Supplement to “Two-Sample Testing for Tail Copulas with an Application to Equity Indices”

Sami Umut Can
Department of Quantitative Economics, University of Amsterdam

John H.J. Einmahl
Department of Econometrics & OR and CentER, Tilburg University

Roger J.A. Laeven
Department of Quantitative Economics, University of Amsterdam

December 1, 2022

Abstract

This text serves as an appendix to the paper “Two-Sample Testing for Tail Copulas with an Application to Equity Indices.” For context, notation and definitions, see that paper. First, we provide the proofs of Theorems 3.1, 4.1 and 5.1. Then, we present simulation results under serial dependence.

Proof of Theorem 3.1

By Skorohod's representation theorem, there is a probability space where probabilistically equivalent versions of all the random elements in Assumption A1 are defined, those in (8) independent of those in (9), and the convergences (8) and (9) hold in probability. All statements in this proof should be understood as statements about random elements in this probability space.

Given $(x, y) \in [0, \infty)^2$, let us define points $(\widehat{x}, \widehat{y})$ and $(\widehat{x}', \widehat{y}')$ by

$$\begin{aligned}
\widehat{x} &= \left[\left(1+\gamma_1\left(\frac{x^{-\widehat{\gamma}_1}-1}{\widehat{\gamma}_1}\cdot\frac{\widehat{a}_1}{a_1}+\frac{\widehat{b}_1-b_1}{a_1}\right)\right)\vee 0\right]^{-1/\gamma_1},\\
\widehat{y} &= \left[\left(1+\gamma_2\left(\frac{y^{-\widehat{\gamma}_2}-1}{\widehat{\gamma}_2}\cdot\frac{\widehat{a}_2}{a_2}+\frac{\widehat{b}_2-b_2}{a_2}\right)\right)\vee 0\right]^{-1/\gamma_2},\\
\widehat{x}' &= \left[\left(1+\gamma_1'\left(\frac{x^{-\widehat{\gamma}_1'}-1}{\widehat{\gamma}_1'}\cdot\frac{\widehat{a}_1'}{a_1'}+\frac{\widehat{b}_1'-b_1'}{a_1'}\right)\right)\vee 0\right]^{-1/\gamma_1'},\\
\widehat{y}' &= \left[\left(1+\gamma_2'\left(\frac{y^{-\widehat{\gamma}_2'}-1}{\widehat{\gamma}_2'}\cdot\frac{\widehat{a}_2'}{a_2'}+\frac{\widehat{b}_2'-b_2'}{a_2'}\right)\right)\vee 0\right]^{-1/\gamma_2'}.
\end{aligned}$$

It follows from Lemma 1.1 in the Appendix of Can et al. (2015) that

$$\begin{aligned}
&\sup_{x\in[\delta,T]}\left|\sqrt{k}(\widehat{x}-x)-\left[f(x,\gamma_1)A_1+g(x,\gamma_1)B_1+h(x,\gamma_1)\Gamma_1\right]\right| \xrightarrow{P} 0,\\
&\sup_{y\in[\delta,T]}\left|\sqrt{k}(\widehat{y}-y)-\left[f(y,\gamma_2)A_2+g(y,\gamma_2)B_2+h(y,\gamma_2)\Gamma_2\right]\right| \xrightarrow{P} 0,\\
&\sup_{x\in[\delta,T]}\left|\sqrt{k'}(\widehat{x}'-x)-\left[f(x,\gamma_1')A_1'+g(x,\gamma_1')B_1'+h(x,\gamma_1')\Gamma_1'\right]\right| \xrightarrow{P} 0,\\
&\sup_{y\in[\delta,T]}\left|\sqrt{k'}(\widehat{y}'-y)-\left[f(y,\gamma_2')A_2'+g(y,\gamma_2')B_2'+h(y,\gamma_2')\Gamma_2'\right]\right| \xrightarrow{P} 0.
\end{aligned} \tag{S.1}$$

Now, let $\widehat{R}_n$ and T_n be as defined in (6) and (7), respectively, and let $\widehat{R}'_{n'}$ and $T'_{n'}$ be their analogues constructed from the second sample. Note that the probability of the event

$$\{\widehat{R}_n(x,y)=T_n(\widehat{x},\widehat{y}) \text{ and } \widehat{R}'_{n'}(x,y)=T'_{n'}(\widehat{x}',\widehat{y}') \text{ for all } (x,y)\in[\delta,T]^2\} \tag{S.2}$$

tends to 1 as $n, n' \to \infty$. Hence, instead of $\eta_{n,n'}$, it will suffice to show convergence for

