}\nu_k^2 < \infty$, we have

$$\lim_{K\to\infty} E[\{\mathcal{H}(Z, Z') - \mathcal{H}^{(K)}(Z, Z')\}^2] = \lim_{K\to\infty}\sum_{k=K+1}^{\infty}\nu_k^2 = 0, \tag{15}$$

which indicates that $\mathcal{H}^{(K)}(Z, Z')$ approximates $\mathcal{H}(Z, Z')$ as $K \to +\infty$. We define $nU_n^* = \frac{1}{n-1}\sum_{i\neq j}\mathcal{H}(Z_i, Z_j)W_i^*W_j^*$ and $nU_n^{(K)*} = \frac{1}{n-1}\sum_{i\neq j}\mathcal{H}^{(K)}(Z_i, Z_j)W_i^*W_j^*$. We first prove that for any $\epsilon > 0$,

$$\lim_{K\to+\infty} \limsup_{n\to\infty} \mathrm{pr}^*(|nU_n^* - nU_n^{(K)*}| > \epsilon) = 0 \tag{16}$$

almost surely. Consider the U-statistic

$$\frac{1}{n(n-1)} \sum_{i\neq j} \{\mathcal{H}(Z_i, Z_j) - \mathcal{H}^{(K)}(Z_i, Z_j)\}^2 = \frac{1}{n(n-1)} \sum_{i\neq j} \left\{ \sum_{k=K+1}^{\infty} \nu_k \phi_k(Z_i)\phi_k(Z_j) \right\}^2$$

with the kernel $\mathcal{H}^*(Z, Z') = \{\mathcal{H}(Z, Z') - \mathcal{H}^{(K)}(Z, Z')\}^2$. As $E\{\mathcal{H}^*(Z, Z')\} = E[\{\mathcal{H}(Z, Z') - \mathcal{H}^{(K)}(Z, Z')\}^2] = \sum_{k=K+1}^{\infty} \nu_k^2 < \infty$, by the strong law of large numbers for U-statistic [see Hoeffding (1961) and Lee (1990)], we obtain

$$\begin{aligned}
E^*[(nU_n^* - nU_n^{(K)*})^2] &= E^* \left[\frac{1}{(n-1)^2} \left\{ \sum_{i\neq j} \sum_{k=K+1}^{\infty} \nu_k \phi_k(Z_i)\phi_k(Z_j) W_i^* W_j^* \right\}^2 \right] \\
&= \frac{1}{(n-1)^2} \sum_{i\neq j} \left\{ \sum_{k=K+1}^{\infty} \nu_k \phi_k(Z_i)\phi_k(Z_j) \right\}^2 \\
&\to^{a.s.} E\left[\left\{ \sum_{k=K+1}^{\infty} \nu_k \phi_k(Z)\phi_k(Z') \right\}^2 \right],
\end{aligned}$$

as $n \to +\infty$. Thus (16) follows from the Markov inequality and (15).

Next we show that for any fixed K,

$$nU_n^{(K)*} \to^{D^*} \sum_{k=1}^{K} \nu_k (N_k^2 - 1) \; a.s., \tag{17}$$

where $(N_k)_{k=1}^K$ iid $N(0, 1)$. First, we can rewrite $nU_n^{(K)*}$ as

$$nU_n^{(K)*} = \frac{1}{n} \sum_i \sum_j \left\{ \sum_{k=1}^{K} \nu_k \phi_k(Z_i)\phi_k(Z_j) W_i^* W_j^* \right\} - \frac{1}{n} \sum_i \sum_{k=1}^{K} \nu_k \{\phi_k(Z_i) W_i^*\}^2 \tag{18}$$

and for convenience, we let the denominator of $nU_n^{(K)*}$ be n instead of $n-1$. By Lemma 2, $E\{\phi_k(Z_i)^4\} < \infty$ which implies that $E\{\sum_{i=1}^{+\infty} \phi_k(Z_i)^4/i^2\} < \infty$, where $\phi_k(\cdot)$ corresponds to $\nu_k \neq 0$. Define the set

$$\mathcal{A}_k := \left\{ w \in \Omega : \sum_{i=1}^{+\infty} \frac{\phi_k(Z_i(w))^4}{i^2} < \infty \text{ and } \frac{1}{n} \sum_{i=1}^{n} \phi_k(Z_i(w))^b \to E\{\phi_k(Z_i)^b\} \text{ for } b = 2, 4 \right\}.$$

