Weighted approximations of tail cupula processes with application to testing the multivariate extreme value condition

Weighted Approximations of Tail Copula Processes with Application to Testing the Multivariate Extreme Value Condition

John H.J. Einmahl — *Tilburg University*
Laurens de Haan — *Erasmus University*
Deyuan Li — *Erasmus University*

3rd August 2004

Abstract. Consider n i.i.d. random vectors on $\mathbb{R}^2$, with unknown, common distribution function F. Under a sharpening of the extreme value condition on F, we derive a weighted approximation of the corresponding tail copula process. Then we construct a test to check whether the extreme value condition holds by comparing two estimators of the limiting extreme value distribution, one obtained from the tail copula process and the other obtained by first estimating the spectral measure which is then used as a building block for the limiting extreme value distribution. We derive the limiting distribution of the test statistic from the aforementioned weighted approximation. This limiting distribution contains unknown functional parameters. Therefore we show that a version with estimated parameters converges weakly to the true limiting distribution. Based on this result, the finite sample properties of our testing procedure are investigated through a simulation study. A real data application is also presented.

Running title: Testing the multivariate EVT condition.

AMS 2000 subject classifications. Primary 62G32, 62G30, 62G10; secondary 60G70, 60F17.

Keywords and phrases. Dependence structure, goodness-of-fit test, multivariate extreme value theory, tail copula process, weighted approximation.

1 Introduction

Let (X,Y), (X_1,Y_1), ..., (X_n,Y_n) be i.i.d. random vectors with continuous distribution function (d.f.) F. Suppose that there exist norming constants $a_n, c_n > 0$ and $b_n, d_n \in \mathbb{R}$ such that the sequence of d.f.'s

$$P\left(\frac{\max_{1\le i\le n} X_i - b_n}{a_n} \le x,\ \frac{\max_{1\le i\le n} Y_i - d_n}{c_n} \le y\right)$$

converges to a limit d.f., say $G(x,y)$, with non-degenerate marginal d.f., that is,

$$\lim_{n\to\infty} F^n(a_n x + b_n,\ c_n y + d_n) = G(x,y) \tag{1.1}$$

for all but countably many x and y. Then, for a suitable choice of a_n, b_n, c_n and d_n, there exist $\gamma_1,\ \gamma_2 \in \mathbb{R}$ such that

$$G(x,\infty) = \exp\left(-(1+\gamma_1 x)^{-1/\gamma_1}\right), \quad G(\infty,y) = \exp\left(-(1+\gamma_2 y)^{-1/\gamma_2}\right).$$

The d.f. G is called an extreme value d.f. and $\gamma_1,\ \gamma_2$ are called the (marginal) extreme value indices.

Any extreme value d.f. G can be represented as

$$G\left(\frac{x^{-\gamma_1}-1}{\gamma_1},\ \frac{y^{-\gamma_2}-1}{\gamma_2}\right) = \exp\left(-\int_0^{\pi/2} (x(1\wedge\tan\theta)) \vee (y(1\wedge\cot\theta))\,\Phi(d\theta)\right), \tag{1.2}$$

with Φ the d.f. of the so-called spectral measure. There is a one-to-one correspondence between extreme value d.f.'s G and finite measures with d.f. Φ that satisfy

$$\int_0^{\pi/2} (1\wedge\tan\theta)\,\Phi(d\theta) = \int_0^{\pi/2} (1\wedge\cot\theta)\,\Phi(d\theta) = 1,$$

via (1.2).

Alternatively one can characterize the extreme value d.f.'s G by: there is a measure Λ on $[0,\infty]^2 \setminus \{(\infty,\infty)\}$ such that, with

$$l(x,y) := -\log G\left(\frac{x^{-\gamma_1}-1}{\gamma_1},\ \frac{y^{-\gamma_2}-1}{\gamma_2}\right), \tag{1.3}$$

we have

(1.4)

1. $l(x,y) = \Lambda\left(\{(u,v)\in[0,\infty]^2 :\ u \le x \textit{ or } v \le y\}\right)$,
2. $l(tx,ty) = t\,l(x,y)$ for $t,x,y > 0$.