$\eta^*_{n,n'}(x,y) := \sqrt{\kappa}[T_n(\widehat{x},\widehat{y}) - T'_{n'}(\widehat{x}',\widehat{y}')]$, which we decompose as follows:

$$\begin{aligned}\eta^*_{n,n'}(x,y) &= \sqrt{\kappa}[T_n(\widehat{x},\widehat{y}) - R_n(\widehat{x},\widehat{y})] + \sqrt{\kappa}[R_n(\widehat{x},\widehat{y}) - R(\widehat{x},\widehat{y})] \\ &\quad + \sqrt{\kappa}[R(\widehat{x},\widehat{y}) - R(x,y)] - \sqrt{\kappa}[T'_{n'}(\widehat{x}',\widehat{y}') - R'_{n'}(\widehat{x}',\widehat{y}')] \\ &\quad - \sqrt{\kappa}[R'_{n'}(\widehat{x}',\widehat{y}') - R(\widehat{x}',\widehat{y}')] - \sqrt{\kappa}[R(\widehat{x}',\widehat{y}') - R(x,y)] \\ &=: \eta^*_{1n}(x,y) + \eta^*_{2n}(x,y) + \eta^*_{3n}(x,y) - \eta^*_{4n'}(x,y) - \eta^*_{5n'}(x,y) - \eta^*_{6n'}(x,y).\end{aligned}$$

The in-probability convergence (8), (S.1) and the continuity of V_R yield

$$\sup_{(x,y)\in[\delta,T]^2} \left|\eta^*_{1n}(x,y) - \sqrt{c}V_R(x,y)\right| \xrightarrow{P} 0. \tag{S.3}$$

From Assumption A3 and (S.1) it also follows that

$$\sup_{(x,y)\in[\delta,T]^2} \left|\eta^*_{2n}(x,y)\right| \xrightarrow{P} 0. \tag{S.4}$$

Moreover, from the Mean Value Theorem we know that

$$\eta^*_{3n}(x,y) = \sqrt{\kappa}[R^{(1)}(\breve{x},\breve{y})(\widehat{x}-x) + R^{(2)}(\breve{x},\breve{y})(\widehat{y}-y),$$

for some $(\breve{x},\breve{y})$ lying on the line segment connecting (x,y) and $(\widehat{x},\widehat{y})$. The convergence (S.1) in combination with Assumption A2 now yields that

$$\begin{aligned}\sup_{(x,y)\in[\delta,T]^2} \Big| &\eta^*_{3n}(x,y) - \sqrt{c}R^{(1)}(x,y)\big[f(x,\gamma_1)A_1 + g(x,\gamma_1)B_1 + h(x,\gamma_1)\Gamma_1\big] \\ &- \sqrt{c}R^{(2)}(x,y)\big[f(y,\gamma_2)A_2 + g(y,\gamma_2)B_2 + h(y,\gamma_2)\Gamma_2\big]\Big| \xrightarrow{P} 0.\end{aligned} \tag{S.5}$$

The analogues of (S.3), (S.4), (S.5) for $\eta^*_{4n'}$, $\eta^*_{5n'}$, $\eta^*_{6n'}$ are proved along similar lines. Combining these six results completes the proof.

Proof of Theorem 4.1

This result essentially follows from the general martingale transformation result in Theorem 3.1 of Can et al. (2015). Instead of arbitrary Borel sets $B \subset [\delta,T]^2$ considered therein, we consider rectangles $[\delta, \delta + x] \times [\delta, \delta + y]$ for $0 \le x, y \le T - \delta$. Furthermore, we use the scanning family $A_u = [\delta, T] \times [\delta, (1-u)\delta + uT]$, $0 \le u \le 1$. Then, from (12) under Assumptions A4–A5 confined to the true $\gamma_1, \gamma_2, \gamma_1', \gamma_2', r, r^{(1)}, r^{(2)}$, we obtain that

$$\begin{aligned} W_R(x,y) &= \int_\delta^{\delta+x}\int_\delta^{\delta+y} \mathrm{d}\eta(s,t) \\ &\quad - \int_\delta^{\delta+x}\int_\delta^{\delta+y} \mathbf{q}(s,t)^\mathsf{T}\Big(\mathbf{I}_{\delta,T}^{-1}(t)\int_\delta^T\int_t^T \mathbf{q}(s',t')\,\mathrm{d}\eta(s',t')\Big) r(s,t)\,\mathrm{d}s\,\mathrm{d}t \end{aligned}$$

is a bivariate Wiener process on $[0,\tau]^2$ for any $\tau \in (\delta, T-\delta)$, with "time" measure $R([\delta, \delta + \cdot\,] \times [\delta, \delta + \cdot\,])$. That is, W_R is a zero-mean Gaussian process with covariance structure

$$E[W_R(x,y)W_R(x',y')] = R([\delta, \delta + x \wedge x'] \times [\delta, \delta + y \wedge y']),$$

for $(x,y),(x',y') \in [0,\tau]^2$. It then follows from the standard theory of multivariate Gaussian processes (see, e.g., the lemma preceding Theorem 3 in Khmaladze (1988)) that the normalized process

$$\begin{aligned} W(x,y) &= \int_\delta^{\delta+x}\int_\delta^{\delta+y} \frac{1}{\sqrt{r(s,t)}}\,\mathrm{d}W_R(s,t) \\ &= \int_\delta^{\delta+x}\int_\delta^{\delta+y} \frac{1}{\sqrt{r(s,t)}}\,\mathrm{d}\eta(s,t) \\ &\quad - \int_\delta^{\delta+x}\int_\delta^{\delta+y} \mathbf{q}(s,t)^\mathsf{T}\Big(\mathbf{I}_{\delta,T}^{-1}(t)\int_\delta^T\int_t^T \mathbf{q}(s',t')\,\mathrm{d}\eta(s',t')\Big)\sqrt{r(s,t)}\,\mathrm{d}s\,\mathrm{d}t \end{aligned}$$

is a standard bivariate Wiener process on $[0,\tau]^2$.