Then $\mathrm{pr}(\cap_{k=1}^{(K)} \mathcal{A}_k) = 1$, where $\cap_{k=1}^{(K)}$ is the intersection of indices where eigenvalues $(\nu_k)_{k=1}^K$ are nonzero. Conditional on $\{Z_i(w)\}$ with $w \in \cap_{k=1}^{(K)} \mathcal{A}_k$, by Corollary 7.4.1 of Resnick (2005), we have

$$\frac{1}{n} \sum_{i=1}^{n} (W_i^{*2} - 1)\phi_k(Z_i)^2 \to^{a.s.} 0,$$

where $\phi_k(\cdot)$ corresponds to $\nu_k \neq 0$. As $\sum_{i=1}^n \phi_k(Z_i)^2/n \to 1$, we have $\frac{1}{n}\sum_{i=1}^n W_i^{*2}\phi_k(Z_i)^2 \to^{a.s.} 1$, which implies

$$\frac{1}{n}\sum_{i=1}^n \sum_{k=1}^K \nu_k \{\phi_k(Z_i) W_i^*\}^2 \to^{a.s.} \sum_{k=1}^K \nu_k.$$

On the other hand, note that the first term in (18) can be rewritten as

$$\sum_{k=1}^K \nu_k \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^n W_i^* \phi_k(Z_i) \right\}^2$$

and

$$\begin{aligned} \mathrm{cov}^* \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^n W_i^* \phi_s(Z_i), \frac{1}{n^{1/2}} \sum_{j=1}^n W_j^* \phi_t(Z_j) \right\} &= \frac{1}{n} \sum_{i=1}^n \phi_s(Z_i)\phi_t(Z_i) \\ &\to^{a.s.} E\{\phi_s(Z)\phi_t(Z)\} = \delta_{st}. \qquad (19) \end{aligned}$$

Similarly, define the set

$$\mathcal{B}_k := \left\{ w \in \Omega : \frac{1}{n} \max_{1 \le i \le n} \phi_k(Z_i(w))^2 \to 0 \right\}.$$

By Lemma 1 and $E\{\phi_k(Z)^2\} < \infty$ for $k = 1, 2, \cdots, K$, we have $\mathrm{pr}(\cap_{k=1}^K \mathcal{B}_k) = 1$ which implies that $\mathrm{pr}\{\cap_{k=1}^{(K)} (\mathcal{A}_k \cap \mathcal{B}_k)\} = 1$. Conditional on $\{Z_i(w)\}$ with $w \in \cap_{k=1}^{(K)} (\mathcal{A}_k \cap \mathcal{B}_k)$, we have

$$\frac{\max_{1\le i\le n} \mathrm{var}^*\{W_i^*\phi_k(Z_i)\}}{\sum_{j=1}^n \mathrm{var}^*\{W_j^*\phi_k(Z_j)\}} = \frac{\frac{1}{n}\max_{1\le i\le n}\phi_k(Z_i)^2}{\frac{1}{n}\sum_{j=1}^n \phi_k(Z_j)^2} \to 0. \qquad (20)$$

□

By Theorem D.19 in Greene (2007) and the Cramer-Wold device,

$$\left(\frac{1}{n^{1/2}} \sum_{i=1}^n W_i^* \phi_{(1)}(Z_i), \ldots, \frac{1}{n^{1/2}} \sum_{i=1}^n W_i^* \phi_{(K)}(Z_i) \right) \to^D N(0, I_{(K)}),$$

for almost every realization of $\{Z_i\}$, where $((1), \cdots, (K))$ are indices that correspond to nonzero eigenvalues $(\nu_k)_{k=1}^K$, $I_{(K)}$ is the $(K) \times (K)$ identity matrix. Hence, $nU_n^{(K)*} \to^{D^*} \sum_{k=1}^K \nu_k (N_k^2 - 1)\ a.s.$