Combining the two characterizations we find

$$l(x,y)=\int_0^{\pi/2}(x(1\wedge\tan\theta))\ \vee\ (y(1\wedge\cot\theta))\ \Phi(d\theta). \tag{1.5}$$

Relation (1.1) implies (cf. Einmahl, de Haan and Piterbarg (2001))

$$\lim_{t\downarrow 0} t^{-1}P\left((1-F_1(X))\wedge(1-F_2(Y))\le t,\ 1-F_2(Y)\le(1-F_1(X))\tan\theta\right)=\Phi(\theta) \tag{1.6}$$

for continuity points $\theta\in(0,\pi/2]$ of Φ, where $F_1(x):=F(x,\infty)$ and $F_2(y):=F(\infty,y)$. Also

$$\lim_{t\downarrow 0} t^{-1}P\left(1-F_1(X)\le tx \text{ or } 1-F_2(Y)\le ty\right)=l(x,y) \tag{1.7}$$

for $(x,y)\in[0,\infty)^2$. More generally

$$\lim_{t\downarrow 0} t^{-1}P\left((1-F_1(X),\,1-F_2(Y))\in tA\right)=\Lambda(A) \tag{1.8}$$

for any Borel set A in $[0,\infty]^2\setminus\{(\infty,\infty)\}$ (with $tA:=\{(tx,ty):(x,y)\in A\}$) provided $\Lambda(\partial A)=0$.

A non-parametric estimator for Φ, suggested by the limit relation (1.6) is (Einmahl *et al.* (2001))

$$\hat\Phi(\theta):=\frac{1}{k}\sum_{i=1}^{n} I_{\{R_i^X\vee R_i^Y\ge n+1-k,\ n+1-R_i^Y\le(n+1-R_i^X)\tan\theta\}} \tag{1.9}$$

where R_i^X is the rank of X_i among $X_1,X_2,...,X_n$, R_i^Y is the rank of Y_i among $Y_1,Y_2,...,Y_n$. Similarly a non-parametric estimator for l, suggested by the limit relation (1.7) is (Huang (1992), see also Drees and Huang (1998))

$$\begin{aligned}\hat l_2(x,y) &:= \frac{1}{k}\sum_{i=1}^{n} I_{\{X_i>X_{n+1-\lceil kx\rceil:n}\ \text{or}\ Y_i>Y_{n+1-\lceil ky\rceil:n}\}}\\ &= \frac{1}{k}\sum_{i=1}^{n} I_{\{R_i^X>n+1-kx\ \text{or}\ R_i^Y>n+1-ky\}},\end{aligned} \tag{1.10}$$

where $X_{1:n}\le\cdots\le X_{n:n}$ are the order statistics of the $X_i,\ i=1,2,\ldots,n$ (similarly for the Y_i), with $\lceil z\rceil$ the smallest integer $\ge z$.

The mentioned papers give asymptotic normality results for $\hat\Phi$ and $\hat l_2$ under certain conditions and with sequences $k=k(n)$ satisfying $k(n)\to\infty$, $k(n)/n\to 0$, as $n\to\infty$. Another way of estimating l is via (1.5) and (1.9):

$$\hat l_1(x,y):=\int_0^{\pi/2}(x(1\wedge\tan\theta))\ \vee\ (y(1\wedge\cot\theta))\,\hat\Phi(d\theta). \tag{1.11}$$

The multivariate extreme value framework that we sketched is the appropriate one when one, e.g., wants to estimate the probability of *extreme sets* i.e., sets outside the range of the observations.; see de Haan and Sinha (1999). Condition (1.1) is fulfilled for many standard distributions but not for all distributions. Hence before using this framework to estimate probabilities of extreme sets, it is important to check whether (1.1) is a reasonable assumption for the data set at hand. And one wants to do this beforehand, without specifying the exact structure of the limiting distribution.

A promising approach to this testing problem seems to be to see if the two estimators $\hat{l}_1$ and $\hat{l}_2$ for l, that have a different background, are not too different. The estimator $\hat{l}_2$ is a natural one mimicking more or less the tail of the distribution itself. But this estimator does not necessarily satisfy condition 2 of (1.4). On the other hand $\hat{l}_1$ does satisfy condition 2 of (1.4) but the estimator itself is of a somewhat more complicated nature. So one can maintain that such a test would check whether condition 2 of (1.4) holds.