Proof of Theorem 5.1

We start by establishing the limit relations

$$\sup_{(x,y)\in[\delta,T]^2} |\widehat{r}(x,y) - r(x,y)| \xrightarrow{P} 0 \quad \text{as } n, n' \to \infty, \tag{S.6}$$

$$\sup_{(x,y)\in[\delta,T]^2} |\widehat{r}^{(12)}(x,y) - r^{(12)}(x,y)| \xrightarrow{P} 0 \quad \text{as } n, n' \to \infty, \tag{S.7}$$

where $\widehat{r}^{(12)} := \partial^2\widehat{r}/(\partial x\partial y)$.

Clearly, for (S.6), it is sufficient to show that

$$\sup_{(x,y)\in[\delta,T]^2} |s_w(x,y) - r(x,y)| \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

We write

$$\begin{aligned} &s_w(x,y) - r(x,y) \\ &\quad = \frac{1}{w^2} \int_{y-w}^{y+w} \int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right) \mathrm{d}(\widehat{R}_n(u,v) - R(u,v)) \\ &\qquad + \frac{1}{w^2} \int_{y-w}^{y+w} \int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right) (r(u,v) - r(x,y))\, \mathrm{d}u\, \mathrm{d}v. \end{aligned} \tag{S.8}$$

Applying bivariate integration by parts (see e.g., Henstock (1973), Theorem 3), we find that the first term on the right-hand side is equal to

$$\frac{1}{w^2} \int_{y-w}^{y+w} \int_{x-w}^{x+w} (\widehat{R}_n(u,v) - R(u,v))\, \mathrm{d}K\left(\frac{x-u}{w}\right) \mathrm{d}K\left(\frac{y-v}{w}\right).$$

From (S.3), (S.4) and (S.5), this expression is $O_P((k^{1/10})^2 k^{-1/2})(\int_{-1}^{1} |\mathrm{d}K(u)|)^2 = o_P(1)$ uniformly on $[\delta,T]^2$. The absolute value of the second term in (S.8) is bounded by

$$\sup_{(u,v)\in[x-w,\,x+w]\times[y-w,\,y+w]} |r(u,v) - r(x,y)|,$$

which by the (uniform) continuity of r tends to 0, uniformly on $[\delta,T]^2$. The convergence (S.6) is thereby established.

For (S.7), it is sufficient to show, with $s_w^{(12)} := \partial^2 s_w/(\partial x \partial y)$, that

$$\sup_{(x,y)\in[\delta,T]^2} |s_w^{(12)}(x,y) - r^{(12)}(x,y)| \xrightarrow{P} 0, \quad \text{as } n \to \infty.$$

We have

$$\begin{aligned} &s_w^{(12)}(x,y) - r^{(12)}(x,y) \\ &\quad = \frac{1}{w^4} \int_{y-w}^{y+w} \int_{x-w}^{x+w} K^{(1)}\left(\frac{x-u}{w}\right) K^{(1)}\left(\frac{y-v}{w}\right) \mathrm{d}(\widehat{R}_n(u,v) - R(u,v)) \\ &\qquad + \frac{1}{w^4} \int_{y-w}^{y+w} \int_{x-w}^{x+w} K^{(1)}\left(\frac{x-u}{w}\right) K^{(1)}\left(\frac{y-v}{w}\right) r(u,v)\, \mathrm{d}u\, \mathrm{d}v - r^{(12)}(x,y). \end{aligned} \tag{S.9}$$

Similar as for the first term in (S.8), it can be shown that the first term of (S.9) is $O_P((k^{1/10})^4 k^{-1/2})(\int_{-1}^1 |\mathrm{d}K^{(1)}(u)|)^2 = o_P(1)$ uniformly on $[\delta, T]^2$. The second term in (S.9) is equal to

$$\begin{aligned}
&\frac{1}{w^2}\int_{y-w}^{y+w}\int_{x-w}^{x+w} r(u,v)\,\mathrm{d}K\left(\frac{x-u}{w}\right)\mathrm{d}K\left(\frac{y-v}{w}\right) - r^{(12)}(x,y)\\
&\quad = \frac{1}{w^2}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right)\mathrm{d}r(u,v) - r^{(12)}(x,y)\\
&\quad = \frac{1}{w^2}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right)(r^{(12)}(u,v) - r^{(12)}(x,y))\,\mathrm{d}u\,\mathrm{d}v,
\end{aligned}$$

where for the first equality again bivariate integration by parts is used. The absolute value of the last expression is bounded by

$$\sup_{(u,v)\in[x-w,\,x+w]\times[y-w,\,y+w]} |r^{(12)}(u,v) - r^{(12)}(x,y)|,$$

which by the (uniform) continuity of $r^{(12)}$ tends to 0, uniformly on $[\delta, T]^2$.

