Center-Outward Ranks and Signs for Testing Conditional Quantile Independence

Kai Xu, Huijun Shi and Daojiang He

School of Mathematics and Statistics, Anhui Normal University

Supplementary Material

The supplementary material contains all technical proofs, and more numerical results on some aspects of limiting distributions and comparison under moderate dimension.

S1. Proof of Proposition 1

Since $\{(Y_i, \mathbf{x}_i), i = 1, \ldots, 4\}$ is a collection of independent and identically distributed random vectors, we obtain

$$E[\{\mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2 + \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2$$

$$- \mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 - \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2\} \mid \varepsilon_{i,\tau}] = 0, \quad \text{(S1.1)}$$

and $$E[\{\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_3)\| + \|\mathbf{G}_\pm(\mathbf{x}_2) - \mathbf{G}_\pm(\mathbf{x}_4)\|$$

$$-\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_4)\| - \|\mathbf{G}_\pm(\mathbf{x}_2) - \mathbf{G}_\pm(\mathbf{x}_3)\|\} \mid \mathbf{x}_i] = 0, \quad \text{(S1.2)}$$

for $i = 1, \ldots, 4$. The null of conditional quantile independence between $Q_\tau(Y \mid \mathbf{x})$ and $\mathbf{x}$ implies that $\psi_\tau(\varepsilon_\tau)$ is independent of $\mathbf{G}_\pm(\mathbf{x})$. It follows from the relations in (S1.1)-(S1.2) and the definition of h that

$$E(h[\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\varepsilon_{2,\tau}), \mathbf{G}_\pm(\mathbf{x}_2)\}, \{\psi_\tau(\varepsilon_{3,\tau}), \mathbf{G}_\pm(\mathbf{x}_3)\}, \{\psi_\tau(\varepsilon_{4,\tau}), \mathbf{G}_\pm(\mathbf{x}_4)\}] \mid \varepsilon_{i,\tau}, \mathbf{x}_i) = 0, \tag{S1.3}$$

$$E(h[\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\varepsilon_{3,\tau}), \mathbf{G}_\pm(\mathbf{x}_3)\}, \{\psi_\tau(\varepsilon_{2,\tau}), \mathbf{G}_\pm(\mathbf{x}_2)\}, \{\psi_\tau(\varepsilon_{4,\tau}), \mathbf{G}_\pm(\mathbf{x}_4)\}] \mid \varepsilon_{i,\tau}, \mathbf{x}_i) = 0, \text{and} \tag{S1.4}$$

$$E(h[\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\varepsilon_{4,\tau}), \mathbf{G}_\pm(\mathbf{x}_4)\}, \{\psi_\tau(\varepsilon_{2,\tau}), \mathbf{G}_\pm(\mathbf{x}_2)\}, \{\psi_\tau(\varepsilon_{3,\tau}), \mathbf{G}_\pm(\mathbf{x}_3)\}] \mid \varepsilon_{i,\tau}, \mathbf{x}_i) = 0, \tag{S1.5}$$

for $i = 1, \ldots, 4$. Due to the i.i.d.'ness of samples, it is not hard to see that

$$E[\{\mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2 + \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 - \mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 - \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2\} \mid \varepsilon_{1,\tau}, \varepsilon_{2,\tau}] = 0, \tag{S1.6}$$

$$E[\{\mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2 + \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 - \mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 - \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2\} \mid \varepsilon_{1,\tau}, \varepsilon_{3,\tau}] = -2\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{3,\tau}), \tag{S1.7}$$

$$E[\{\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_3)\| + \|\mathbf{G}_\pm(\mathbf{x}_2) - \mathbf{G}_\pm(\mathbf{x}_4)\| - \|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_4)\| - \|\mathbf{G}_\pm(\mathbf{x}_2) - \mathbf{G}_\pm(\mathbf{x}_3)\|\} \mid \mathbf{x}_1, \mathbf{x}_2] = 0, \quad \text{and} \tag{S1.8}$$

$$E[\{\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_3)\| + \|\mathbf{G}_\pm(\mathbf{x}_2) - \mathbf{G}_\pm(\mathbf{x}_4)\| - \|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_4)\| - \|\mathbf{G}_\pm(\mathbf{x}_2) - \mathbf{G}_\pm(\mathbf{x}_3)\|\} \mid \mathbf{x}_1, \mathbf{x}_3] = D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_3)\}, \tag{S1.9}$$

where the second assertion also applies the fact that $E\{\psi_\tau(\varepsilon_{2,\tau})\} = E\{\psi_\tau(\varepsilon_{4,\tau})\} = 0$. By using the relations in (S1.6), (S1.7), (S1.8) and (S1.9), the conditional quantile independence yields

$$E(h[\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\varepsilon_{2,\tau}), \mathbf{G}_\pm(\mathbf{x}_2)\}, \{\psi_\tau(\varepsilon_{3,\tau}), \mathbf{G}_\pm(\mathbf{x}_3)\},$$
$$\{\psi_\tau(\varepsilon_{4,\tau}), \mathbf{G}_\pm(\mathbf{x}_4)\}] \mid \varepsilon_{1,\tau}, \mathbf{x}_1, \varepsilon_{2,\tau}, \mathbf{x}_2) = 0, \tag{S1.10}$$

$$E(h[\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\varepsilon_{3,\tau}), \mathbf{G}_\pm(\mathbf{x}_3)\}, \{\psi_\tau(\varepsilon_{2,\tau}), \mathbf{G}_\pm(\mathbf{x}_2)\},$$
$$\{\psi_\tau(\varepsilon_{4,\tau}), \mathbf{G}_\pm(\mathbf{x}_4)\}] \mid \varepsilon_{1,\tau}, \mathbf{x}_1, \varepsilon_{2,\tau}, \mathbf{x}_2)$$
$$= -4^{-1}\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}, \text{and} \tag{S1.11}$$

$$E(h[\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\varepsilon_{4,\tau}), \mathbf{G}_\pm(\mathbf{x}_4)\}, \{\psi_\tau(\varepsilon_{2,\tau}), \mathbf{G}_\pm(\mathbf{x}_2)\},$$
$$\{\psi_\tau(\varepsilon_{3,\tau}), \mathbf{G}_\pm(\mathbf{x}_3)\}] \mid \varepsilon_{1,\tau}, \mathbf{x}_1, \varepsilon_{2,\tau}, \mathbf{x}_2)$$
$$= -4^{-1}\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}. \tag{S1.12}$$

Moreover, by the law of iterated expectations, the conditional variables ε_τ and $\mathbf{x}$ can be replaced by the transformed variables $\psi_\tau(\varepsilon_\tau)$ and $\mathbf{G}_\pm(\mathbf{x})$. Recall the definition of the symmetric kernel $\widetilde{h}[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4]$. By invoking (S1.3), (S1.4) and (S1.5), we are able to conclude that $\widetilde{h}_1\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\} = 3^{-1}(0+0+0) = 0$. A direct application of (S1.10), (S1.11) and (S1.12) yields that

$$\begin{aligned}
&\widetilde{h}_2[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^2] \\
= \;& 3^{-1}(0 - 4^{-1} - 4^{-1})\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\} \\
= \;& -6^{-1}\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}. \qquad \text{(S1.13)}
\end{aligned}$$

As $\mathbf{G}_{\pm}(\mathbf{x})$ is not a constant, it is immediate that

$$\text{var}(\widetilde{h}_2[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}_{i=1}^2]) = 36^{-1}\tau^2(1-\tau)^2\text{var}[D\{\mathbf{G}_{\pm}(\mathbf{x}_1), \mathbf{G}_{\pm}(\mathbf{x}_2)\}] > 0.$$

It follows (Serfling, 1980, Problem 5.P.3(i)) by independence of $Q_\tau(Y \mid \mathbf{x})$ and $\mathbf{x}$ that

$$\begin{aligned}
& \text{var}(\widetilde{h}_2[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}_{i=1}^2]) \\
\leq\; & \text{var}(\widetilde{h}_3[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}_{i=1}^3]) \leq \text{var}(\widetilde{h}[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}_{i=1}^4]) \\
\leq\; & 3^{-1}8^{-2}\text{var}[\{\mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2 + \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 \\
& - \mid \psi_\tau(\varepsilon_{1,\tau}) - \psi_\tau(\varepsilon_{4,\tau}) \mid^2 - \mid \psi_\tau(\varepsilon_{2,\tau}) - \psi_\tau(\varepsilon_{3,\tau}) \mid^2\}] \\
& \times\, \text{var}[\{\|\mathbf{G}_{\pm}(\mathbf{x}_1) - \mathbf{G}_{\pm}(\mathbf{x}_3)\| + \|\mathbf{G}_{\pm}(\mathbf{x}_2) - \mathbf{G}_{\pm}(\mathbf{x}_4)\| \\
& -\|\mathbf{G}_{\pm}(\mathbf{x}_1) - \mathbf{G}_{\pm}(\mathbf{x}_4)\| - \|\mathbf{G}_{\pm}(\mathbf{x}_2) - \mathbf{G}_{\pm}(\mathbf{x}_3)\|\}] \\
=\; & 3^{-1}4^{-1}\tau^2(1-\tau)^2 E[\{\|\mathbf{G}_{\pm}(\mathbf{x}_1) - \mathbf{G}_{\pm}(\mathbf{x}_3)\| + \|\mathbf{G}_{\pm}(\mathbf{x}_2) - \mathbf{G}_{\pm}(\mathbf{x}_4)\| \\
& -\|\mathbf{G}_{\pm}(\mathbf{x}_1) - \mathbf{G}_{\pm}(\mathbf{x}_4)\| - \|\mathbf{G}_{\pm}(\mathbf{x}_2) - \mathbf{G}_{\pm}(\mathbf{x}_3)\|\}]^2.
\end{aligned}$$

Using the relations $\|\mathbf{G}_{\pm}(\mathbf{x})\| = \|J\{\|\mathbf{F}_{\pm}(\mathbf{x})\|\}\{\mathbf{F}_{\pm}(\mathbf{x})/\|\mathbf{F}_{\pm}(\mathbf{x})\|\}I(\|\mathbf{F}_{\pm}(\mathbf{x})\| \neq 0)\| =\mid J\{\|\mathbf{F}_{\pm}(\mathbf{x})\|\} \mid I(\|\mathbf{F}_{\pm}(\mathbf{x})\| \neq 0)$, $\mid \|\mathbf{G}_{\pm}(\mathbf{x}_1) - \mathbf{G}_{\pm}(\mathbf{x}_3)\| - \|\mathbf{G}_{\pm}(\mathbf{x}_2) - \mathbf{G}_{\pm}(\mathbf{x}_3)\| \mid \leq \|\mathbf{G}_{\pm}(\mathbf{x}_1)\| + \|\mathbf{G}_{\pm}(\mathbf{x}_2)\|$ and $\mid \|\mathbf{G}_{\pm}(\mathbf{x}_2) - \mathbf{G}_{\pm}(\mathbf{x}_4)\| - \|\mathbf{G}_{\pm}(\mathbf{x}_1) - \mathbf{G}_{\pm}(\mathbf{x}_4)\| \mid \leq \|\mathbf{G}_{\pm}(\mathbf{x}_1)\| + \|\mathbf{G}_{\pm}(\mathbf{x}_2)\|$, we have $\text{var}(\widetilde{h}_c[\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}_{i=1}^c]), c =$

2, 3, 4 are bounded by

$$
\begin{aligned}
& 3^{-1}2\tau^2(1-\tau)^2[E \mid J\{\|\mathbf{F}_{\pm}(\mathbf{x}_1)\|\} \mid^2 + E \mid J\{\|\mathbf{F}_{\pm}(\mathbf{x}_2)\|\} \mid^2] \\
= \; & 3^{-1}4\tau^2(1-\tau)^2 E \mid J(\|\mathbf{w}_{1,q}\|) \mid^2 = 3^{-1}4\tau^2(1-\tau)^2 \int_0^1 J^2(x)dx < \infty.
\end{aligned}
$$

The first identity follows from the fact that $\mathbf{F}_{\pm}(\mathbf{x}_1)$ and $\mathbf{F}_{\pm}(\mathbf{x}_2)$ have the same distribution that of $\mathbf{w}_{1,q}$. The second identity holds trivially because the spherical uniform measure W_q is the product of the uniform measures on $[0, 1)$ and on $\mathcal{S}_{q-1}$, which leads to $\|\mathbf{w}_{1,q}\| \sim$ uniform(0, 1). The third identity follows by invoking the assumption that $\int_0^1 J^2(x)dx < \infty$.