Finally, since (16) and (17) are both satisfied, we can apply Theorem 2 in Dehling et al. (2009) to conclude that

$$nU_n^* \to^{D^*} \sum_{k=1}^\infty \nu_k (N_k^2 - 1)\ a.s.$$

Proof of Theorem 5. Recall that $J(Z_i, Z_j) = U(X_i, X_j)V(Y_i, Y_j)$ for $Z_i = (X_i, Y_i)$; see Theorem 1. We first show that

$$\begin{aligned} &\mathrm{var}^* \left[\frac{1}{(n-3)} \sum_{i \neq j} \{\widetilde{A}_{ij}\widetilde{B}_{ij} - J(Z_i, Z_j)\}\eta_i\eta_j \right] \\ =&\frac{1}{(n-3)^2} \sum_{i\neq j} \{\widetilde{A}_{ij}\widetilde{B}_{ij} - J(Z_i, Z_j)\}^2 \to^{a.s.} 0. \end{aligned}$$

For the ease of notation, write $U_{ij} = U(X_i, X_j)$ and $V_{ij} = V(Y_i, Y_j)$. Notice that

$$\begin{aligned}
\sum_{i\neq j}(\widetilde{A}_{ij}\widetilde{B}_{ij} - U_{ij}V_{ij})^2 =& \sum_{i\neq j}(\widetilde{A}_{ij}\widetilde{B}_{ij} - U_{ij}\widetilde{B}_{ij} + U_{ij}\widetilde{B}_{ij} - U_{ij}V_{ij})^2 \\
\leq& 2\sum_{i\neq j}(\widetilde{A}_{ij} - U_{ij})^2\widetilde{B}_{ij}^2 + 2\sum_{i\neq j}U_{ij}^2(\widetilde{B}_{ij} - V_{ij})^2 \\
\leq& 4\sum_{i\neq j}(\widetilde{A}_{ij} - U_{ij})^2(\widetilde{B}_{ij} - V_{ij})^2 + 2\sum_{i\neq j}U_{ij}^2(\widetilde{B}_{ij} - V_{ij})^2 \\
&+ 4\sum_{i\neq j}(\widetilde{A}_{ij} - U_{ij})^2 V_{ij}^2 \\
\leq& 4\left\{\sum_{i\neq j}(\widetilde{A}_{ij} - U_{ij})^4\right\}^{1/2}\left\{\sum_{i\neq j}(\widetilde{B}_{ij} - V_{ij})^4\right\}^{1/2} \\
&+ 2\left(\sum_{i\neq j}U_{ij}^4\right)^{1/2}\left\{\sum_{i\neq j}(\widetilde{B}_{ij} - V_{ij})^4\right\}^{1/2} \\
&+ 4\left(\sum_{i\neq j}V_{ij}^4\right)^{1/2}\left\{\sum_{i\neq j}(\widetilde{A}_{ij} - U_{ij})^4\right\}^{1/2}.
\end{aligned}$$

Under the assumption $E[|Y|^8 + |X|^4] < \infty$, we have

$$\frac{1}{n^2}\sum_{i\neq j}U_{ij}^4 \to^{a.s.} E(U_{12}^4), \quad \frac{1}{n^2}\sum_{i\neq j}V_{ij}^4 \to^{a.s.} E(V_{12}^4).$$

Thus we only need to show that

$$\frac{1}{n^2}\sum_{i\neq j}(\widetilde{A}_{ij} - U_{ij})^4 \to^{a.s.} 0, \tag{21}$$

$$\frac{1}{n^2}\sum_{i\neq j}(\widetilde{B}_{ij} - V_{ij})^4 \to^{a.s.} 0. \tag{22}$$

We only prove (22) as the proof for the other one is similar. Some algebra shows that

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i\neq j}(\widetilde{B}_{ij} - V_{ij})^4 \\
\leq& \frac{C}{n^2}\sum_{i\neq j}\left[\left\{\frac{1}{n}\sum_{l=1}^{n}(b_{il} - E(b_{il} \mid Y_i))\right\}^4 + \left\{\frac{1}{n^2}\sum_{k,l=1}^{n}(b_{kl} - E(b_{kl}))\right\}^4\right] + o_{a.s.}(1),
\end{aligned}$$

for some constant C. Under the assumption $E(|Y|^8) < \infty$, by the strong law of large numbers for V-statistics, we have