Next, we establish that for $j = 1, 2$,

$$\sup_{(x,y)\in[\delta,T]^2} |\widehat{r}^{(j)}(x,y) - r^{(j)}(x,y)| \xrightarrow{P} 0 \quad \text{as } n, n' \to \infty, \tag{S.10}$$

$$\sup_{(x,y)\in[\delta,T]^2} |\widehat{r}^{(j12)}(x,y) - r^{(j12)}(x,y)| \xrightarrow{P} 0 \quad \text{as } n, n' \to \infty, \tag{S.11}$$

where $\widehat{r}^{(112)} := \partial^2 \widehat{r}^{(1)}/(\partial x \partial y)$ and $\widehat{r}^{(212)} := \partial^2 \widehat{r}^{(2)}/(\partial x \partial y)$.

For (S.10), it is sufficient to show that

$$\sup_{(x,y)\in[\delta,T]^2} \left| \frac{1}{kw^3}\sum_{i=1}^n K^{(1)}\left(\frac{x - \widehat{X}_i(n/k)}{w}\right) K\left(\frac{y - \widehat{Y}_i(n/k)}{w}\right) - r^{(1)}(x,y)\right| \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

The expression inside the absolute-value signs is equal to

$$\begin{aligned}
&\frac{1}{w^3}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K^{(1)}\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right)\mathrm{d}(\widehat{R}_n(u,v) - R(u,v))\\
&\quad + \frac{1}{w^3}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K^{(1)}\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right) r(u,v)\,\mathrm{d}u\,\mathrm{d}v - r^{(1)}(x,y).
\end{aligned} \tag{S.12}$$

From (S.3)–(S.5) again, the first term is now $O_P((k^{1/12})^3 k^{-1/2}) = o_P(1)$ uniformly on $[\delta, T]^2$.

The second term in (S.12) is equal to

$$\begin{aligned}
&\frac{1}{w^2}\int_{y-w}^{y+w} K\left(\frac{y-v}{w}\right)\int_{x-w}^{x+w} r(u,v)\,\mathrm{d}K\left(\frac{x-u}{w}\right)\mathrm{d}v - r^{(1)}(x,y)\\
&\quad= \frac{1}{w^2}\int_{y-w}^{y+w} K\left(\frac{y-v}{w}\right)\int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) r^{(1)}(u,v)\,\mathrm{d}u\,\mathrm{d}v - r^{(1)}(x,y)\\
&\quad= \frac{1}{w^2}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right)(r^{(1)}(u,v) - r^{(1)}(x,y))\,\mathrm{d}u\,\mathrm{d}v,
\end{aligned}$$

which tends to 0, uniformly on $[\delta,T]^2$, by the (uniform) continuity of $r^{(1)}$. The limit relation (S.10) is thereby established. (S.11) follows by very similar reasoning, using that, with $K^{(2)}$ the second derivative of K,

$$\begin{aligned}
&\frac{1}{kw^5}\sum_{i=1}^{n} K^{(2)}\left(\frac{x-\widehat{X}_i(n/k)}{w}\right) K^{(1)}\left(\frac{y-\widehat{Y}_i(n/k)}{w}\right)\\
&\quad= \frac{1}{w^5}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K^{(2)}\left(\frac{x-u}{w}\right) K^{(1)}\left(\frac{y-v}{w}\right)\mathrm{d}(\widehat{R}_n(u,v) - R(u,v))\\
&\qquad+ \frac{1}{w^2}\int_{y-w}^{y+w}\int_{x-w}^{x+w} K\left(\frac{x-u}{w}\right) K\left(\frac{y-v}{w}\right)(r^{(112)}(u,v) - r^{(112)}(x,y))\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}$$

Now, by Theorem 3.1 and Skorohod's representation theorem, there is a probability space where probabilistically equivalent versions of $\eta_{n,n'}$ and η are defined, and these satisfy $\|\eta_{n,n'} - \eta\|_{[\delta,T]^2} \to 0$ a.s., with $\|\cdot\|_S := \sup_S|\cdot|$ for $S \subset [0,\infty)^2$. We will show that in this probability space,

$$\|W_{n,n'} - W\|_{[0,\tau]^2} \xrightarrow{P} 0, \tag{S.13}$$

with W as defined in (15). In view of Theorem 4.1, this will suffice for the proof.

Given $a < b \in \mathbb{R}$ and a function $\varphi : [a,b]^2 \to \mathbb{R}$, we will let $V^{\mathrm{HK}}_{[a,b]^2}(\varphi)$ denote the *Hardy-Krause variation* of this function over the square $[a,b]^2$. That is,

$$V^{\mathrm{HK}}_{[a,b]^2}(\varphi) = V^{(2)}_{[a,b]^2}(\varphi) + V^{(1)}_{[a,b]}(\varphi(\,\cdot\,,a)) + V^{(1)}_{[a,b]}(\varphi(\,\cdot\,,b)) + V^{(1)}_{[a,b]}(\varphi(a,\,\cdot\,)) + V^{(1)}_{[a,b]}(\varphi(b,\,\cdot\,)),$$

where $V^{(1)}_{[a,b]}$ denotes the univariate total variation over the interval $[a,b]$ and $V^{(2)}_{[a,b]}$ denotes the bivariate Vitali total variation over $[a,b]^2$. Note that if the partial derivatives $\varphi^{(1)}(x,a) := \partial\varphi(x,a)/\partial x$ and $\varphi^{(2)}(a,y) := \partial\varphi(a,y)/\partial y$, as well as the analogously defined $\varphi^{(1)}(x,b)$ and