An argument parallel to that of Shi et al. (2022, Proposition 3.1) shows that $\widehat{U}_{\tau}^{\natural}$ and $\widehat{V}_{\tau}^{\natural}$ equal to

$$
\begin{aligned}
\widehat{U}_{\tau}^{\natural} \; = \; & 24\{n(n-1)(n-2)(n-3)\}^{-1} \sum_{1\leq i_1<i_2<i_3<i_4\leq n} \widetilde{h}[\{\psi_{\tau}(\varepsilon_{i_1,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_1})\}, \\
& \{\psi_{\tau}(\varepsilon_{i_2,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_2})\}, \{\psi_{\tau}(\varepsilon_{i_3,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_3})\}, \{\psi_{\tau}(\varepsilon_{i_4,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_4})\}], \text{and} \\
\widehat{V}_{\tau}^{\natural} \; = \; & n^{-4} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n} \widetilde{h}[\{\psi_{\tau}(\varepsilon_{i_1,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_1})\}, \{\psi_{\tau}(\varepsilon_{i_2,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_2})\}, \\
& \{\psi_{\tau}(\varepsilon_{i_3,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_3})\}, \{\psi_{\tau}(\varepsilon_{i_4,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_{i_4})\}],
\end{aligned}
$$

respectively. Having verified that $\text{var}[\widetilde{h}_1\{\psi_{\tau}(\varepsilon_{1,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_1)\}] = 0$ and $0 < \text{var}(\widetilde{h}_c[\{\psi_{\tau}(\varepsilon_{i,\tau}),\, \mathbf{G}_{\pm}(\mathbf{x}_i)\}_{i=1}^{c}]) < \infty, c = 2, 3, 4$ previously, an application of Hájek's projection method (van der Vaart, 1998, Chapter 11), with Lemma 5.7.3 of Serfling (1980), yields

$$
n(\widehat{U}_{\tau}^{\natural} - \widehat{U}_{2,\tau}^{\natural}) = O_p(n^{-1/2}), \text{ and } n(\widehat{V}_{\tau}^{\natural} - \widehat{V}_{2,\tau}^{\natural}) = O_p(n^{-1/2}), \qquad \text{(S1.14)}
$$

where $\widehat{U}^{\natural}_{2,\tau} = 12\{n(n-1)\}^{-1} \sum_{1\le i_1<i_2\le n} \widetilde{h}_2[\{\psi_\tau(\varepsilon_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \{\psi_\tau(\varepsilon_{i_2,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_2})\}]$ and $\widehat{V}^{\natural}_{2,\tau} = 6n^{-2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \widetilde{h}_2[\{\psi_\tau(\varepsilon_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \{\psi_\tau(\varepsilon_{i_2,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_2})\}]$. By definition, the function $-D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}$ is symmetric, continuous, non-negative definite (Lyons, 2013, page 3291) and satisfies $E[-D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\} \mid \mathbf{x}_i] = 0, i = 1, 2$ and $\mathrm{var}[D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}] < \infty$. Using the Hilbert-Schmidt theorem (Simon, 2015a, Theorem 3.2.1, Example 3.1.15), $D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}$ admits the following eigenfunction expansion

$$D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\} = -\sum_{k=1}^{\infty} \lambda_k \phi_k\{\mathbf{G}_\pm(\mathbf{x}_1)\}\phi_k\{\mathbf{G}_\pm(\mathbf{x}_2)\}. \tag{S1.15}$$

The sequence of numbers $\lambda_k > 0, k \ge 1$ is the non-zero eigenvalues of the integral equation

$$E[D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}\phi_k\{\mathbf{G}_\pm(\mathbf{x}_1)\} \mid \mathbf{x}_2] = -\lambda_k \phi_k\{\mathbf{G}_\pm(\mathbf{x}_2)\},$$

and orthonormal eigenfunctions $\phi_k\{\mathbf{G}_\pm(\mathbf{x})\}, k \ge 1$ are such that

$$E[\phi_{k_1}\{\mathbf{G}_\pm(\mathbf{x})\}\phi_{k_2}\{\mathbf{G}_\pm(\mathbf{x})\}] = I(k_1 = k_2). \tag{S1.16}$$

Putting the two pieces (S1.13) and (S1.14) together and by Theorem 4.11.8

in Simon (2015b), we may write

$$
\begin{aligned}
(n-1)\widehat{U}^{\natural}_{2,\tau} &= 6\sum_{k=1}^{\infty} n^{-1} \sum_{1\le i_1\ne i_2\le n} \lambda_{k,\tau}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_1,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_1})\} \\
&\quad \phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_2,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_2})\}, \\
n\widehat{V}^{\natural}_{2,\tau} &= 6\sum_{k=1}^{\infty} n^{-1} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \lambda_{k,\tau}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_1,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_1})\} \\
&\quad \phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_2,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_2})\},
\end{aligned}
$$

where $\lambda_{k,\tau} = 6^{-1}\tau(1-\tau)\lambda_k$ and $\phi_{k,\tau}\{\psi_\tau(\varepsilon_\tau),\mathbf{G}_\pm(\mathbf{x})\} = \{\tau(1-\tau)\}^{-1/2}\psi_\tau(\varepsilon_\tau)$ $\phi_k\{\mathbf{G}_\pm(\mathbf{x})\}$ satisfying

$$
\begin{aligned}
& E[\phi_{k_1,\tau}\{\psi_\tau(\varepsilon_\tau),\mathbf{G}_\pm(\mathbf{x})\}\phi_{k_2,\tau}\{\psi_\tau(\varepsilon_\tau),\mathbf{G}_\pm(\mathbf{x})\}] \\
= & \{\tau(1-\tau)\}^{-1}\mathrm{var}\{\psi_\tau(\varepsilon_\tau)\}E[\phi_{k_1}\{\mathbf{G}_\pm(\mathbf{x})\}\phi_{k_2}\{\mathbf{G}_\pm(\mathbf{x})\}] \\
= & I(k_1 = k_2). \qquad \text{(S1.17)}
\end{aligned}
$$

The last line follows from the independence of $Q_\tau(Y \mid \mathbf{x})$ and $\mathbf{x}$, and the relation in (S1.16).

For each integer K, we define the truncated versions of $\widehat{U}^{\natural}_{2,\tau}$ and $\widehat{V}^{\natural}_{2,\tau}$

$$
\begin{aligned}
(n-1)\widehat{U}^{\natural}_{K,2,\tau} &= 6\sum_{k=1}^{K} n^{-1} \sum_{1\le i_1\ne i_2\le n} \lambda_{k,\tau}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_1,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_1})\} \\
&\quad \phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_2,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_2})\}, \text{ and} \\
n\widehat{V}^{\natural}_{K,2,\tau} &= 6\sum_{k=1}^{K} n^{-1} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \lambda_{k,\tau}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_1,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_1})\} \\
&\quad \phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_2,\tau}),\mathbf{G}_\pm(\mathbf{x}_{i_2})\}.
\end{aligned}
$$

Let $\imath = (-1)^{1/2}$ be the imaginary unit. For any $x \in \mathbb{R}$ and any $\delta > 0$, choose and fix K large enough that $\mid x \mid \{2\tau^2(1-\tau)^2 \sum_{k=K+1}^{\infty} \lambda_k^2\}^{1/2} < \delta/3$. Using similar arguments to those in the derivation of Serfling (1980, Theorem 5.5.2), we have

$$
\begin{aligned}
& \mid E[\exp\{\imath x(n-1)\widehat{U}^{\natural}_{2,\tau}\}] - E[\exp\{\imath x(n-1)\widehat{U}^{\natural}_{K,2,\tau}\}] \mid \\
\leq & \mid x \mid \{2\tau^2(1-\tau)^2 \sum_{k=K+1}^{\infty} \lambda_k^2\}^{1/2} < \delta/3, \text{and} \qquad \text{(S1.18)} \\
& \mid E[\exp\{\imath x 6 \sum_{k=1}^{\infty} \lambda_{k,\tau}(N_k^2-1)\}] - E[\exp\{\imath x 6 \sum_{k=1}^{K} \lambda_{k,\tau}(N_k^2-1)\}] \mid \\
\leq & \mid x \mid \{2\tau^2(1-\tau)^2 \sum_{k=K+1}^{\infty} \lambda_k^2\}^{1/2} < \delta/3, \qquad \text{(S1.19)}
\end{aligned}
$$

for all $n \geq 2$. By multivariate central limit theorem, for any fixed K and any $\delta > 0$,

$$
\mid E[\exp\{\imath x(n-1)\widehat{U}^{\natural}_{K,2,\tau}\}] - E[\exp\{\imath x 6 \sum_{k=1}^{K} \lambda_{k,\tau}(N_k^2-1)\}] \mid < \delta/3, \qquad \text{(S1.20)}
$$

for all n sufficiently large. Combining (S1.18), (S1.19) and (S1.20), we have, for any x and any $\delta > 0$, and for all n sufficiently large,

$$
\mid E[\exp\{\imath x(n-1)\widehat{U}^{\natural}_{2,\tau}\}] - E[\exp\{\imath x 6 \sum_{k=1}^{\infty} \lambda_{k,\tau}(N_k^2-1)\}] \mid < \delta. \qquad \text{(S1.21)}
$$

Apply Slutsky's theorem, (S1.14) and (S1.21) to yield the weak convergence of $n\widehat{U}^{\natural}_{\tau}$.

In addition, we observe that

$$
\begin{aligned}
n\widehat{V}^{\natural}_{2,\tau} &= (n-1)\widehat{U}^{\natural}_{2,\tau} + 6\sum_{k=1}^{\infty} n^{-1}\sum_{i=1}^{n}\lambda_{k,\tau}\phi^2_{k,\tau}\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}, \text{and} \\
n\widehat{V}^{\natural}_{K,2,\tau} &= (n-1)\widehat{U}^{\natural}_{K,2,\tau} + 6\sum_{k=1}^{K} n^{-1}\sum_{i=1}^{n}\lambda_{k,\tau}\phi^2_{k,\tau}\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}.
\end{aligned}
$$

Since $\phi_{k,\tau}\{\psi_\tau(\varepsilon_\tau), \mathbf{G}_{\pm}(\mathbf{x})\} = \{\tau(1-\tau)\}^{-1/2}\psi_\tau(\varepsilon_\tau)\phi_k\{\mathbf{G}_{\pm}(\mathbf{x})\}$, a straightforward application of (S1.16) and (S1.17) yields

$$
\begin{aligned}
& E(n\widehat{V}^{\natural}_{2,\tau} - n\widehat{V}^{\natural}_{K,2,\tau})^2 \\
\leq\ & 2E\{(n-1)\widehat{U}^{\natural}_{2,\tau} - (n-1)\widehat{U}^{\natural}_{K,2,\tau}\}^2 \\
& +2\tau^2(1-\tau)^2 E[\sum_{k=K+1}^{\infty} n^{-1}\sum_{i=1}^{n}\lambda_k\phi^2_{k,\tau}\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_{\pm}(\mathbf{x}_i)\}]^2 \\
=\ & 4(n-1)n^{-1}\tau^2(1-\tau)^2\sum_{k=K+1}^{\infty}\lambda_k^2 + 2(n-1)n^{-1}\tau^2(1-\tau)^2(\sum_{k=K+1}^{\infty}\lambda_k)^2 \\
& +2n^{-1}\sum_{k_1=K+1}^{\infty}\sum_{k_2=K+1}^{\infty}\lambda_{k_1}\lambda_{k_2}E\{\psi^4_\tau(\varepsilon_\tau)\}E[\phi^2_{k_1}\{\mathbf{G}_{\pm}(\mathbf{x})\}\phi^2_{k_2}\{\mathbf{G}_{\pm}(\mathbf{x})\}].
\end{aligned}
\tag{S1.22}
$$