The proposed test statistic is of Anderson-Darling type:

$$L_n := \iint_{0<x,y\leq 1} \left(\hat{l}_1(x,y) - \hat{l}_2(x,y)\right)^2 (x \vee y)^{-\beta}\, dxdy \tag{1.12}$$

for certain $\beta \geq 0$. The test statistic is similar to those used for testing a parametric null hypothesis (like testing for normality), where the empirical distribution function is compared with the true distribution function with estimated parameters. Here, however, the estimated parameter Φ is a function (and we only deal with the tail of the distribution). Also note that our methods allow us to deal with other test statistics than L_n as well.

Note that this test checks whether the dependence structure is of the right type. It is only based on the relative positions (ranks) of the data and completely independent of the marginal distributions of F for which tests have been developed already in Drees, de Haan and Li (2004) and Dietrich, de Haan and Hüsler (2002).

We shall establish the asymptotic distribution of kL_n as $n \to \infty$ under (1.1) and some extra conditions stemming from Huang (1992) and Einmahl *et al.* (2001), thus providing a basis for applying a test.

Note that the test statistic L_n is based on observations for which at least one component exceeds a certain threshold. Since the estimators depend on this threshold, one can plot L_n as a function of k. This plot can be used as an exploratory tool for determining from which threshold on the two estimators $\hat{l}_1$ and $\hat{l}_2$ are close to each other suggesting that the approximations (1.6) and (1.7) can be trusted, and hence yields a heuristic procedure for determining k. So this a second use of the test statistic L_n.

The weak convergence of kL_n is stated in Theorem 2.3. For the proof of this theorem the known asymptotic normality result for $\hat{\Phi}$ (Einmahl *et al.* (2001)) is sufficient but not the *known* one for $\hat{l}_2$ (Huang (1992)). Hence as a preliminary but important result, we first develop a Gaussian approximation for the weighted tail copula process on $(0,1]^2$

$$\sqrt{k}\left(\hat{l}_2(x,y) - l(x,y)\right)/(x \vee y)^\eta, \quad 0 \le \eta < 1/2,$$

thus extending significantly the result of Huang (1992) where $\eta = 0$. This result, which seems to be useful in other contexts as well, is stated in Theorem 2.2. The proofs are given in section 3.

The limiting random variable in Theorem 2.3 is determined as an integral of a combination of Gaussian processes. They are parametrized by functions which can be estimated consistently. In section 4 it is proved that the probability distribution of the limiting random variable with these functions estimated converges to the distribution of the limiting random variable with these functions equal to the actual ones, which makes the procedure applicable in practice. In section 5 simulation results and an application to real data are reported.

2 Main results

Before stating the main results, we introduce some notation. Define W_Λ to be a Wiener process indexed by the Borel sets in $[0,\infty]^2 \setminus \{(\infty,\infty)\}$, depending on the parameter Λ from (1.4), which is a measure and we assume it has a density λ, in the following way: W_Λ is a centered Gaussian process and for Borel sets C and $\tilde{C}$: $EW_\Lambda(C)W_\Lambda(\tilde{C}) = \Lambda(C \cap \tilde{C})$. Define the sets C_θ by

$$C_\theta = \{(x,y) \in [0,\infty]^2 : \; x \wedge y \le 1, \, y \le x \tan\theta\}, \qquad \theta \in [0, \pi/2],$$

and the process Z by

$$
\begin{aligned}
Z(\theta) &= \int_0^{1\vee\frac{1}{\tan\theta}} \lambda(x, x\tan\theta)(W_1(x)\tan\theta - W_2(x\tan\theta))\,dx \\
&\quad -W_2(1)\int_{1\vee\frac{1}{\tan\theta}}^{\infty} \lambda(x,1)\,dx - I_{(\pi/4,\pi/2]}(\theta)W_1(1)\int_1^{\tan\theta}\lambda(1,y)\,dy, \quad \theta\in[0,\pi/2), \\
Z\left(\frac{\pi}{2}\right) &= -W_2(1)\int_1^{\infty}\lambda(x,1)\,dx \;-\; W_1(1)\int_1^{\infty}\lambda(1,y)\,dy,
\end{aligned}
\tag{2.1}
$$

where λ is the density of Λ, with $W_1(x) = W_\Lambda([0,x]\times[0,\infty])$ and $W_2(y) = W_\Lambda([0,\infty]\times[0,y])$.