$\varphi^{(2)}(b,y)$ exist on $[a,b]$, and the mixed partial derivative $\partial\varphi(x,y)/(\partial x\partial y)$ exists on $[a,b]^2$, then

$$V^{\mathrm{HK}}_{[a,b]^2}(\varphi) \le \int_a^b\int_a^b |\varphi^{(12)}(x,y)|\,\mathrm{d}x\,\mathrm{d}y + \int_a^b |\varphi^{(1)}(x,a)|\,\mathrm{d}x + \int_a^b |\varphi^{(1)}(x,b)|\,\mathrm{d}x + \int_a^b |\varphi^{(2)}(a,y)|\,\mathrm{d}y + \int_a^b |\varphi^{(2)}(b,y)|\,\mathrm{d}y.$$

We refer to Owen (2005) for an overview of various concepts of multivariate variation.

Throughout the proof, we will let $A_\delta(x,y)$ denote the rectangle $[\delta,\delta+x]\times[\delta,\delta+y]$ for $(x,y)\in[0,\tau]^2$. Note that (S.13) will follow from

$$\left\|\int_{A_\delta(x,y)} \frac{1}{\sqrt{\widehat{r}(s,t)}}\,\mathrm{d}\eta_{n,n'}(s,t) - \int_{A_\delta(x,y)} \frac{1}{\sqrt{r(s,t)}}\,\mathrm{d}\eta(s,t)\right\|_{[0,\tau]^2} \xrightarrow{P} 0 \tag{S.14}$$

and

$$\begin{aligned}\Bigg\|\int_{A_\delta(x,y)} \widehat{\mathbf{q}}(s,t)^\mathsf{T}\left(\widehat{\mathbf{I}}^{-1}_{\delta,T}(t)\int_\delta^T\int_t^T \widehat{\mathbf{q}}(s',t')\,\mathrm{d}\eta_{n,n'}(s',t')\right)\sqrt{\widehat{r}(s,t)}\,\mathrm{d}s\,\mathrm{d}t \\ - \int_{A_\delta(x,y)} \mathbf{q}(s,t)^\mathsf{T}\left(\mathbf{I}^{-1}_{\delta,T}(t)\int_\delta^T\int_t^T \mathbf{q}(s',t')\,\mathrm{d}\eta(s',t')\right)\sqrt{r(s,t)}\,\mathrm{d}s\,\mathrm{d}t\Bigg\|_{[0,\tau]^2} \xrightarrow{P} 0.\end{aligned} \tag{S.15}$$

We will prove (S.14) first. Define $\Delta_{n,n'} := \eta_{n,n'} - \eta$, $\sigma(x,y) := r(x,y)^{-1/2}$, $\widehat{\sigma}(x,y) := \widehat{r}(x,y)^{-1/2}$, and $\Delta\sigma(x,y) := \sigma(x,y) - \widehat{\sigma}(x,y)$. Then (S.14) will follow from

$$\left\|\int_{A_\delta(x,y)} \Delta\sigma(s,t)\,\mathrm{d}\eta(s,t)\right\|_{[0,\tau]^2} \xrightarrow{P} 0, \qquad \left\|\int_{A_\delta(x,y)} \widehat{\sigma}(s,t)\,\mathrm{d}\Delta_{n,n'}(s,t)\right\|_{[0,\tau]^2} \xrightarrow{P} 0. \tag{S.16}$$

Applying bivariate integration by parts to the first integral term in (S.16), we obtain the following bound:

$$\begin{aligned}\left|\int_{A_\delta(x,y)} \Delta\sigma(s,t)\,\mathrm{d}\eta(s,t)\right| &\le \sum_{(u,v)\in\mathcal{V}_\delta(x,y)} \big|\Delta\sigma(u,v)\,\eta(u,v)\big| + \|\eta\|_{A_\delta(x,y)} V^{\mathrm{HK}}_{A_\delta(x,y)}(\Delta\sigma) \\ &\le \|\eta\|_{[\delta,T]^2}\big(4\|\Delta\sigma\|_{[\delta,T]^2} + V^{\mathrm{HK}}_{[\delta,T]^2}(\Delta\sigma)\big),\end{aligned} \tag{S.17}$$

where $\mathcal{V}_\delta(x,y)$ denotes the set of the four vertices of the rectangle $A_\delta(x,y)$. Now, Assumption A2 ensures that η is continuous (hence bounded) on $[\delta,T]^2$, (S.6) ensures that $|\Delta\sigma|$ is $o_P(1)$ uniformly over $[\delta,T]^2$, and (S.7) together with a similar result for the first-order partial derivatives of $\widehat{r}$ (cf. (S.10)) ensures that $V^{\mathrm{HK}}_{[\delta,T]^2}(\Delta\sigma)$ is $o_P(1)$ as well. It follows that the far right-hand side of (S.17) vanishes in probability, and the first convergence in (S.16) is proved. The second convergence in (S.16) follows from a similar integration by parts argument:

$$\left|\int_{A_\delta(x,y)} \widehat{\sigma}(s,t)\,\mathrm{d}\Delta_{n,n'}(s,t)\right| \le \|\Delta_{n,n'}\|_{[\delta,T]^2}\big(4\|\widehat{\sigma}\|_{[\delta,T]^2} + V^{\mathrm{HK}}_{[\delta,T]^2}(\widehat{\sigma})\big),$$

where the right-hand side is $o_P(1)$ since $\|\Delta_{n,n'}\|_{[\delta,T]^2}$ is $o_P(1)$ and $\|\widehat{\sigma}\|_{[\delta,T]^2}$ as well as $V^{\mathrm{HK}}_{[\delta,T]^2}(\widehat{\sigma})$ are $O_P(1)$ terms.