By the definition of ψ_τ, we have $E\{\psi^4_\tau(\varepsilon_\tau)\} = \tau(1-\tau)\{\tau^3 + (1-\tau)^3\}$. The boundedness of $D\{\mathbf{G}_{\pm}(\cdot), \mathbf{G}_{\pm}(\cdot)\}$ and its orthogonal expansion in (S1.15) together imply that

$$
\begin{aligned}
& \sum_{k=1}^{\infty}\lambda_k^2 = E[D^2\{\mathbf{G}_{\pm}(\mathbf{x}_1), \mathbf{G}_{\pm}(\mathbf{x}_2)\}] < \infty, \sum_{k=1}^{\infty}\lambda_k = -E[D\{\mathbf{G}_{\pm}(\mathbf{x}), \mathbf{G}_{\pm}(\mathbf{x})\}] < \infty, \\
& \sum_{k_1=1}^{\infty}\sum_{k_2=1}^{\infty}\lambda_{k_1}\lambda_{k_2}E[\phi^2_{k_1}\{\mathbf{G}_{\pm}(\mathbf{x})\}\phi^2_{k_2}\{\mathbf{G}_{\pm}(\mathbf{x})\}] = E[D^2\{\mathbf{G}_{\pm}(\mathbf{x}), \mathbf{G}_{\pm}(\mathbf{x})\}] < \infty.
\end{aligned}
$$

Moreover, for all $n \geq 2$ and $K \geq 1$, we have $(n-1)n^{-1} \leq 1$, $n^{-1} \leq 1$ and

$$E(6\sum_{k=K+1}^{\infty}\lambda_{k,\tau}N_k^2)^2 = \tau^2(1-\tau)^2\{(\sum_{k=K+1}^{\infty}\lambda_k)^2 + 2\sum_{k=K+1}^{\infty}\lambda_k^2\} \leq 3\tau^2(1-\tau)^2(\sum_{k=K+1}^{\infty}\lambda_k)^2. \quad \text{(S1.23)}$$

For any $x \in \mathbb{R}$ and any $\delta > 0$, choose and fix K large enough that

$$| x | \{3\tau^2(1-\tau)^2(\sum_{k=K+1}^{\infty}\lambda_k)^2\}^{1/2} < \delta/3. \quad \text{(S1.24)}$$

By the inequality $| \exp(\imath z) - 1 | \leq | z |$ and using the foregoing considerations of (S1.22), (S1.23) and (S1.24), we have $| E\{\exp(\imath x n\widehat{V}_{2,\tau}^{\natural})\} - E\{\exp(\imath x n\widehat{V}_{K,2,\tau}^{\natural})\} | \leq | x | E^{1/2}(n\widehat{V}_{2,\tau}^{\natural} - n\widehat{V}_{K,2,\tau}^{\natural})^2 < \delta/3$, $| E\{\exp(\imath x 6\sum_{k=1}^{\infty}\lambda_{k,\tau}N_k^2)\} - E\{\exp(\imath x 6\sum_{k=1}^{K}\lambda_{k,\tau}N_k^2)\} | \leq | x | E^{1/2}(6\sum_{k=K+1}^{\infty}\lambda_{k,\tau}N_k^2)^2 < \delta/3$, for all $n \geq 2$. By the continuous mapping theorem and the central limit theorem, the distribution of $n\widehat{V}_{K,2,\tau}^{\natural}$, which is fixed at K, converges to the distribution of $6\sum_{k=1}^{K}\lambda_{k,\tau}N_k^2$. This observation, combined with Lévy's theorem (van der Vaart, 1998, Theorem 2.13), yields that for any fixed K and any $\delta > 0$, $| E\{\exp(\imath x n\widehat{V}_{K,2,\tau}^{\natural})\} - E\{\exp(\imath x 6\sum_{k=1}^{K}\lambda_{k,\tau}N_k^2)\} | < \delta/3$, for all n sufficiently large. Thus, by the triangle inequality, we obtain for any x and any $\delta > 0$, $| E\{\exp(\imath x n\widehat{V}_{2,\tau}^{\natural})\} - E\{\exp(\imath x 6\sum_{k=1}^{\infty}\lambda_{k,\tau}N_k^2)\} | < \delta$, as $n \to \infty$. The relation (S1.14), together with Slutsky's theorem, gives that $n\widehat{V}_{\tau}^{\natural}$ converges in distribution to $\tau(1-\tau)\sum_{k=1}^{\infty}\lambda_k N_k^2$.

□

S2. Proof of Theorem 1

We begin by proving that $\widehat{U}_\tau$ and $\widehat{V}_\tau$ are asymptotically equivalent to their oracle versions, that is, $n\widehat{U}_\tau - n\widehat{U}_\tau^\natural = o_p(1)$ and $n\widehat{V}_\tau - n\widehat{V}_\tau^\natural = o_p(1)$. Write $\widehat{U}_\tau^\natural = \widehat{U}_\tau^\natural\{Q_\tau(Y)\}$ and $\widehat{V}_\tau^\natural = \widehat{V}_\tau^\natural\{Q_\tau(Y)\}$. According to the convention of Sherman (1994), we then say that $\widehat{U}_\tau^\natural$ and $\widehat{V}_\tau^\natural$ are respectively the MDD-based processes of the 4th-order U-type and V-type, indexed by $\theta = Q_\tau(Y)$. Moreover, the kernel associated with the two processes is $\widetilde{h}[\{\psi_\tau(Y_i-\theta), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4]$. By the boundness of the function $\psi_{\tau(\cdot)}$, we have

$$\begin{aligned} & \sup_\theta \mid \widetilde{h}[\{\psi_\tau(Y_i-\theta), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4] \mid \\ \leq \; & \|\mathbf{G}_\pm(\mathbf{x}_1)\| + \|\mathbf{G}_\pm(\mathbf{x}_2)\| + \|\mathbf{G}_\pm(\mathbf{x}_3)\| + \|\mathbf{G}_\pm(\mathbf{x}_4)\|. \end{aligned}$$

Therefore, $\widetilde{h}[\{\psi_\tau(Y_i-\theta), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4]$ has an envelope $\sum_{i=1}^4 \|\mathbf{G}_\pm(\mathbf{x}_i)\|$. Since the score function J satisfies $\int_0^1 J^2(x)dx < \infty$, it follows from the definition of $\mathbf{G}_\pm(\mathbf{x})$ that $E(\sum_{i=1}^4 \|\mathbf{G}_\pm(\mathbf{x}_i)\|)^2 \leq 16E(\|\mathbf{G}_\pm(\mathbf{x})\|^2) = 16\int_0^1 J^2(x)dx < \infty$. For $\delta_0 > 0$ sufficiently small, it is straightforward to show that

$$\begin{aligned} & \sup_{\theta_1,\theta_2 \in [Q_\tau(Y)-\delta_0, Q_\tau(Y)+\delta_0]} E \mid \widetilde{h}[\{\psi_\tau(Y_i-\theta_1), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4] \\ & \qquad\qquad -\widetilde{h}[\{\psi_\tau(Y_i-\theta_2), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4] \mid \\ \leq \; & (1+\tau)E[\{\sum_{i=1}^4 I(\theta_2 < Y_i \leq \theta_1) + \sum_{i=1}^4 I(\theta_1 < Y_i \leq \theta_2)\}(\sum_{i=1}^4 \|\mathbf{G}_\pm(\mathbf{x}_i)\|)] \\ \leq \; & 2^6(1+\tau)\delta_0\{\int_0^1 J^2(x)dx\}^{1/2}\{\sup_{y \in [Q_\tau(Y)-\delta_0, Q_\tau(Y)+\delta_0]} f_0(y)\} = O(\delta_0), \end{aligned}$$

where f_0 is the density function of Y. The first and the second inequalities follow due to that in a small neighborhood of $Q_\tau = Q_\tau(Y)$, the cumulative distribution function of Y is continuously differentiable, and $(Y \le Q_\tau)$ is independent of $\mathbf{x}$ under the conditional quantile independence. The main corollary 8 of Sherman (1994) indicates that

$$\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\} - \widehat{U}_\tau^\natural\{Q_\tau(Y)\} = o_p(n^{-1}), \qquad \text{(S2.25)}$$

when $\widetilde{h}[\{\psi_\tau(Y_i - \theta), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4]$ with $\theta = Q_\tau(Y)$ is degenerate under the conditional quantile independence.

According to Section 5.7.3 of Serfling (1980), $\widehat{V}_\tau^\natural\{Q_\tau(Y)\}$ can be decomposed as

$$\begin{aligned} \widehat{V}_\tau^\natural\{Q_\tau(Y)\} &= (1-n^{-1})(1-2n^{-1})(1-3n^{-1})\widehat{U}_\tau^\natural\{Q_\tau(Y)\} \\ &\quad + n^{-1}(6-11n^{-1}+6n^{-2})\widehat{R}_\tau^\natural\{Q_\tau(Y)\}, \qquad \text{(S2.26)} \end{aligned}$$

where $\widehat{R}_\tau^\natural\{Q_\tau(Y)\}$ is the average of all terms $\widetilde{h}[\{\psi_\tau(Y_{i_1}-\theta), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \{\psi_\tau(Y_{i_2}-\theta), \mathbf{G}_\pm(\mathbf{x}_{i_2})\}, \{\psi_\tau(Y_{i_3}-\theta), \mathbf{G}_\pm(\mathbf{x}_{i_3})\}, \{\psi_\tau(Y_{i_4}-\theta), \mathbf{G}_\pm(\mathbf{x}_{i_4})\}]$ with $\theta = Q_\tau(Y)$ and at least one equality $i_a = i_b, a \neq b$. Apply the main corollary of Sherman (1994) with $d = 1$ and Jensen's inequality to obtain that for any $0 < \varpi < 1$,

$$\begin{aligned} & E[\sup_{\theta \in [Q_\tau(Y)-\delta_0, Q_\tau(Y)+\delta_0]} \mid \widehat{U}_\tau^\natural(\theta) - \widehat{U}_\tau^\natural\{Q_\tau(Y)\} \mid] \\ = \;& O\{n^{-1}(n^{-1/2}+\delta_0)^{\varpi/2}\}. \qquad \text{(S2.27)} \end{aligned}$$

By a similar device, we have

$$
\begin{aligned}
& E[\sup_{\theta\in[Q_\tau(Y)-\delta_0,Q_\tau(Y)+\delta_0]} \mid \widehat{R}_\tau^\natural(\theta) - \widehat{R}_\tau^\natural\{Q_\tau(Y)\} \mid] \\
= \; & O\{n^{-1/2}(n^{-1/2}+\delta_0)^{\varpi/2}\}, \qquad \text{(S2.28)}
\end{aligned}
$$

because $\widehat{R}_\tau^\natural\{Q_\tau(Y)\}$ can be written as a nondegenerate U-process plus negligible terms. It is noted that $(1-n^{-1})(1-2n^{-1})(1-3n^{-1}) = 1 + O(n^{-1})$ and $n^{-1}(6-11n^{-1}+6n^{-2}) = 6n^{-1} + O(n^{-2})$. A further application of (S2.26), (S2.27) and (S2.28), with Markov's inequality, yields

$$
\begin{aligned}
& \widehat{V}_\tau^\natural\{\widehat{Q}_\tau(Y)\} - \widehat{V}_\tau^\natural\{Q_\tau(Y)\} \\
= \; & O_p\{n^{-1}(n^{-1/2}+\delta_0)^{\varpi/2}\} + O_p\{n^{-3/2}(n^{-1/2}+\delta_0)^{\varpi/2}\} \\
= \; & o_p(n^{-1}). \qquad \text{(S2.29)}
\end{aligned}
$$

By combining the equations in (S2.25) and (S2.29) with Slutsky's theorem, it thus suffices to show that $n\widehat{U}_\tau - n\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\} = o_p(1)$ and $n\widehat{V}_\tau - n\widehat{V}_\tau^\natural\{\widehat{Q}_\tau(Y)\} = o_p(1)$.