Define for $x, y > 0$

$$
W_R(x,y) = W_\Lambda([0,x]\times[0,y]), \quad R(x,y) = \Lambda([0,x]\times[0,y]) \tag{2.2}
$$

and

$$
R_1(x,y) = \partial R(x,y)/\partial x, \quad R_2(x,y) = \partial R(x,y)/\partial y. \tag{2.3}
$$

Theorem 2.1. *Assume that condition (1.8) and Conditions 1 and 2 of Einmahl et al. (2001) hold, and that Λ has a continuous density λ on $[0,\infty)^2\setminus\{(0,0)\}$. Then for a special construction*

$$
\sup_{0<x,y\le 1}\frac{\left|\sqrt{k}(\hat{l}_1(x,y) - l(x,y)) - A(x,y)\right|}{x\vee y} \xrightarrow{P} 0
$$

as $n\to\infty$, where

$$
A(x,y) := \begin{cases} x(W_\Lambda(C_{\frac{\pi}{2}}) + Z(\frac{\pi}{2})) + y\int_{\pi/4}^{\arctan\frac{y}{x}} \frac{1}{\sin^2\theta}(W_\Lambda(C_\theta) + Z(\theta))d\theta, & \text{if } y\ge x, \\ x(W_\Lambda(C_{\frac{\pi}{2}}) + Z(\frac{\pi}{2})) - x\int_{\arctan\frac{y}{x}}^{\pi/4} \frac{1}{\cos^2\theta}(W_\Lambda(C_\theta) + Z(\theta))d\theta, & \text{if } y< x. \end{cases}
$$

Let

$$
U_i = 1 - F_1(X_i), \quad V_i = 1 - F_2(Y_i), \quad i = 1,2,...,n. \tag{2.4}
$$

Let $C(x,y)$ is the distribution function of (U_i, V_i). By (1.8) and (2.2) we have $R(x,y) = \lim_{t\downarrow 0} t^{-1}C(tx,ty)$. We assume, as in Huang (1992), that for some $\alpha > 0$

$$
t^{-1}C(tx,ty) - R(x,y) = O(t^\alpha) \quad \text{as } t\downarrow 0, \tag{2.5}
$$

uniformly for $x\vee y\le 1$, $x, y\ge 0$.

Theorem 2.2. *Assume that conditions (1.8) and (2.5) hold and that* $k = o\left(n^{\frac{2\alpha}{1+2\alpha}}\right)$. *If* R_1 *and* R_2 *are continuous, then we have for* $0 \le \eta < 1/2$ *and for a special construction*

$$\sup_{0<x,y\le 1} \frac{\left|\sqrt{k}(\hat{l}_2(x,y) - l(x,y)) + B(x,y)\right|}{(x \vee y)^\eta} \xrightarrow{P} 0$$

as $n \to \infty$, *where*

$$B(x,y) := W_R(x,y) - R_1(x,y)W_1(x) - R_2(x,y)W_2(y).$$

Theorem 2.3. *Assume the conditions of Theorems 2.1 and 2.2 hold. Then for each* $0 \le \beta < 3$

$$\iint_{0<x,y\le 1} \frac{k\left(\hat{l}_1(x,y) - \hat{l}_2(x,y)\right)^2}{(x \vee y)^\beta} dxdy \xrightarrow{d} \iint_{0<x,y\le 1} \frac{(A(x,y) + B(x,y))^2}{(x \vee y)^\beta} dxdy \tag{2.6}$$

as $n \to \infty$, *and the limit is finite almost surely.*

Remark 2.1. *The case* $\beta = 0$ *is similar to the Cramér-von Mises test. Note that for* $\beta < 2$, *Theorem 2.3 easily follows from an unweighted approximation in Theorems 2.1 and 2.2. Therefore the case* $\beta = 2(!)$ *is similar to the Anderson-Darling test.*

Remark 2.2. *Note that we do not merely test the multivariate extreme value condition but also the refined conditions of Theorem 2.3. Hence we actually test a smaller null hypothesis. But such a smaller hypothesis is needed for statistical applications, since these refined conditions are the ones that yield that the normalized tail of* F *is sufficiently close to* G.