We have thus established (S.14), and it remains to prove (S.15). For ease of notation, we let

$$\begin{aligned}
H(s,t) &= \mathbf{q}(s,t)^{\mathsf{T}}\mathbf{I}^{-1}_{\delta,t}(t)\int_\delta^T\int_t^T \mathbf{q}(s',t')\,\mathrm{d}\eta(s',t'),\\
H_{n,n'}(s,t) &= \mathbf{q}(s,t)^{\mathsf{T}}\mathbf{I}^{-1}_{\delta,T}(t)\int_\delta^T\int_t^T \mathbf{q}(s',t')\,\mathrm{d}\eta_{n,n'}(s',t'),\\
\widehat{H}(s,t) &= \widehat{\mathbf{q}}(s,t)^{\mathsf{T}}\widehat{\mathbf{I}}^{-1}_{\delta,T}(t)\int_\delta^T\int_t^T \widehat{\mathbf{q}}(s',t')\,\mathrm{d}\eta(s',t'),\\
\widehat{H}_{n,n'}(s,t) &= \widehat{\mathbf{q}}(s,t)^{\mathsf{T}}\widehat{\mathbf{I}}^{-1}_{\delta,T}(t)\int_\delta^T\int_t^T \widehat{\mathbf{q}}(s',t')\,\mathrm{d}\eta_{n,n'}(s',t').
\end{aligned}$$

Then (S.15) can be written succinctly as

$$\left\|\int_{A_\delta(x,y)}\Big(\widehat{H}_{n,n'}(s,t)\sqrt{\widehat{r}(s,t)} - H(s,t)\sqrt{r(s,t)}\Big)\,\mathrm{d}s\,\mathrm{d}t\right\|_{[0,\tau]^2} \xrightarrow{P} 0,$$

which can be proved by showing

$$\big\|H\big(\sqrt{\widehat{r}}-\sqrt{r}\big)\big\|_{A_\delta(\tau,\tau)} \xrightarrow{P} 0, \qquad \big\|\big(\widehat{H}_{n,n'}-H\big)\sqrt{\widehat{r}}\,\big\|_{A_\delta(\tau,\tau)} \xrightarrow{P} 0. \tag{S.18}$$

The first convergence in (S.18) follows easily from the continuity (hence boundedness) of H over $A_\delta(\tau,\tau)$ and (S.6). As for the second convergence in (S.18), since $\|\sqrt{\widehat{r}}\|_{A_\delta(\tau,\tau)} = O_P(1)$, we need to show that $\|\widehat{H}_{n,n'} - H\|_{A_\delta(\tau,\tau)} \xrightarrow{P} 0$. We will do this by proving

$$\|H_{n,n'} - H\|_{A_\delta(\tau,\tau)} \xrightarrow{P} 0, \qquad \|\widehat{H}_{n,n'} - H_{n,n'}\|_{A_\delta(\tau,\tau)} \xrightarrow{P} 0. \tag{S.19}$$

Consider the first convergence in (S.19). We have

$$\|H_{n,n'} - H\|_{A_\delta(\tau,\tau)} = \left\| \mathbf{q}(s,t)^\mathsf{T} \mathbf{I}_{\delta,T}^{-1}(t) \int_\delta^T \int_t^T \mathbf{q}(s',t')\, \mathrm{d}\Delta_{n,n'}(s',t') \right\|_{A_\delta(\tau,\tau)},$$

with $\Delta_{n,n'} = \eta_{n,n'} - \eta$, as before. The term $|\mathbf{q}(s,t)^\mathsf{T} \mathbf{I}_{\delta,T}^{-1}(t)|$ is component-wise bounded on $A_\delta(\tau,\tau)$ by continuity, so we need to show that

$$\sup_{t\in[\delta,\delta+\tau]} \left| \int_\delta^T \int_t^T q_i(s',t')\, \mathrm{d}\Delta_{n,n'}(s',t') \right| \xrightarrow{P} 0, \qquad i = 1,\ldots,8. \tag{S.20}$$

Applying integration by parts as before, we obtain

$$\left| \int_\delta^T \int_t^T q_i(s',t')\, \mathrm{d}\Delta_{n,n'}(s',t') \right| \le \|\Delta_{n,n'}\|_{[\delta,T]^2} \big(4\|q_i\|_{[\delta,T]^2} + V^{\mathrm{HK}}_{[\delta,T]^2}(q_i)\big),$$

where the right-hand side is $o_P(1)$ since $\|\Delta_{n,n'}\|_{[\delta,T]^2} = o_P(1)$, $\|q_i\|_{[\delta,T]^2} < \infty$ by continuity, and $V^{\mathrm{HK}}_{[\delta,T]^2}(q_i) < \infty$ by Assumptions A4 and A6. Hence (S.20) is established and it remains to prove the second convergence in (S.19).