We first deal with $n\widehat{U}_\tau - n\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\}$. Let $C(n,d)$ denote the number of all combinations of d distinct elements from $\{1,\ldots,n\}$. For an arbitrary event $\mathcal{A}$, we define

$$\begin{aligned}
\widehat{U}_\tau(\mathcal{A}) &= \sum_{c=1}^{4} \widehat{H}_{c,\tau}(\mathcal{A}) \\
&\overset{\text{def}}{=} \sum_{c=1}^{4} C(4,c)C(n,c)^{-1} \sum_{1\le i_1<\cdots<i_c\le n} \\
&\quad \widetilde{h}_c^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\widehat{\varepsilon}_{i_c,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_c})\}, \mathcal{A}],
\end{aligned}$$

and

$$\begin{aligned}
\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y), \mathcal{A}\} &= \sum_{c=1}^{4} \widehat{H}_{c,\tau}^\natural(\mathcal{A}) \\
&\overset{\text{def}}{=} \sum_{c=1}^{4} C(4,c)C(n,c)^{-1} \sum_{1\le i_1<\cdots<i_c\le n} \\
&\quad \widetilde{h}_c^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\widehat{\varepsilon}_{i_c,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_c})\}, \mathcal{A}].
\end{aligned}$$

The notations $\widehat{H}_{c,\tau}$ and $\widehat{H}_{c,\tau}^\natural$ are denoted in an obvious way. Moreover, we write

$$\begin{aligned}
&\widetilde{h}_c^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^c, \mathcal{A}] = \widetilde{h}_c[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^c, \mathcal{A}] \\
&\quad - \sum_{j=1}^{c-1} \sum_{1\le i_1<\cdots<i_j\le c} \widetilde{h}_j^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\widehat{\varepsilon}_{i_j,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_j})\}, \mathcal{A}], \\
&\widetilde{h}_c^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^c, \mathcal{A}] = \widetilde{h}_c[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^c, \mathcal{A}] \\
&\quad - \sum_{j=1}^{c-1} \sum_{1\le i_1<\cdots<i_j\le c} \widetilde{h}_j^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\widehat{\varepsilon}_{i_j,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_j})\}, \mathcal{A}],
\end{aligned}$$

$\widetilde{h}_c[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^c, \mathcal{A}] = E[I(\mathcal{A})\widetilde{h}[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^4] \mid \{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^c]$ and $\widetilde{h}_c[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^c, \mathcal{A}] = E[I(\mathcal{A})\widetilde{h}[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^4] \mid \{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^c]$, for $c = 1, \ldots, 4$. When $\mathcal{A}$ is a certain event, $\widehat{U}_\tau(\mathcal{A})$ and $\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y), \mathcal{A}\}$ reduce to $\widehat{U}_\tau$ and $\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\}$.

The following proof is divided into three steps. The first step shows that $n\widehat{H}_{1,\tau} = o_p(1)$ and $n\widehat{H}^{\natural}_{1,\tau} = o_p(1)$, the second step that $n\widehat{H}_{2,\tau} - n\widehat{H}^{\natural}_{2,\tau} = o_p(1)$. The third step verifies that $n\widehat{H}_{c,\tau}$ and $n\widehat{H}^{\natural}_{c,\tau}, c = 3, 4$ all are $o_p(1)$ terms.

Step I. Let $\mathcal{A}_\delta = \{\widehat{Q}_\tau(Y) - Q_\tau(Y) > \delta\} \cup \{\widehat{Q}_\tau(Y) - Q_\tau(Y) \le -\delta\}$. For any $\eta > 0$, choose $0 < \delta < \delta_0$ such that $\mathrm{pr}(\mathcal{A}_\delta) \le \eta/2$ for large enough n. As claimed in (32) of Yao et al. (2018, Appendix),

$$\begin{aligned} & E[I\{Y_{i_1} \le C_1, \dots, Y_{i_j} \le C_j, -\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \le \delta\} \mid \mathbf{x}_1, \dots, \mathbf{x}_n] \\ = \; & E[I\{Y_{i_1} \le C_1, \dots, Y_{i_j} \le C_j, -\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \le \delta\}], \qquad \text{(S2.30)} \end{aligned}$$

for small enough δ, $C_1, \dots, C_k = \widehat{Q}_\tau(Y)$ or $Q_\tau(Y)$, and $j = 1, \dots, 4$. Clearly, $\{\psi_\tau(\widehat{\varepsilon}_{1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\}, \dots, \{\psi_\tau(\widehat{\varepsilon}_{4,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\}$ are dependent but identically distributed random vectors. By symmetry, we obtain

$$\begin{aligned} & E[\{\mid \psi_\tau(\widehat{\varepsilon}_{1,\tau}) - \psi_\tau(\widehat{\varepsilon}_{3,\tau}) \mid^2 + \mid \psi_\tau(\widehat{\varepsilon}_{2,\tau}) - \psi_\tau(\widehat{\varepsilon}_{4,\tau}) \mid^2 \\ & - \mid \psi_\tau(\widehat{\varepsilon}_{1,\tau}) - \psi_\tau(\widehat{\varepsilon}_{4,\tau}) \mid^2 - \mid \psi_\tau(\widehat{\varepsilon}_{2,\tau}) - \psi_\tau(\widehat{\varepsilon}_{3,\tau}) \mid^2\} \\ & I\{-\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \le \delta\} \mid \psi_\tau(\widehat{\varepsilon}_{i,\tau})] = 0, \text{ and} \qquad \text{(S2.31)} \\ & E[\{\|\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\| + \|\widehat{\mathbf{G}}_\pm(\mathbf{x}_2) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\| - \|\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\| \\ & -\|\widehat{\mathbf{G}}_\pm(\mathbf{x}_2) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\|\} \mid \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)] = 0, \qquad \text{(S2.32)} \end{aligned}$$

for any $\delta > 0$ and $i = 1, \dots, 4$. It follows from the relations in (S2.30), (S2.31) and (S2.32) and the law of iterated expectations that

$$E\{h[\{\psi_\tau(\widehat{\varepsilon}_{1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\widehat{\varepsilon}_{2,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\},$$
$$\{\psi_\tau(\widehat{\varepsilon}_{3,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\}, \{\psi_\tau(\widehat{\varepsilon}_{4,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\}]$$
$$I\{-\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \leq \delta\} \mid \psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\} = 0, \quad \text{(S2.33)}$$
$$E\{h[\{\psi_\tau(\widehat{\varepsilon}_{1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\widehat{\varepsilon}_{3,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\},$$
$$\{\psi_\tau(\widehat{\varepsilon}_{2,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\}, \{\psi_\tau(\widehat{\varepsilon}_{4,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\}]$$
$$I\{-\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \leq \delta\} \mid \psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\} = 0, \text{and} \quad \text{(S2.34)}$$
$$E\{h[\{\psi_\tau(\widehat{\varepsilon}_{1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\}, \{\psi_\tau(\widehat{\varepsilon}_{4,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\},$$
$$\{\psi_\tau(\widehat{\varepsilon}_{2,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\}, \{\psi_\tau(\widehat{\varepsilon}_{3,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\}]$$
$$I\{-\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \leq \delta\} \mid \psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\} = 0, \quad \text{(S2.35)}$$

for small enough δ and $i = 1, \ldots, 4$. In view of the definition of $\widetilde{h}$, the relations in (S2.30), (S2.31) and (S2.32) confirm that for any $\eta > 0$,

$$\begin{aligned} \text{pr}(\mid n\widehat{H}_{1,\tau} \mid > \eta) &\leq \text{pr}(\mid n\widehat{H}_{1,\tau} \mid > \eta, \overline{\mathcal{A}_\delta}) + \text{pr}(\mathcal{A}_\delta) \\ &\leq \text{pr}\{\mid n\widehat{H}_{1,\tau}(\overline{\mathcal{A}_\delta}) \mid > \eta\} + \eta/2 \\ &\rightarrow \eta/2, \end{aligned} \quad \text{(S2.36)}$$

for small enough δ and as $n \to \infty$. The last line follows since conditional on $\{-\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \leq \delta\}$, $\widehat{H}_{1,\tau}(\overline{\mathcal{A}_\delta}) = 0$ for small enough δ. Combining (S1.2), (S2.30) and (S2.31), we have that (S2.30), (S2.31) and (S2.32) hold similarly with $\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}$'s replaced by $\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}$'s. As a result, it is immediately clear that for any $\eta > 0$,

$$\begin{aligned} \mathrm{pr}(\mid n\widehat{H}^{\natural}_{1,\tau} \mid > \eta) &\leq \mathrm{pr}\{\mid n\widehat{H}^{\natural}_{1,\tau}(\overline{\mathcal{A}_\delta}) \mid > \eta\} + \eta/2 \\ &\to \eta/2, \end{aligned} \tag{S2.37}$$

for small enough δ and as $n \to \infty$. Therefore, the asserted claims in Step I follow from (S2.36) and (S2.37).

Step II. By symmetry, (S2.30) and Zhang et al. (2018, Appendix 1.2), it is straightforward to verify that given the event $\{-\delta < \widehat{Q}_\tau(Y) - Q_\tau(Y) \leq \delta\}$ with δ small enough,

$$\begin{aligned} & \widetilde{h}^*_2[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^2, \overline{\mathcal{A}_\delta}] \\ = & -6^{-1}\psi_\tau(\widehat{\varepsilon}_{1,\tau})\psi_\tau(\widehat{\varepsilon}_{2,\tau})D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\}I(\overline{\mathcal{A}_\delta}) \\ & +o_p(n^{-1}), \end{aligned} \tag{S2.38}$$

$$\begin{aligned} & \widetilde{h}^*_2[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}_{i=1}^2, \overline{\mathcal{A}_\delta}] \\ = & -6^{-1}\psi_\tau(\widehat{\varepsilon}_{1,\tau})\psi_\tau(\widehat{\varepsilon}_{2,\tau})D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}I(\overline{\mathcal{A}_\delta}) \\ & +o_p(n^{-1}), \end{aligned} \tag{S2.39}$$

where o_p is uniform in i and $D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\} = \|\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\| - E\{\|\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\| \mid \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\} - E\{\|\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\| \mid \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\} + E\{\|\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) - \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\|\}$. An application of the triangle inequality and

Markov's inequality implies

$$\begin{aligned}\mathrm{pr}(\mid n\widehat{H}_{2,\tau}-n\widehat{H}^{\natural}_{2,\tau}\mid>\eta) &\le \mathrm{pr}(\mid n\widehat{H}_{2,\tau}-n\widehat{H}^{\natural}_{2,\tau}\mid>\eta,\overline{\mathcal{A}_\delta})+\mathrm{pr}(\mathcal{A}_\delta)\\ &\le \mathrm{pr}\{\mid n\widehat{H}_{2,\tau}(\overline{\mathcal{A}_\delta})-n\widehat{H}^{\natural}_{2,\tau}(\overline{\mathcal{A}_\delta})\mid>\eta\}+\eta/2\\ &\le \eta^{-2}E\mid n\widehat{H}_{2,\tau}(\overline{\mathcal{A}_\delta})-n\widehat{H}^{\natural}_{2,\tau}(\overline{\mathcal{A}_\delta})\mid^2+\eta/2.\end{aligned}$$