Remark 2.3. *The random variable on the right in Theorem 2.3 has a continuous distribution function. This follows from a property of Gaussian measures on Banach spaces: the measure of a closed ball is a continuous function of its radius, see, e.g., Paulauskas and Račkauskas (1989), Chapter 4, Theorem 1.2.*

Remark 2.4. *Since* $x \vee y \le l(x,y) \le x + y \le 2(x \vee y)$, *(2.6) remains true with* $x \vee y$ *replaced with* $l(x,y)$ *or* $x + y$, *but when choosing* $l(x,y)$, *the left-hand-side of (2.6) is not a statistic and* l *has to be estimated.*

3 Proofs

Before proving Theorem 2.1, we first present two lemmas and a proposition.

Lemma 3.1.

$$l(x,y)=\begin{cases} x\Phi(\frac{\pi}{2})+y\int_{\pi/4}^{\arctan\frac{y}{x}}\frac{1}{\sin^2\theta}\Phi(\theta)d\theta, & \text{if } \; y\geq x,\\ x\Phi(\frac{\pi}{2})-x\int_{\arctan\frac{y}{x}}^{\pi/4}\frac{1}{\cos^2\theta}\Phi(\theta)d\theta, & \text{if } \; y<x.\end{cases}$$

Proof. Since

$$\begin{aligned} l(x,y) &= \int_0^{\pi/2}(x(1\wedge\tan\theta))\,\vee\,(y(1\wedge\cot\theta))\;\Phi(d\theta)\\ &= \int_0^{\pi/4}(x\tan\theta)\,\vee\,y\;\Phi(d\theta)+\int_{\pi/4}^{\pi/2}x\,\vee\,(y\cot\theta)\;\Phi(d\theta)\end{aligned}$$

and

$$x\tan\theta>y\Leftrightarrow x>y\cot\theta\Leftrightarrow\theta>\arctan\frac{y}{x},$$

then

$$\begin{aligned} l(x,y) &= \int_0^{\frac{\pi}{4}\wedge\arctan\frac{y}{x}}y\,\Phi(d\theta)+\int_{\frac{\pi}{4}\wedge\arctan\frac{y}{x}}^{\frac{\pi}{4}}x\tan\theta\,\Phi(d\theta)\\ &\quad+\int_{\pi/4}^{\frac{\pi}{4}\vee\arctan\frac{y}{x}}y\cot\theta\,\Phi(d\theta)+\int_{\frac{\pi}{4}\vee\arctan\frac{y}{x}}^{\pi/2}x\,\Phi(d\theta)\\ &= \begin{cases}\int_0^{\pi/4}y\,\Phi(d\theta)+\int_{\pi/4}^{\arctan\frac{y}{x}}y\cot\theta\,\Phi(d\theta)+\int_{\arctan\frac{y}{x}}^{\pi/2}x\,\Phi(d\theta), & \text{if } y\geq x,\\ \int_0^{\arctan\frac{y}{x}}y\,\Phi(d\theta)+\int_{\arctan\frac{y}{x}}^{\pi/4}x\tan\theta\,\Phi(d\theta)+\int_{\pi/4}^{\pi/2}x\,\Phi(d\theta), & \text{if } y<x.\end{cases}\end{aligned}$$

In case of $y\geq x$, via integration by parts, one has

$$\begin{aligned} l(x,y) &= y\Phi(\frac{\pi}{4})-y\Phi(0)+y\cot(\arctan\frac{y}{x})\Phi(\arctan\frac{y}{x})-y\cot\frac{\pi}{4}\Phi(\frac{\pi}{4})\\ &\quad-y\int_{\pi/4}^{\arctan\frac{y}{x}}\Phi(\theta)(-\frac{1}{\sin^2\theta})d\theta+x\Phi(\frac{\pi}{2})-x\Phi(\arctan\frac{y}{x})\\ &= x\Phi(\frac{\pi}{2})+y\int_{\pi/4}^{\arctan\frac{y}{x}}\frac{1}{\sin^2\theta}\Phi(\theta)d\theta.\end{aligned}$$