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{n}\sum_{l=1}^{n}\{b_{il}-E(b_{il}\mid Y_i)\}\right]^4\\ =&\frac{1}{n^5}\sum_{i=1}^{n}\sum_{l_1,l_2,l_3,l_4=1}^{n}\prod_{k=1}^{4}\{b_{il_k}-E(b_{il_k}\mid Y_i)\}\rightarrow^{a.s.}0\end{aligned}$$

due to the fact that $E[\prod_{k=1}^{4}\{b_{il_k}-E(b_{il_k}\mid Y_i)\}]=0$, when (l_1,l_2,l_3,l_4) are distinct indices. Similarly,

$$\frac{1}{n^2}\sum_{i\neq j}\left[\frac{1}{n^2}\sum_{k,l=1}^{n}\{b_{kl}-E(b_{kl})\}\right]^4\leq\left[\frac{1}{n^2}\sum_{k,l=1}^{n}\{b_{kl}-E(b_{kl})\}\right]^4\rightarrow^{a.s.}0.$$

Therefore, we have

$$\frac{1}{n-3}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}_{ij}\eta_i\eta_j=\frac{1}{n-3}\sum_{i\neq j}J(Z_i,Z_j)\eta_i\eta_j+o_p^*(1)\ a.s.,\tag{23}$$

and the conclusion follows from Theorem 4. $\square$

Remark 1. Since (23) is shown only with the assumption $E(|Y|^8+|X|^4)<\infty$, (23) is valid under the local alternative with the assumption $E\{|\epsilon|^8+|g(X)|^8+|X|^4\}<\infty$ and under the fixed alternative with the assumption $E(|Y|^8+|X|^4)<\infty$.

Proof of Theorem 6. Under the local alternative and the assumption $E\{|\epsilon|^8+|g(X)|^8+|X|^4\}<\infty$, (23) remains valid and we further show that

$$\frac{1}{n-3}\sum_{i\neq j}\widetilde{A}_{ij}\widetilde{B}_{ij}\eta_i\eta_j=\frac{1}{n-3}\sum_{i\neq j}U(X_i,X_j)V(\epsilon_i,\epsilon_j)\eta_i\eta_j+o_p^*(1)\ a.s.$$

Then we are left to show

$$\frac{1}{n-3}\sum_{i\neq j}J(Z_i,Z_j)\eta_i\eta_j=\frac{1}{n-3}\sum_{i\neq j}U(X_i,X_j)V(\epsilon_i,\epsilon_j)\eta_i\eta_j+o_p^*(1)\ a.s.$$

Similar to the proof of Theorem 5, let us consider

$$
\begin{aligned}
& \operatorname{var}^* \left[\frac{1}{n-3} \sum_{i \neq j} \{ J(Z_i, Z_j) - U(X_i, X_j) V(\epsilon_i, \epsilon_j) \} \eta_i \eta_j \right] \\
= \quad & \frac{1}{(n-3)^2} \sum_{i \neq j} U(X_i, X_j)^2 \\
\times \quad & \left[\frac{1}{n^{2a}} V(g(X_i), g(X_j)) + \frac{1}{n^a} \left\{ V(g(X_i)+\epsilon_i, g(X_j)+\epsilon_j) - V(g(X_i), g(X_j)) - V(\epsilon_i, \epsilon_j) \right\} \right]^2 \\
\leq \quad & O(n^{-2(1+2a)}) \left\{ \sum_{i \neq j} U(X_i, X_j)^4 \right\}^{1/2} \left\{ \sum_{i \neq j} V(g(X_i), g(X_j))^4 \right\}^{1/2} \\
+ \quad & O(n^{-2(1+a)}) \left\{ \sum_{i \neq j} U(X_i, X_j)^4 \right\}^{1/2} \left\{ \sum_{i \neq j} V(g(X_i)+\epsilon_i, g(X_j)+\epsilon_j)^4 \right\}^{1/2} \\
+ \quad & O(n^{-2(1+a)}) \left\{ \sum_{i \neq j} U(X_i, X_j)^4 \right\}^{1/2} \left\{ \sum_{i \neq j} V(g(X_i), g(X_j))^4 \right\}^{1/2} \\
+ \quad & O(n^{-2(1+a)}) \left\{ \sum_{i \neq j} U(X_i, X_j)^4 \right\}^{1/2} \left\{ \sum_{i \neq j} V(\epsilon_i, \epsilon_j)^4 \right\}^{1/2} \\
\longrightarrow^{a.s.} & \; 0.
\end{aligned} \tag{24}
$$