To prove that $n\widehat{H}_{2,\tau}-n\widehat{H}^{\natural}_{2,\tau}=o_p(1)$, by Slutsky's theorem, it suffices to show that

$$\begin{aligned}&E[(n-1)^{-1}\sum_{i_1\neq i_2}^{n}\psi_\tau(\widehat{\varepsilon}_{i_1,\tau})\psi_\tau(\widehat{\varepsilon}_{i_2,\tau})I(\overline{\mathcal{A}_\delta})D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}),\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2})\}\\ &-(n-1)^{-1}\sum_{i_1\neq i_2}^{n}\psi_\tau(\widehat{\varepsilon}_{i_1,\tau})\psi_\tau(\widehat{\varepsilon}_{i_2,\tau})I(\overline{\mathcal{A}_\delta})D\{\mathbf{G}_\pm(\mathbf{x}_{i_1}),\mathbf{G}_\pm(\mathbf{x}_{i_2})\}]^2\\ =\ &o(1).\end{aligned}\tag{S2.40}$$

Similar as in (A.7) of Shi et al. (2022), the left-hand side of (S2.40) is equal to

$$\begin{aligned}&2n(n-1)^{-1}A^{(n)}_{1,\tau}B^{(n)}_1+4n(n-2)(n-1)^{-1}A^{(n)}_{2,\tau}B^{(n)}_2\\ &+n(n-2)(n-3)(n-1)^{-1}A^{(n)}_{3,\tau}B^{(n)}_3,\end{aligned}\tag{S2.41}$$

where we write

$$
\begin{aligned}
A_{1,\tau}^{(n)} &= E\{\psi_\tau(\widehat{\varepsilon}_{1,\tau})\psi_\tau(\widehat{\varepsilon}_{2,\tau})I(\overline{\mathcal{A}_\delta})\}^2, \\
A_{2,\tau}^{(n)} &= E\{\psi_\tau^2(\widehat{\varepsilon}_{1,\tau})\psi_\tau(\widehat{\varepsilon}_{2,\tau})\psi_\tau(\widehat{\varepsilon}_{3,\tau})I(\overline{\mathcal{A}_\delta})\}, \\
A_{3,\tau}^{(n)} &= E\{\psi_\tau(\widehat{\varepsilon}_{1,\tau})\psi_\tau(\widehat{\varepsilon}_{2,\tau})\psi_\tau(\widehat{\varepsilon}_{3,\tau})\psi_\tau(\widehat{\varepsilon}_{4,\tau})I(\overline{\mathcal{A}_\delta})\}, \\
B_1^{(n)} &= E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\} - D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}]^2, \\
B_2^{(n)} &= E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\} - D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}] \\
&\quad [D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\} - D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_3)\}], \\
\text{and } B_3^{(n)} &= E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\} - D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}] \\
&\quad [D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_3), \widehat{\mathbf{G}}_\pm(\mathbf{x}_4)\} - D\{\mathbf{G}_\pm(\mathbf{x}_3), \mathbf{G}_\pm(\mathbf{x}_4)\}].
\end{aligned}
$$

The assertion that $\widehat{\varepsilon}_\tau \to \varepsilon_\tau$ and $I(\overline{\mathcal{A}_\delta}) \to 1$ almost surely follows from an application of Serfling (1980, Theorem 2.3.2). Since $\psi_\tau(\cdot)$ and $I(\cdot)$ are bounded, the dominated convergence theorem implies that

$$
\begin{aligned}
A_{1,\tau}^{(n)} &\to E\{\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})\}^2 = \tau^2(1-\tau)^2, && \text{(S2.42)} \\
A_{2,\tau}^{(n)} &\to E[\psi_\tau^2(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})\psi_\tau(\varepsilon_{3,\tau})] = 0, && \text{(S2.43)} \\
\text{and } nA_{3,\tau}^{(n)} &\to -3E\{\psi_\tau^2(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})\psi_\tau(\varepsilon_{3,\tau})\} = 0. && \text{(S2.44)}
\end{aligned}
$$

Under the assumption that the score J is square-integrable, we have, since

$E\{\mathbf{G}_\pm(\mathbf{x})\} = 0$, that

$$\begin{aligned}
& E[D^2\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}] \\
= \;& E\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_2)\|^2 + E^2\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_2)\| \\
& -2E\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_2)\|\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_3)\| \\
\leq \;& 2E\|\mathbf{G}_\pm(\mathbf{x}_1) - \mathbf{G}_\pm(\mathbf{x}_2)\|^2 \\
= \;& 4\int_0^1 J^2(x)dx < \infty. \qquad \text{(S2.45)}
\end{aligned}$$

By Hallin et al. (2021, Proposition 2.3), it follows that $\widehat{\mathbf{G}}_\pm(\mathbf{x}_1) \to \mathbf{G}_\pm(\mathbf{x}_1)$ almost surely. Using (S2.45) and Vitali's theorem (Shorack, 2017, Chapter 3, Theorem 5.5) yields the L_2-convergence relation $E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\} - D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}]^2 \to 0$. This, together with (S2.42) and the fact that $n(n-1)^{-1} = 1 + o(1)$, entails immediately that

$$n(n-1)^{-1}A_{1,\tau}^{(n)}B_1^{(n)} = \{1 + o(1)\}\tau^2(1-\tau)^2B_1^{(n)} \to 0. \qquad \text{(S2.46)}$$

Using similar arguments to those for dealing with (A.10)-(A.12) in Shi et al. (2022), we get

$$\begin{aligned}
B_2^{(n)} \;=\;& E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\}D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_3)\}] \\
& -2E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\}D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_3)\}] \\
=\;& -(n-2)^{-1}E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\}D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_2)\}] \\
& +2(n-2)^{-1}E[D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_1), \widehat{\mathbf{G}}_\pm(\mathbf{x}_1)\}D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\}]
\end{aligned}$$

$$+2(n-2)^{-1}E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_2)\}D\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_2)\}] \quad \text{(S2.47)}$$

$$-(n-2)^{-1}E[D^2\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_2)\}]$$

$$= n^{-1}\{1+o(1)\}E[D^2\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_2)\}]. \quad \text{(S2.48)}$$

Applying Vitali's theorem together with (S2.45) proves that

$$\begin{aligned}
B_3^{(n)} &= E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_2)\}D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_3),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_4)\}] \\
&\quad -2E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_2)\}D\{\mathbf{G}_{\pm}(\mathbf{x}_3),\mathbf{G}_{\pm}(\mathbf{x}_4)\}] \\
&= -(n-3)^{-1}E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1)\}D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_3),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_4)\}] \\
&\quad +2(n-3)^{-1}E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_3)\}D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_3),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_4)\}] \\
&\quad -(n-3)^{-1}E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1)\}D\{\mathbf{G}_{\pm}(\mathbf{x}_3),\mathbf{G}_{\pm}(\mathbf{x}_4)\}] \\
&\quad +2(n-3)^{-1}E[D\{\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_1),\widehat{\mathbf{G}}_{\pm}(\mathbf{x}_3)\}D\{\mathbf{G}_{\pm}(\mathbf{x}_3),\mathbf{G}_{\pm}(\mathbf{x}_4)\}] \\
&= -2n^{-1}\{1+o(1)\}E[D\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_1)\}D\{\mathbf{G}_{\pm}(\mathbf{x}_3),\mathbf{G}_{\pm}(\mathbf{x}_4)\}] \\
&\quad +4n^{-1}\{1+o(1)\}E[D\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_3)\}D\{\mathbf{G}_{\pm}(\mathbf{x}_3),\mathbf{G}_{\pm}(\mathbf{x}_4)\}] \\
&= o(n^{-1}). \qquad \text{(S2.49)}
\end{aligned}$$

In the last step we employ the useful property of the double centred distance, that is, $E[D\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_2)\} \mid \mathbf{x}_1] = E[D\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_2)\} \mid \mathbf{x}_2] = 0$. The use of (S2.43) and (S2.48) immediately yields

$$\begin{aligned}
&n(n-2)(n-1)^{-1}A_{2,\tau}^{(n)}B_2^{(n)} \\
\to\ &E[\psi_\tau^2(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})\psi_\tau(\varepsilon_{3,\tau})]E[D^2\{\mathbf{G}_{\pm}(\mathbf{x}_1),\mathbf{G}_{\pm}(\mathbf{x}_2)\}] = 0 \quad \text{(S2.50)}
\end{aligned}$$

It follows from (S2.44) and (S2.49) that

$$
\begin{aligned}
& n(n-2)(n-3)(n-1)^{-1} A_{3,\tau}^{(n)} B_3^{(n)} \\
= \; & -3\{1+o(1)\} E\{\psi_\tau^2(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau})\psi_\tau(\varepsilon_{3,\tau})\} n B_3^{(n)} = o(1). \quad \text{(S2.51)}
\end{aligned}
$$

The proof of (S2.40) is completed by plugging (S2.46), (S2.50) and (S2.51) into (S2.41).

Step III. Write

$$
\begin{aligned}
& \widetilde{h}_{31}^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^3, \overline{\mathcal{A}_\delta}] \\
= \; & -24^{-1} \sum_{(i_1,i_2,i_3)}^{3} \psi_\tau(\widehat{\varepsilon}_{i_1,\tau})\psi_\tau(\widehat{\varepsilon}_{i_2,\tau}) I(\overline{\mathcal{A}_\delta})[2D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2})\} \\
& -D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_3})\} - D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_3})\}] \\
& +3^{-1} \sum_{(i_1,i_2)}^{3} \psi_\tau(\widehat{\varepsilon}_{i_1,\tau})\psi_\tau(\widehat{\varepsilon}_{i_2,\tau}) I(\overline{\mathcal{A}_\delta}) D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2})\}, \text{and} \\
& \widetilde{h}_{41}^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^4, \overline{\mathcal{A}_\delta}] \\
= \; & -24^{-1} \sum_{(i_1,i_2,i_3,i_4)}^{4} \psi_\tau(\widehat{\varepsilon}_{i_1,\tau})\psi_\tau(\widehat{\varepsilon}_{i_2,\tau}) I(\overline{\mathcal{A}_\delta})[2D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2})\} \\
& +2D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_3}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_4})\} - D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_3})\} - D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_4})\} \\
& -D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_3})\} - D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_4})\}] - \sum_{(i_1,i_2,i_3)}^{4} \widetilde{h}_{31}^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \\
& \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1})\}, \{\psi_\tau(\widehat{\varepsilon}_{i_2,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2})\}, \{\psi_\tau(\widehat{\varepsilon}_{i_3,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_3})\}, \overline{\mathcal{A}_\delta}] \\
& +3^{-1} \sum_{(i_1,i_2)}^{4} \psi_\tau(\widehat{\varepsilon}_{i_1,\tau})\psi_\tau(\widehat{\varepsilon}_{i_2,\tau}) I(\overline{\mathcal{A}_\delta}) D\{\widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_2})\},
\end{aligned}
$$

where we denote summation over mutually different subscripts shown. Using similar arguments to those in the derivation of (S2.38) and (S2.39), we obtain $\widetilde{h}_3^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm\ (\mathbf{x}_i)\}_{i=1}^3, \overline{\mathcal{A}_\delta}] = \widetilde{h}_{31}^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^3, \overline{\mathcal{A}_\delta}] + o_p(n^{-1})$ and $\widetilde{h}_4^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^4, \overline{\mathcal{A}_\delta}] = \widetilde{h}_{41}^*[\{\psi_\tau(\widehat{\varepsilon}_{i,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_i)\}_{i=1}^4, \overline{\mathcal{A}_\delta}] + o_p(n^{-1})$. Along the same steps as the proofs of Step II and Shi et al. (2022, Theorem 4.2), we deduce that $\{(n-1)\cdots(n-c+1)\}^{-1} \sum_{(i_1,\ldots,i_c)}^{n} \widetilde{h}_{c1}^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\widehat{\varepsilon}_{i_c,\tau}), \widehat{\mathbf{G}}_\pm(\mathbf{x}_{i_c})\}, \overline{\mathcal{A}_\delta}]$ and $\{(n-1)\cdots(n-c+1)\}^{-1} \sum_{(i_1,\ldots,i_c)}^{n} \widetilde{h}_{c1}^*[\{\psi_\tau(\widehat{\varepsilon}_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\widehat{\varepsilon}_{i_c,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_c})\}, \overline{\mathcal{A}_\delta}], c = 3, 4$ all are $o_p(1)$ terms. Step III is completed.