By virtue of the first convergence in (S.19), and an analogous result for $\widehat{H}_{n,n'}$ and $\widehat{H}$, it will suffice to prove $\|\widehat{H} - H\|_{A_\delta(\tau,\tau)} \xrightarrow{P} 0$. Note that

$$\begin{aligned} |\widehat{H}(s,t) - H(s,t)| \le{}& \big|\widehat{\mathbf{q}}(s,t)^\mathsf{T} \widehat{\mathbf{I}}_{\delta,T}^{-1}(t) - \mathbf{q}(s,t)^\mathsf{T} \mathbf{I}_{\delta,T}^{-1}(t)\big| \cdot \left| \int_\delta^T \int_t^T \mathbf{q}(s',t')\, \mathrm{d}\eta(s',t') \right| \\ &+ \big|\widehat{\mathbf{q}}(s,t)^\mathsf{T} \widehat{\mathbf{I}}_{\delta,T}^{-1}(t)\big| \cdot \left| \int_\delta^T \int_t^T \big(\widehat{\mathbf{q}}(s,t') - \mathbf{q}(s',t')\big)\, \mathrm{d}\eta(s',t') \right|, \end{aligned} \tag{S.21}$$

where absolute values should be interpreted component-wise, as usual. Consider the first summand on the right-hand side of (S.21). Our assumptions about the various estimators and (S.6) and (S.10) ensure that the difference $\big|\widehat{\mathbf{q}}(s,t)^\mathsf{T} \widehat{\mathbf{I}}_{\delta,T}^{-1}(t) - \mathbf{q}(s,t)^\mathsf{T} \mathbf{I}_{\delta,T}^{-1}(t)\big|$ is $o_P(1)$ uniformly over $(s,t) \in A_\delta(\tau,\tau)$. Moreover, an integration by parts argument as before yields that

$$\left| \int_\delta^T \int_t^T q_i(s',t')\, \mathrm{d}\eta(s',t') \right| \le \|\eta\|_{[\delta,T]^2} \big(4\|q_i\|_{[\delta,T]^2} + V^{\mathrm{HK}}_{[\delta,T]^2}(q_i)\big),$$

for $i = 1,\ldots,8$, where the right-hand side is $O_P(1)$. So the first summand on the right-hand side of (S.21) is $o_P(1)$ uniformly over $(s,t) \in A_\delta(\tau,\tau)$. The second summand there can be handled similarly: the term $\left|\widehat{\mathbf{q}}(s,t)^{\top}\widehat{\mathbf{I}}_{\delta,T}^{-1}(t)\right|$ is $O_P(1)$, and integration by parts yields

$$\left|\int_{\delta}^{T}\int_{t}^{T} q_i(s',t') - \widehat{q}_i(s',t')\,\mathrm{d}\eta(s',t')\right| \leq \|\eta\|_{[\delta,T]^2}\big(4\|q_i - \widehat{q}_i\|_{[\delta,T]^2} + V^{\mathrm{HK}}_{[\delta,T]^2}(q_i - \widehat{q}_i)\big)$$

for $i = 1, \ldots, 8$, where the right-hand side is $o_P(1)$, using (S.6), (S.7), (S.10), and (S.11), in conjunction with Assumptions A4 and A6.

Both convergences in (S.18) are thereby established, which in turn proves (S.15).

Simulations under serial dependence

In Section 7, we argue that our testing procedure is anticonservative in the presence of serial dependence, that is, it will make more Type I errors when there is component-wise serial dependence than when the samples are i.i.d. To provide some empirical support for this claim, we generate 1000 sample pairs from Model III of Section 6.1, but this time with component-wise serial dependence in each sample. We observe that the rejection rates at the 5% level indeed go up for each of the three test statistics, in agreement with our heuristic argument.

To be more precise, we generate a component-wise serially dependent sample from the distribution F in Model III in the following way; the construction is analogous for F'. First we generate two serially dependent sequences of Uniform(0,1) random variables, say $U_1, \ldots, U_n$ and $V_1, \ldots, V_n$, with $n = 1500$. These sequences are generated, independently from each other, following the "sum of uniforms method" described in Section 3 of Willemain and Desautels (1993). This ensures the presence of serial dependence in each sequence, and a parameter c controls the strength of the auto-correlation. We take $c = 1/3$ for a rather strong auto-correlation of about 0.95 at lag 1 in both sequences. At this point, we have a sample $(U_1, V_1), \ldots, (U_n, V_n)$ from the independence copula, with serial dependence in each component. Next, this sample is transformed into a sample from the Clayton(1) survival copula via the conditional distribution method as explained in, e.g., Section 2.9 of

Nelsen (2006). Finally, the Pareto(3) quantile function is applied component-wise to this copula sample, to obtain a serially dependent sample from F.

The marginal auto-correlation plots of one sample generated this way from F can be seen in Fig. 1 below. Clearly, serial dependence is still present in both components after the transformation to the Clayton(1) survival copula and the subsequent marginal transformations to Pareto(3).