Here (24) is due to the fact that

$$
\begin{aligned}
& \frac{1}{n^2} \sum_{i \neq j} U(X_i, X_j)^4 \to^{a.s.} E\{U(X, X')^4\}, \\
& \frac{1}{n^2} \sum_{i \neq j} V(g(X_i)+\epsilon_i, g(X_j)+\epsilon_j)^4 \to^{a.s.} E\{V(g(X)+\epsilon, g(X')+\epsilon')^4\}, \\
& \frac{1}{n^2} \sum_{i \neq j} V(g(X_i), g(X_j))^4 \to^{a.s.} E\{V(g(X), g(X'))^4\}, \\
& \frac{1}{n^2} \sum_{i \neq j} V(\epsilon_i, \epsilon_j)^4 \to^{a.s.} E\{V(\epsilon, \epsilon')^4\},
\end{aligned}
$$

since $E\{|g(X)|^4 + |\epsilon|^4 + |X|^4\} < \infty$. Thus, under the local alternative, we have

$$
\frac{1}{n-3} \sum_{i \neq j} \widetilde{A}_{ij} \widetilde{B}_{ij} \eta_i \eta_j = \frac{1}{n-3} \sum_{i \neq j} U(X_i, X_j) V(\epsilon_i, \epsilon_j) \eta_i \eta_j + o_p^*(1) \; a.s.
$$

Applying Theorem 4 to $\frac{1}{n-3} \sum_{i \neq j} U(X_i, X_j) V(\epsilon_i, \epsilon_j) \eta_i \eta_j$, we have

$$
T_n^* \to^{D^*} \mathcal{G}_0 \; a.s.
$$

Similarly, by (23) and applying Theorem 4 to $\frac{1}{n-3}\sum_{i\neq j} U(X_i, X_j)V(Y_i, Y_j)\eta_i\eta_j$, under the fixed alternative and the same assumptions in Theorem 5, we have

$$T_n^* \to^{D^*} \widetilde{\mathcal{G}}_0 := \sum_{k=1}^{\infty} \widetilde{\lambda}_k(\widetilde{G}_k^2 - 1)\ a.s.,$$

where $(\widetilde{\lambda}_k)$ is a sequence of eigenvalues corresponding to orthonormal eigenfunctions of J under the fixed alternative and $(\widetilde{G}_k)$ is a sequence of iid $N(0,1)$ random variables. Note that under the fixed alternative, FMDD$(Y \mid X)$ is a fixed positive integer which implies that nFMDD$_n(Y \mid X) \to^{a.s.} +\infty$.

Furthermore, under the local and fixed alternatives, we can show that

$$Q^*_{(1-\alpha),n} \to^p Q_{(1-\alpha),\mathcal{G}_0} \text{ and } Q^*_{(1-\alpha),n} \to^p Q_{(1-\alpha),\widetilde{\mathcal{G}}_0}, \tag{25}$$

□

respectively, where $Q_{(1-\alpha),\widetilde{\mathcal{G}}_0}$ is the $(1-\alpha)$th quantile of $\widetilde{\mathcal{G}}_0$. Here (25) is shown by using (ii) of Lemma 11.2.1 in Lehmann and Romano (2005) and the fact that $\mathcal{G}_0,\ \widetilde{\mathcal{G}}_0$ are continuous random variables which can be shown under $\sum_{k=1}\lambda_k^2 \neq 0$, $\sum_{k=1}\widetilde{\lambda}_k^2 \neq 0$ and these are implied by the assumptions in $H_{1,n}$, H_1. Finally, the conclusions follow from Theorems 2 and 3.

1.2. *Additional Simulation Examples*

Example 1.
In this example, we generate the functional response Y by a quadratic form of the covariate X which is also considered in Patilea et al. (2016).

$$Y_i(t) = c \cdot \{X_i(t)^2 - 1\} + \epsilon_i(t),$$

where X_i and ϵ_i are independent Brownian motion and Brownian bridge and $c = 0, 0.5$. Furthermore, other settings including user-chosen parameters for the existing two tests are the same as Example 4.