By (S2.26), it is clear that

$$\begin{aligned} \widehat{V}_\tau - \widehat{V}_\tau^\natural\{Q_\tau(Y)\} &= (1-n^{-1})(1-2n^{-1})(1-3n^{-1})[\widehat{U}_\tau - \widehat{U}_\tau^\natural\{Q_\tau(Y)\}] \\ &\quad + n^{-1}(6 - 11n^{-1} + 6n^{-2})[\widehat{R}_\tau - \widehat{R}_\tau^\natural\{Q_\tau(Y)\}]. \end{aligned}$$

Employing arguments similar to those for dealing with $\widehat{U}_\tau - \widehat{U}_\tau^\natural\{Q_\tau(Y)\}$, we have $\widehat{R}_\tau - \widehat{R}_\tau^\natural\{Q_\tau(Y)\} = o_p(1)$. Notice that $(1-n^{-1})(1-2n^{-1})(1-3n^{-1}) = 1 + O(n^{-1})$ and $n^{-1}(6 - 11n^{-1} + 6n^{-2}) = 6n^{-1} + O(n^{-2})$. By the Cramér-Wold device and Slutsky's theorem, the proof is thus complete.

□

S3. Proof of Theorem 2

We can decompose

$$\begin{aligned} n^{1/2}[\widehat{U}_\tau - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}] &= n^{1/2}[\widehat{U}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}] \\ &\quad + n^{1/2}(\widehat{U}_\tau - \widehat{U}_\tau^\natural), \text{and} \\ n^{1/2}[\widehat{V}_\tau - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}] &= n^{1/2}[\widehat{V}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}] \\ &\quad + n^{1/2}(\widehat{V}_\tau - \widehat{V}_\tau^\natural). \end{aligned}$$

Having established $Eh^2[\{\psi_\tau(\varepsilon_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}, \ldots, \{\psi_\tau(\varepsilon_{i_4,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_4})\}] \le O(1)$ $\int_0^1 J^2(x)\, dx < \infty$ for all $1 \le i_1, \ldots, i_4 \le 4$, we apply Serfling (1980, Lemma 5.7.3) in the case $r = 2$ to yield $n^{1/2}(\widehat{U}_\tau^\natural - \widehat{V}_\tau^\natural) = o_p(1)$, in which case $n^{1/2}[\widehat{U}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}]$ and $n^{1/2}[\widehat{V}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}]$ have the same limit distribution. Invoking the Hoeffding decomposition in technical appendix 1.2 of Zhang et al. (2018), we have

$$n^{1/2}[\widehat{U}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}] = 4n^{-1/2} \sum_{i=1}^{n} \widetilde{h}_1\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\} + o_p(1),$$

under the fixed alternative $\mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\} > 0$, where $\widetilde{h}_1\{\psi_\tau(\varepsilon_{i,\tau}),$ $\mathbf{G}_\pm(\mathbf{x}_i)\}, i = 1, \ldots, n$ are independent and identically distributed and

$$\begin{aligned} & \widetilde{h}_1\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\} \\ = \; & -2^{-1} E[\psi_\tau(\varepsilon_{1,\tau})\psi_\tau(\varepsilon_{2,\tau}) D\{\mathbf{G}_\pm(\mathbf{x}_1), \mathbf{G}_\pm(\mathbf{x}_2)\} \mid \psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)] \\ & + 2^{-1}\mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}. \end{aligned}$$

By Slutsky's theorem and central limit theorem, it is immediate that as $n \to \infty$, both $n^{1/2}[\widehat{V}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}]/4$ and $n^{1/2}[\widehat{U}_\tau^\natural - \mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\}]/4$ converge in distribution to a normal distribution with mean zero and variance $\mathrm{var}[\widetilde{h}_1\{\psi_\tau(\varepsilon_{1,\tau}), \mathbf{G}_\pm(\mathbf{x}_1)\}]$. The same arguments employed in dealing with (S2.25) and (S2.29) immediately yield

$$\begin{aligned}
\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\} - \widehat{U}_\tau^\natural\{Q_\tau(Y)\} &= O\{n^{-1/2}(n^{-1/2}+\delta_0)^{\varpi/2}\} = o_p(n^{-1/2}), \\
\widehat{V}_\tau^\natural\{\widehat{Q}_\tau(Y)\} - \widehat{V}_\tau^\natural\{Q_\tau(Y)\} &= O\{n^{-1/2}(n^{-1/2}+\delta_0)^{\varpi/2}\} = o_p(n^{-1/2}),
\end{aligned}$$

uniformly over $\mid \widehat{Q}_\tau(Y) - Q_\tau(Y) \mid < \delta_0$ and $0 < \varpi < 1$, when $\mathrm{MDD}\{\psi_\tau(\varepsilon_\tau) \mid \mathbf{G}_\pm(\mathbf{x})\} \neq 0$. Using Vitali's theorem, the Hoeffding decomposition of $\widehat{U}_\tau - \widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\}$ and the connection between U- and V-statistics, it holds that $n^{1/2}\widehat{U}_\tau - n^{1/2}\widehat{U}_\tau^\natural\{\widehat{Q}_\tau(Y)\} = o_p(1)$ and $n^{1/2}\widehat{V}_\tau - n^{1/2}\widehat{V}_\tau^\natural\{\widehat{Q}_\tau(Y)\} = o_p(1)$. The theorem follows by putting the above together.

□

S4. Proof of Theorem 3

Define

$$\begin{aligned}
\widehat{\Lambda}_\tau &= \sum_{i=1}^n \log\{p_\tau(Y_i, \mathbf{x}_i \mid \theta_0 n^{-1/2})/p_\tau(Y_i, \mathbf{x}_i \mid 0)\}, \text{and} \\
\widehat{T}_\tau &= \theta_0 n^{-1/2} \sum_{i=1}^n \eta_\tau(Y_i, \mathbf{x}_i \mid 0).
\end{aligned}$$

By Lehmann and Romano (2005, Example 12.3.7), we clarify that $p_\tau(Y_i, \mathbf{x}_i \mid \theta_0 n^{-1/2})$ is contiguous to $p_\tau(Y_i, \mathbf{x}_i \mid 0)$ in order for Le Cam's third lemma van der Vaart (1998, Theorem 6.6) to be applicable. We proceed in two steps. First, we derive the joint limiting null distributions of $(n\widehat{U}_\tau, \widehat{\Lambda}_\tau)$ and $(n\widehat{V}_\tau, \widehat{\Lambda}_\tau)$ under the null hypothesis. Next, we employ Le Cam's third lemma to obtain their asymptotic distributions under contiguous alternatives.

Step I. In view of the proof of Theorem 1 and Proposition 1, we have $n\widehat{U}_\tau = n\widehat{V}^\natural_{2,\tau} - \tau(1-\tau)\sum_{k=1}^{\infty}\lambda_k + o_p(1)$ and $n\widehat{V}_\tau = n\widehat{V}^\natural_{2,\tau} + o_p(1)$, where $\widehat{V}^\natural_{2,\tau}$ is defined in (S1.14). By (S1.15), we write

$$\widehat{V}^\natural_{2,\tau} = \tau(1-\tau)n^{-2}\sum_{i_1,i_2=1}^{n}\sum_{k=1}^{\infty}\lambda_k\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_2,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_2})\},$$

where $\phi_{k,\tau}\{\psi_\tau(\varepsilon_\tau), \mathbf{G}_\pm(\mathbf{x})\} = \{\tau(1-\tau)\}^{-1/2}\psi_\tau(\varepsilon_\tau)\phi_k\{\mathbf{G}_\pm(\mathbf{x})\}$ satisfying (S1.17). By an application of Slutsky's theorem, it suffices to derive the limiting null distribution of $(n\widehat{V}^\natural_{2,\tau}, \widehat{\Lambda}_\tau)$. As in van der Vaart (1998, Theorem 7.2),

$$\widehat{\Lambda}_\tau - \widehat{T}_\tau + 2^{-1}\theta_0^2\mathcal{I}_\tau(0) \to 0, \tag{S4.52}$$

in probability. In order to apply Le Cam's third lemma, we therefore need to study the limiting joint distribution of $(n\widehat{V}^\natural_{2,\tau}, \widehat{T}_\tau)$.

For each positive integer K, consider the "truncated" V-statistic $\widehat{V}^\natural_{K,2,\tau} = \tau(1-\tau)n^{-2}\sum_{i_1,i_2=1}^{n}\sum_{k=1}^{K}\lambda_k\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_1,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_1})\}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i_2,\tau}), \mathbf{G}_\pm(\mathbf{x}_{i_2})\}$. Apparently, $n\widehat{V}_{2,\tau}^{\natural}$ and $n\widehat{V}_{K,2,\tau}^{\natural}$ can be written as $n\widehat{V}_{2,\tau}^{\natural} = \tau(1-\tau)\sum_{k=1}^{\infty}\lambda_k[n^{-1/2}\sum_{i=1}^{n}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}]^2$ and $n\widehat{V}_{K,2,\tau}^{\natural} = \tau(1-\tau)\sum_{k=1}^{K}\lambda_k[n^{-1/2}\sum_{i=1}^{n}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}]^2$. To obtain the limiting null distribution of $(n\widehat{V}_{2,\tau}^{\natural}, \widehat{T}_\tau)$, first consider the limiting null distribution, for fixed K, of $(n\widehat{V}_{K,2,\tau}^{\natural}, \widehat{T}_\tau)$. Let $\widehat{W}_{k,\tau}$ be a shorthand for $n^{-1/2}\sum_{i=1}^{n}\phi_{k,\tau}\{\psi_\tau(\varepsilon_{i,\tau}), \mathbf{G}_\pm(\mathbf{x}_i)\}$ and observe that

$$E(\widehat{W}_{k,\tau}) = E(\widehat{T}_\tau) = 0, \mathrm{var}(\widehat{W}_{k,\tau}) = 1,$$

$$\mathrm{var}(\widehat{T}_\tau) = \theta_0^2\mathcal{I}_\tau(0), \mathrm{cov}(\widehat{W}_{k_1,\tau}, \widehat{W}_{k_2,\tau}) = I(k_1 \neq k_2),$$

$$\text{and } \mathrm{cov}(\widehat{W}_{k,\tau}, \widehat{T}_\tau) = \theta_0 E[\phi_{k,\tau}\{\psi_\tau(\varepsilon_\tau), \mathbf{G}_\pm(\mathbf{x})\}\eta_\tau(Y, \mathbf{x} \mid 0)].$$

Since the score function $\eta_\tau(Y, \mathbf{x} \mid 0)$ is not additively separable, there exists at least one $k \geq 1$ such that $v_{k,\tau} = E[\phi_{k,\tau}\{\psi_\tau(\varepsilon_\tau), \mathbf{G}_\pm(\mathbf{x})\}\eta_\tau(Y, \mathbf{x} \mid 0)] \neq 0$. Applying the multivariate central limit theorem, we deduce

$$(\widehat{W}_{1,\tau}, \ldots, \widehat{W}_{K,\tau}, \widehat{T}_\tau)^{\mathrm{T}} \to (N_1, \ldots, N_K, T_\tau)^{\mathrm{T}},$$

in distribution, where $(N_1, \ldots, N_K, T_\tau)^{\mathrm{T}} \sim N_{K+1}\left(\mathbf{0}_{K+1}, \begin{pmatrix} \mathbf{I}_K & \theta_0\mathbf{v}_{K,\tau} \\ \theta_0\mathbf{v}_{K,\tau}^{\mathrm{T}} & \theta_0^2\mathcal{I}_\tau(0) \end{pmatrix}\right)$ and $\mathbf{v}_{K,\tau} = (v_{1,\tau}, \ldots, v_{K,\tau})^{\mathrm{T}}$. Then, by the continuous mapping theorem and Slutsky's theorem,

$$\begin{aligned}(n\widehat{V}_{K,2,\tau}^{\natural}, \widehat{T}_\tau)^{\mathrm{T}} \quad \to \quad & (\tau(1-\tau)\sum_{k=1}^{K}\lambda_k N_k^2, \theta_0\sum_{k=1}^{K}v_{k,\tau}N_k \\ & +\theta_0(\mathcal{I}_\tau(0) - \sum_{k=1}^{K}v_{k,\tau}^2)^{1/2}N_0)^{\mathrm{T}}, \qquad \text{(S4.53)}\end{aligned}$$

in distribution, where N_0 is standard Gaussian, independent of $N_1, \ldots, N_K$.