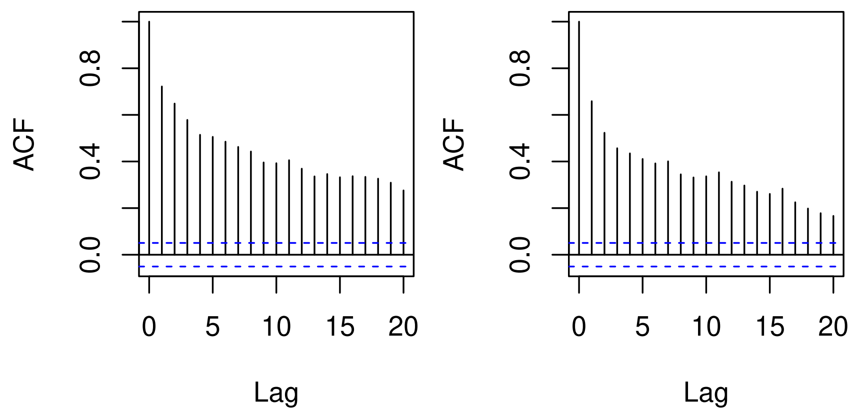

Figure 1: *Marginal auto-correlation plots of a serially dependent sample from F*

Computing the three test statistics in (17) from 1000 sample pairs generated this way from Model III, we observe rejection counts of 88, 99, 120 for κ_n, ω_n^2, A_n^2, respectively, at the 5% level. Thus the empirical sizes are all above the nominal size of 5%, as expected. The PP-plots for the test statistics are provided in Fig. 2 below. The deviation from the 45-degree line is clearly visible for each statistic, especially in comparison with the i.i.d. cases in Fig. 6.1. The plots suggest that the test statistics tend to take larger values under serial dependence than under independence, which means that their critical values would have to be revised *upwards* when accounting for serial dependence in the samples. Thus if the null hypothesis is not rejected under the i.i.d. assumption, it will not be rejected when serial dependence is taken into account.

We also note that changing the dependence parameter c from 1/3 to 1/4 at the initial

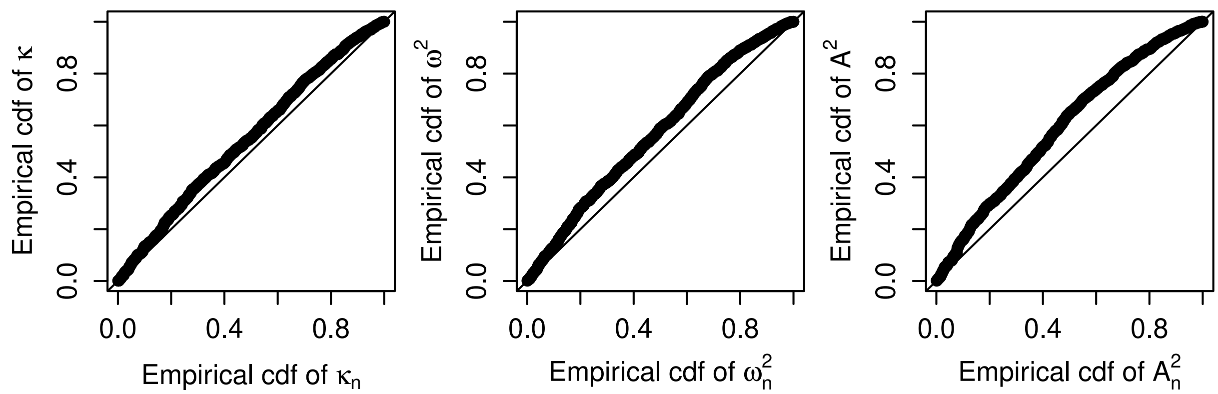

Figure 2: *PP-plots for the three test statistics constructed from 1000 sample pairs with serial dependence*

stage, which leads to even stronger serial dependence in the Uniform(0,1) sequences generated via the sum of uniforms method, results in higher rejection rates for each of the three test statistics: 115, 143, 169 out of 1000 for κ_n, ω_n^2, A_n^2, respectively, at the 5% level. This lends further support to the heuristic idea that higher serial dependence leads to higher estimation errors and therefore higher critical values for the test statistics.

References

Can, S. U., Einmahl, J. H. J., Khmaladze, E. V., and Laeven, R. J. A. (2015). Asymptotically distribution-free goodness-of-fit testing for tail copulas. *Ann. Statist.*, 43(2):878–902.

Henstock, R. (1973). Integration by parts. *Aeq. Math.*, 9:1–18.

Khmaladze, E. V. (1988). An innovation approach to goodness-of-fit tests in R^m. *Ann. Statist.*, 16(4):1503–1516.

Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer.

Owen, A. B. (2005). Multidimensional variation for quasi-Monte Carlo. In Fan, J. and

Li, G., editors, *Contemporary Multivariate Analysis and Design of Experiments*, pages 49–74. World Scientific Publishing.

Willemain, R. T. and Desautels, A. P. (1993). A method to generate autocorrelated uniform random numbers. *J. Stat. Comput. Simul.*, 45(1-2):23–31.