Table 1 reports the empirical sizes and powers for three tests. By comparison, our FMDD-based test appears to outperform the other two tests in terms of a higher empirical power especially when $n = 40$ except for $\alpha = 1\%$. Moreover, the KMSZ test appears inferior to the PSS and the FMDD-based counterparts with respect to the size and power, presumably due to its inability of capturing the nonlinear dependence between Y and X.

Following the suggestion of a referee, we also consider a model with high frequency effect below.

Example 2.

$$X_i(t) = \frac{4}{\pi}\sum_{k=1,3,\ldots,21} Z_{i,k}\sin(2\pi kt),$$
$$Y_i(t) = \frac{4}{\pi}\sum_{k=3,5,7,9} Z_{i,k}^2\sin(2\pi kt) + 4\epsilon_i(t),$$

where $Z_{i,k}$ for $k = 1, \ldots, 21$ are iid $N(0,1)$ random variables and $\epsilon_i(t)$ is a standard Brownian bridge on $[0,1]$.

From Table 2, we observe that FMDD and PSS deliver reasonable power when $n = 100$ due to their ability to detect the nonlinear dependence between X and Y. KMSZ suffers from some

Table 1: Percentage of rejections of the three tests for Example 1. The PSS, KMSZ, and FMDD refer to the test in Patilea et al. (2016), the test in Kokoszka et al. (2008), and the proposed test, respectively.

	$\alpha = 10\%$		$\alpha = 5\%$		$\alpha = 1\%$	
Size	$n = 40$	$n = 100$	$n = 40$	$n = 100$	$n = 40$	$n = 100$
FMDD	11.1	11.5	6.2	6.0	1.3	1.1
PSS	9.2	9.7	4.6	5.3	1.0	1.2
KMSZ	8.1	12.0	3.6	6.8	0.6	1.2
Power	$n = 40$	$n = 100$	$n = 40$	$n = 100$	$n = 40$	$n = 100$
FMDD	70.4	100.0	40.8	99.5	7.5	69.3
PSS	50.5	99.4	35.7	98.8	12.1	93.3
KMSZ	33.1	38.2	20.5	27.5	7.3	12.6

Table 2: Percentage of rejections of the three tests for Example 2. The PSS, KMSZ, and FMDD refer to the test in Patilea et al. (2016), the test in Kokoszka et al. (2008), and the proposed test, respectively.

	$\alpha = 10\%$		$\alpha = 5\%$		$\alpha = 1\%$	
Power	$n = 40$	$n = 100$	$n = 40$	$n = 100$	$n = 40$	$n = 100$
FMDD	54.7	99.9	35.5	97.0	8.5	67.1
PSS	21.7	77.8	12.7	69.9	2.6	52.1
KMSZ	44.6	51.8	31.6	40.3	11.2	23.1

power loss as the linear model assumption is not satisfied. For $n = 40$, our test still maintains higher power except for the case of $\alpha = 1\%$.

REFERENCES

Dehling, H., Durieu, O., and Volny, D. (2009). New techniques for empirical processes of dependent data. *Stochastic Processes and their Applications*, **119**, 3699-3718.

Dunford, N. and Schwartz, J. T. (1963). *Linear Operators, Spectral Theory, Self Adjoint Operators in Hilbert Space, Part 2*. Wiley-Interscience.

Greene, W. (2007). *Econometric Analysis*. Pearson Education India.

Hoeffding, W. (1961). The strong law of large numbers for U-statistics. Inst. Statist. Univ. of North Carolina, Mimeo Report, No. 302.

Lee, A. J. (1990). *U-Statistics: Theory and Practice*. Marcel Dekker, Inc.

Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses*. 3rd edition, Springer, New York.

Lyons, R. (2013). Distance covariance in metric spaces. *The Annals of Probability*, **41**, 3284-3305.

Resnick, S. I. (2005). *A Probability Path*, Springer Science & Business Media.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons.

Zhang, X., Yao, S. and Shao, X. (2018). Conditional mean and quantile dependence testing in high dimension. *The Annals of Statistics*, 46(1), 219-246.