For any $x, y \in \mathbb{R}$, we have

$$\begin{aligned}
& \mid E\exp[\imath x \sum_{k=1}^{\infty} \lambda_k N_k^2 + \imath y\{\sum_{k=1}^{\infty} v_{k,\tau} N_k + (\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} v_{k,\tau}^2)^{1/2} N_0\}] \\
& -E\exp[\imath x \sum_{k=1}^{K} \lambda_k N_k^2 + \imath y\{\sum_{k=1}^{K} v_{k,\tau} N_k + (\mathcal{I}_\tau(0) - \sum_{k=1}^{K} v_{k,\tau}^2)^{1/2} N_0\}] \mid \\
\leq \quad & \mid x \mid E^{1/2}(\sum_{k=K+1}^{\infty} \lambda_k N_k^2)^2 + \mid y \mid E^{1/2}\{\sum_{k=K+1}^{\infty} v_{k,\tau} N_k \\
& +(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} v_{k,\tau}^2)^{1/2} N_0 - (\mathcal{I}_\tau(0) - \sum_{k=1}^{K} v_{k,\tau}^2)^{1/2} N_0\}^2 \\
= \quad & \mid x \mid \{2\sum_{k=K+1}^{\infty} \lambda_k^2 - (\sum_{k=K+1}^{\infty} \lambda_k)^2\}^{1/2} + \mid y \mid [\sum_{k=K+1}^{\infty} v_{k,\tau}^2 \\
& +\{(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} v_{k,\tau}^2)^{1/2} - (\mathcal{I}_\tau(0) - \sum_{k=1}^{K} v_{k,\tau}^2)^{1/2}\}^2]^{1/2} \\
\leq \quad & \mid x \mid \{2\sum_{k=K+1}^{\infty} \lambda_k^2 + (\sum_{k=K+1}^{\infty} \lambda_k)^2\}^{1/2} + \mid y \mid (2\sum_{k=K+1}^{\infty} v_{k,\tau}^2)^{1/2} \\
\to \quad & 0, \qquad \text{(S4.54)}
\end{aligned}$$

as $K \to \infty$. The last line follows from the fact that

$$\begin{aligned}
& \mid (\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} v_{k,\tau}^2)^{1/2} - (\mathcal{I}_\tau(0) - \sum_{k=1}^{K} v_{k,\tau}^2)^{1/2} \mid \leq (\sum_{k=K+1}^{\infty} v_{k,\tau}^2)^{1/2}, \\
& \sum_{k=1}^{\infty} \lambda_k^2 \leq 4\int_0^1 J^2(x)dx < \infty, \sum_{k=1}^{\infty} \lambda_k \leq 2\{\int_0^1 J^2(x)dx\}^{1/2} < \infty, \\
& \text{and} \sum_{k=1}^{\infty} v_{k,\tau}^2 \leq \mathcal{I}_\tau(0) < \infty.
\end{aligned}$$

A similar calculation shows that for any $x, y \in \mathbb{R}$,

$$
\begin{aligned}
& \mid E \exp(\imath x n \widehat{V}^{\natural}_{2,\tau} + \imath y \widehat{T}_\tau) - E \exp(\imath x n \widehat{V}^{\natural}_{K,2,\tau} + \imath y \widehat{T}_\tau) \mid \\
\leq \;& \mid x \mid \{4(n-1)n^{-1}\tau^2(1-\tau)^2 \sum_{k=K+1}^{\infty} \lambda_k^2 + 2(n-1)n^{-1}\tau^2(1-\tau)^2 (\sum_{k=K+1}^{\infty} \lambda_k)^2 \\
& +2n^{-1} \sum_{k_1=K+1}^{\infty} \sum_{k_2=K+1}^{\infty} \lambda_{k_1}\lambda_{k_2} E\{\psi_\tau^4(\varepsilon_\tau)\} E[\phi_{k_1}^2\{\mathbf{G}_\pm(\mathbf{x})\}\phi_{k_2}^2\{\mathbf{G}_\pm(\mathbf{x})\}]\} \\
\to \;& 0, \qquad \text{(S4.55)}
\end{aligned}
$$

as $K \to \infty$. Combining (S4.53), (S4.54) and (S4.55), we deduce that

$$
(n\widehat{V}^{\natural}_{2,\tau}, \widehat{T}_\tau)^{\mathrm{T}} \to (\tau(1-\tau)\sum_{k=1}^{\infty} \lambda_k N_k^2, \theta_0 \sum_{k=1}^{\infty} \upsilon_{k,\tau} N_k + \theta_0(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} \upsilon_{k,\tau}^2)^{1/2} N_0)^{\mathrm{T}},
$$

in distribution. The relation (S4.52) and Slutsky's theorem together imply that

$$
\begin{aligned}
(n\widehat{U}_\tau, \widehat{\Lambda}_\tau)^{\mathrm{T}} \;\to\;& (\tau(1-\tau)\sum_{k=1}^{\infty} \lambda_k (N_k^2 - 1), \theta_0 \sum_{k=1}^{\infty} \upsilon_{k,\tau} N_k \\
& +\theta_0(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} \upsilon_{k,\tau}^2)^{1/2} N_0 - 2^{-1}\theta_0^2 \mathcal{I}_\tau(0))^{\mathrm{T}}, \text{and} \\
(n\widehat{V}_\tau, \widehat{\Lambda}_\tau)^{\mathrm{T}} \;\to\;& (\tau(1-\tau)\sum_{k=1}^{\infty} \lambda_k N_k^2, \theta_0 \sum_{k=1}^{\infty} \upsilon_{k,\tau} N_k \\
& +\theta_0(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty} \upsilon_{k,\tau}^2)^{1/2} N_0 - 2^{-1}\theta_0^2 \mathcal{I}_\tau(0))^{\mathrm{T}},
\end{aligned}
$$

in distribution.

Step II. By appealing to Le Cam's third Lemma, under contiguous alternatives,

$$
\begin{aligned}
\mathrm{pr}(n\widehat{U}_\tau \le x) &\to E[I\{\tau(1-\tau)\sum_{k=1}^{\infty}\lambda_k(N_k^2-1) \le x\} \\
&\quad \exp\{\theta_0\sum_{k=1}^{\infty}\upsilon_{k,\tau}N_k + \theta_0(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty}\upsilon_{k,\tau}^2)^{1/2}N_0 - 2^{-1}\theta_0^2\mathcal{I}_\tau(0)\}], \\
&\stackrel{\mathrm{def}}{=} F_{1,\tau}(x), \text{and} \\
\mathrm{pr}(n\widehat{V}_\tau \le y) &\to E[I\{\tau(1-\tau)\sum_{k=1}^{\infty}\lambda_k N_k^2 \le y\} \\
&\quad \exp\{\theta_0\sum_{k=1}^{\infty}\upsilon_{k,\tau}N_k + \theta_0(\mathcal{I}_\tau(0) - \sum_{k=1}^{\infty}\upsilon_{k,\tau}^2)^{1/2}N_0 - 2^{-1}\theta_0^2\mathcal{I}_\tau(0)\}] \\
&\stackrel{\mathrm{def}}{=} F_{2,\tau}(y),
\end{aligned}
$$

for any $x, y \in \mathbb{R}$. Moreover, it is not hard to verify that $6\sum_{k=1}^{\infty}\lambda_{k,\tau}\{(N_k + \theta_0\{\tau(1-\tau)\}^{-1/2}\mathrm{cov}[\psi_\tau(\varepsilon_\tau)\phi_k\{\mathbf{G}_\pm(\mathbf{x})\}, \eta_\tau(Y, \mathbf{x} \mid 0)])^2 - 1\}$ and $6\sum_{k=1}^{\infty}\lambda_{k,\tau}(N_k + \theta_0\{\tau(1-\tau)\}^{-1/2}\mathrm{cov}[\psi_\tau(\varepsilon_\tau)\phi_k\{\mathbf{G}_\pm(\mathbf{x})\}, \eta_\tau(Y, \mathbf{x} \mid 0)])^2$ have probability measures $F_{1,\tau}(x)$ and $F_{2,\tau}(x)$, respectively. This completes the proof of Theorem 3.

□

S5. Proof of Theorem 4

An application of the proof of Theorem 1 gives

$$\begin{aligned} n\ \widehat{U}_{\tau_l} &= 6\sum_{k=1}^{\infty} n^{-1} \sum_{1\le i_1\ne i_2\le n} \lambda_{k,\tau_l}\phi_{k,\tau_l}\{\psi_{\tau_l}(\varepsilon_{i_1,\tau_l}), \mathbf{G}_{\pm}(\mathbf{x}_{i_1})\} \\ &\quad \phi_{k,\tau_l}\{\psi_{\tau_l}(\varepsilon_{i_2,\tau_l}), \mathbf{G}_{\pm}(\mathbf{x}_{i_2})\} + o_p(1) \\ &= \sum_{k=1}^{\infty} \lambda_k \tau_l(1-\tau_l)[n^{-1/2}\sum_{i_1=1}^{n}\{\tau_l(1-\tau_l)\}^{-1/2}\psi_{\tau_l}(\varepsilon_{i_1,\tau_l})\phi_k\{\mathbf{G}_{\pm}(\mathbf{x}_{i_1})\}]^2 \\ &\quad -\tau_l(1-\tau_l)\sum_{k=1}^{\infty}\lambda_k + o_p(1), \text{ and} \\ n\ \widehat{V}_{\tau_l} &= 6\sum_{k=1}^{\infty} n^{-1} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \lambda_{k,\tau_l}\phi_{k,\tau_l}\{\psi_{\tau_l}(\varepsilon_{i_1,\tau_l}), \mathbf{G}_{\pm}(\mathbf{x}_{i_1})\} \\ &\quad \phi_{k,\tau_l}\{\psi_{\tau_l}(\varepsilon_{i_2,\tau_l}), \mathbf{G}_{\pm}(\mathbf{x}_{i_2})\} + o_p(1) \\ &= \sum_{k=1}^{\infty} \lambda_k \tau_l(1-\tau_l)[n^{-1/2}\sum_{i_1=1}^{n}\{\tau_l(1-\tau_l)\}^{-1/2}\psi_{\tau_l}(\varepsilon_{i_1,\tau_l})\phi_k\{\mathbf{G}_{\pm}(\mathbf{x}_{i_1})\}]^2 \\ &\quad +o_p(1), \end{aligned}$$

for $l = 1, \dots, L$, when the null hypothesis is true. Since $\sum_{k=1}^{\infty} \lambda_k^2 \le 4\int_0^1 J^2(x)dx < \infty$ and $\sum_{k=1}^{\infty} \lambda_k \le 2\{\int_0^1 J^2(x)dx\}^{1/2} < \infty$, it follows that $\sum_{k=K+1}^{\infty} \lambda_k^2 \to 0$ and $\sum_{k=K+1}^{\infty} \lambda_k \to 0$ as $K \to \infty$. By the characteristic function approach used in the proof of Proposition 1, it suffices to show that for every integer K, the joint limiting distribution of the sequences of $\sum_{k=1}^{K} \lambda_k \tau_l(1-\tau_l)[n^{-1/2}\sum_{i_1=1}^{n}\{\tau_l(1-\tau_l)\}^{-1/2}\psi_{\tau_l}(\varepsilon_{i_1,\tau_l})\phi_k\{\mathbf{G}_{\pm}(\mathbf{x}_{i_1})\}]^2, l = 1, \dots, L$ converge in distribution to $\{\tau_l(1-\tau_l)\sum_{k=1}^{K}\lambda_k N_{kl}^2\}_{l=1}^{L}$. The λ_k are defined as in Proposition 1, and $(N_{1l})_{l=1}^{L}, (N_{2l})_{l=1}^{L}, \dots, (N_{Kl})_{l=1}^{L}$ are mutually independent and identically distributed, each with the multivariate normal distribution with mean $\mathbf{0} \in \mathbb{R}^L$ and covariance matrix

$$\mathbf{\Sigma}_0 = [\{\tau_{l_1}\tau_{l_2}(1-\tau_{l_1})(1-\tau_{l_2})\}^{-1/2}\{\min(\tau_{l_1},\tau_{l_2}) - \tau_{l_1}\tau_{l_2}\}]_{l_1,l_2=1}^{L} \in \mathbb{R}^{L\times L}.$$

According to the multivariate central limit theorem, the random vector $[n^{-1/2}\sum_{i_1=1}^{n}\{\tau_l(1-\tau_l)\}^{-1/2}\psi_{\tau_l}(\varepsilon_{i_1,\tau_l})\phi_k\{\mathbf{G}_\pm(\mathbf{x}_{i_1})\}]_{l=1}^{L}$ is asymptotically normal $N(\mathbf{0},\mathbf{\Sigma}_0)$-distributed. Therefore, by the continuous mapping theorem and the Cramér-Wold device, we obtain the required limit.

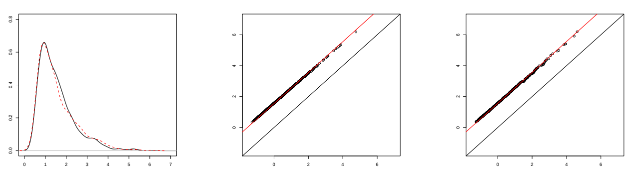
□

S6. Some Aspects of Limiting Distributions

Combined with Proposition 1, Theorem 1 implies that when the null hypothesis is true, as $n \to \infty$, $\{\tau(1-\tau)\}^{-1}n\ \widehat{V}_\tau \to \sum_{k=1}^{\infty}\lambda_k N_k^2$ and $\{\tau(1-\tau)\}^{-1}n\ \widehat{U}_\tau \to \sum_{k=1}^{\infty}\lambda_k(N_k^2-1)$ in distribution. Moreover, the values of λ_k's are independent of the quantile level τ. We give a toy example to illustrate this point. Consider the following mixture distribution:

$$Y = (1+Z)\epsilon - (1+X_1+X_2)(1-\epsilon),$$

where ϵ is a Bernoulli random variable with success probability 0.75, (Z, X_1, X_2) has three components independently generated from Gamma(2, 2), and ϵ is independent of (Z, X_1, X_2). It is not difficult to see that Y is conditionally quantile independent of $\mathbf{x} = (X_1, X_2)^{\mathrm{T}}$ only when $\tau \geq 0.25$. As a graphical illustration, we appreciate the limiting random variables of $\{\tau(1-\tau)\}^{-1}n\,\widehat{V}_\tau$ and $\{\tau(1-\tau)\}^{-1}n\,\widehat{U}_\tau$ with van der Waerden score function by means of a small simulation study for the sample size $n = 100$, and two different values of $\tau = 0.5, 0.75$. As expected, it is shown from Figure 1 that different quantile levels do not alter the asymptotic null distributions of $\{\tau(1-\tau)\}^{-1}n\,\widehat{V}_\tau$ and $\{\tau(1-\tau)\}^{-1}n\,\widehat{U}_\tau$.



(A) (B) (C)

Figure 1: (A) Density curves of 1000 realizations of van der Waerden score-based statistics $\{\tau(1-\tau)\}^{-1}n\,\widehat{V}_\tau I(\tau = 0.5)$ (black solid line) and $\{\tau(1-\tau)\}^{-1}n\,\widehat{V}_\tau I(\tau = 0.75)$ (red dashed line). (B) Quantile-quantile plots of 1000 realizations of $\{\tau(1-\tau)\}^{-1}n\,\widehat{U}_\tau I(\tau = 0.5)$ versus $\{\tau(1-\tau)\}^{-1}n\,\widehat{V}_\tau I(\tau = 0.5)$. (C) Quantile-quantile plots of 1000 realizations of $\{\tau(1-\tau)\}^{-1}n\,\widehat{U}_\tau I(\tau = 0.75)$ versus $\{\tau(1-\tau)\}^{-1}n\,\widehat{V}_\tau I(\tau = 0.75)$.

S7. Comparison under Moderate Dimension

We next examine the empirical performance of our center-outward sign- and rank-based tests under moderate dimension by employing

$$\text{a heteroscedastic model}: Y = \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_1 + \exp(\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_2 + \epsilon). \qquad \text{(S7.56)}$$

In the heteroscedastic model (S7.56), we set $q = 7$, $\boldsymbol{\beta}_1 = (0, 1/3, 0, \ldots, 0)^{\mathrm{T}}$ and $\boldsymbol{\beta}_2 = (0, 0, 1/3, \ldots, 1/3)^{\mathrm{T}}$. The covariates $X_1, \ldots, X_7$ are distributed independently and identically according to the standard normal distribution. Independently of the covariates, the error ϵ follows the standard Cauchy distribution. The error distribution is heavy-tailed in this example. We consider two values of moderately large to small sample sizes $n = 200, 100$. Tables S1 and S2 report the empirical sizes and powers of the tests considered in our study. The proposed center-outward score tests and the SZ, ZYS tests are robust to the presence of heavy-tailed errors. It might not be surprising to see that the aforementioned tests control the Type-I error well around the nominal level. Since the ZYS test is designed for large dimension, it has reasonably good powers at a particular quantile level $\tau \in \{0.25, 0.5\}$. Inspection of the two tables further reveals that when the sample size is sufficiently large, our center-outward score tests across multiple quantiles deliver comparable power results to the competitors. The displeasure exception occurs when the dimension is relatively large and the sample size is small. Such phenomenon can be attributed to the slow convergence rate of empirical center-outward ranks and signs when the dimension gets large (Shi et al., 2025, page 5, lines 45-46).

Table S1: The empirical sizes and powers of the proposed tests with different score functions and the SZ, ZYS tests under different quantiles at the nominal level 5% for Model (S7.56) when $n = 200$. (With a fixed seed, the symbol "$--$" indicates that the value at this position is the same as the previous line.)

score function	$\widehat{V}_{0.25}$	$\widehat{U}_{0.25}$	$\mathrm{SZ}_{0.25}$	$\mathrm{ZYS}_{0.25}$	$\widehat{V}_{0.5}$	$\widehat{U}_{0.5}$	$\mathrm{SZ}_{0.5}$	$\mathrm{ZYS}_{0.5}$
				Size				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.056	0.060	0.058	0.062	0.055	0.057	0.051	0.054
x	0.061	0.054	$--$	$--$	0.051	0.052	$--$	$--$
1	0.054	0.055	$--$	$--$	0.055	0.057	$--$	$--$
				Power				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.856	0.854	0.813	0.976	0.832	0.829	0.687	0.930
x	0.845	0.851	$--$	$--$	0.815	0.814	$--$	$--$
1	0.829	0.823	$--$	$--$	0.803	0.801	$--$	$--$
	$\widehat{V}_{0.75}$	$\widehat{U}_{0.75}$	$\mathrm{SZ}_{0.75}$	$\mathrm{ZYS}_{0.75}$	$\widehat{S}_{1,\mathfrak{T}}$	$\widehat{S}_{2,\mathfrak{T}}$	$\widehat{M}_{1,\mathfrak{T}}$	$\widehat{M}_{2,\mathfrak{T}}$
				Size				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.048	0.050	0.042	0.048	0.056	0.059	0.054	0.055
x	0.038	0.038	$--$	$--$	0.050	0.052	0.053	0.059
1	0.036	0.039	$--$	$--$	0.039	0.037	0.043	0.044
				Power				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.390	0.380	0.234	0.516	0.901	0.897	0.861	0.880
x	0.361	0.355	$--$	$--$	0.892	0.895	0.844	0.864
1	0.349	0.343	$--$	$--$	0.885	0.884	0.838	0.849

Table S2: The empirical sizes and powers of the proposed tests with different score functions and the SZ, ZYS tests under different quantiles at the nominal level 5% for Model (S7.56) when $n = 100$.

score function	$\widehat{V}_{0.25}$	$\widehat{U}_{0.25}$	$\mathrm{SZ}_{0.25}$	$\mathrm{ZYS}_{0.25}$	$\widehat{V}_{0.5}$	$\widehat{U}_{0.5}$	$\mathrm{SZ}_{0.5}$	$\mathrm{ZYS}_{0.5}$
				Size				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.032	0.038	0.054	0.052	0.058	0.062	0.052	0.060
x	0.049	0.048	$--$	$--$	0.059	0.061	$--$	$--$
1	0.050	0.055	$--$	$--$	0.052	0.058	$--$	$--$
				Power				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.427	0.435	0.340	0.695	0.394	0.408	0.277	0.570
x	0.421	0.430	$--$	$--$	0.380	0.406	$--$	$--$
1	0.414	0.426	$--$	$--$	0.377	0.385	$--$	$--$
	$\widehat{V}_{0.75}$	$\widehat{U}_{0.75}$	$\mathrm{SZ}_{0.75}$	$\mathrm{ZYS}_{0.75}$	$\widehat{S}_{1,\mathfrak{T}}$	$\widehat{S}_{2,\mathfrak{T}}$	$\widehat{M}_{1,\mathfrak{T}}$	$\widehat{M}_{2,\mathfrak{T}}$
				Size				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.034	0.036	0.060	0.054	0.052	0.052	0.057	0.061
x	0.048	0.046	$--$	$--$	0.051	0.054	0.054	0.059
1	0.048	0.055	$--$	$--$	0.056	0.060	0.055	0.062
				Power				
$F^{-1}_{(\chi^2_q)^{1/2}}(x)$	0.172	0.175	0.114	0.232	0.462	0.470	0.432	0.467
x	0.161	0.170	$--$	$--$	0.445	0.461	0.418	0.453
1	0.156	0.168	$--$	$--$	0.441	0.452	0.409	0.450

References

Hallin, M., E. del Barrio, J. Cuesta-Albertos, and C. Matrán (2021). Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach. *The Annals of Statistics 49*, 1139–1165.

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

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

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

Sherman, R. (1994). Maximal inequalities for degenerate u-processes with applications to optimization estimators. *The Annals of Statistics 22*, 439–459.

Shi, H., M. Drton, M. Hallin, and F. Han (2025). Supplement to "distribution-free tests of multivariate independence based on center-outward quadrant, spearman, kendall, and van der waerden statistics". *Bernoulli 31*, 106–129.

Shi, H., M. Drton, and F. Han (2022). Distribution-free consistent independence tests via center-outward ranks and signs. *Journal of the American Statistical Association 117*, 395–410.

Shi, H., M. Hallin, M. Drton, and F. Han (2022). On universally consistent and fully distribution-free rank tests of vector independence. *The Annals of Statistics 50*, 1933–1959.

Shorack, G. R. (2017). *Probability For Statisticians.* Springer, Cham, Switzerland.

Simon, B. (2015a). *Operator Theory.* A Comprehensive Course in Analysis, Part 4. American Mathematical Society, Providence, RI.

Simon, B. (2015b). *Real Analysis.* A Comprehensive Course in Analysis, Part 1. American Mathematical Society, Providence, RI.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

Yao, S., X. Zhang, X. Shao, et al. (2018). Testing mutual independence in high dimension via distance covariance. *Journal of the Royal Statistical Society Series B 80*, 455–480.

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