In case of $y < x$, via integration by parts again, one has

$$\begin{aligned} l(x,y) &= y\Phi(\arctan\frac{y}{x}) - y\Phi(0) + x\tan\frac{\pi}{4}\Phi(\frac{\pi}{4}) - x\tan(\arctan\frac{y}{x})\Phi(\arctan\frac{y}{x}) \\ &\quad - x\int_{\arctan\frac{y}{x}}^{\pi/4}\Phi(\theta)\frac{1}{\cos^2\theta}d\theta + x\Phi(\frac{\pi}{2}) - x\Phi(\frac{\pi}{4}) \\ &= x\Phi(\frac{\pi}{2}) - x\int_{\arctan\frac{y}{x}}^{\pi/4}\frac{1}{\cos^2\theta}\Phi(\theta)d\theta. \end{aligned}$$

□

Write

$$R_n(x,y) = \frac{n}{k}C\left(\frac{kx}{n},\frac{ky}{n}\right), \quad T_n(x,y) = \frac{1}{k}\sum_{i=1}^{n} I_{\{U_i<\frac{kx}{n},V_i<\frac{ky}{n}\}} \tag{3.1}$$

$$v_n(x,y) = \sqrt{k}(T_n(x,y) - R_n(x,y)), \quad v_{n,\eta}(x,y) = \frac{v_n(x,y)}{(x\vee y)^\eta} \tag{3.2}$$

and

$$v_{n,\eta,1}(x) = \frac{v_n(x,\infty)}{x^\eta}, \quad v_{n,\eta,2}(y) = \frac{v_n(\infty,y)}{y^\eta}, \quad v_{n,j} = v_{n,0,j},\ j=1,2. \tag{3.3}$$

Proposition 3.1. *Let $T > 0$. For $0 \le \eta < 1/2$*

$$(v_{n,\eta}(x,y),\ x,y\in(0,T], \quad v_{n,\eta,1}(x),\ x\in(0,T], \quad v_{n,\eta,2}(y),\ y\in(0,T])$$

converges in distribution to

$$\left(\frac{W_R(x,y)}{(x\vee y)^\eta},\ x,y\in(0,T], \quad \frac{W_1(x)}{x^\eta},\ x\in(0,T], \quad \frac{W_2(y)}{y^\eta},\ y\in(0,T]\right)$$

as $n\to\infty$.

Proof. Define

$$Z_{n,i} = \frac{1}{\sqrt{k}}\delta_{(\frac{n}{k}U_i,\,\frac{n}{k}V_i)}$$

and for all $0 < x, y \le T$ define the functions

$$f_{x,y} = I_{[0,x)\times[0,y)}/(x\vee y)^\eta, \quad f_x^{(1)} = I_{[0,x)\times[0,\infty]}/x^\eta, \quad f_y^{(2)} = I_{[0,\infty]\times[0,y)}/y^\eta.$$

All these f's form the class $\mathcal{F}$. We equip $\mathcal{F}$ with the semi-metric d defined by

$$d(f_{x,y},\ f_{u,v}) = \sqrt{E\left(\frac{W_R(x,y)}{(x\vee y)^\eta} - \frac{W_R(u,v)}{(u\vee v)^\eta}\right)^2},$$

$$d(f_{x,y},\, f_u^{(1)}) = \sqrt{E\left(\frac{W_R(x,y)}{(x\vee y)^\eta} - \frac{W_1(u)}{u^\eta}\right)^2},$$

etc.

For any $\varepsilon > 0$, the bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, L_2^n)$ is the minimal number of sets N_ε in a partition $\mathcal{F} = \bigcup_{j=1}^{N_\varepsilon} \mathcal{F}_{\varepsilon j}$ of the index set into sets $\mathcal{F}_{\varepsilon j}$ such that, for every partitioning set $\mathcal{F}_{\varepsilon j}$

$$\sum_{i=1}^n E^* \sup_{f,g\in\mathcal{F}_{\varepsilon j}} |\mathcal{Z}_{n,i}(f) - \mathcal{Z}_{n,i}(g)|^2 \le \varepsilon^2. \tag{3.4}$$

We will use Theorem 2.11.9 in van der Vaart and Wellner (1996): For each n, let $\mathcal{Z}_{n,1}, \mathcal{Z}_{n,2}, \ldots, \mathcal{Z}_{n,n}$ be independent stochastic processes with finite second moments indexed by a totally bounded semimetric space $(\mathcal{F}, d)$. Suppose

$$\sum_{i=1}^n E^* \|\mathcal{Z}_{n,i}\|_{\mathcal{F}} 1_{\{\|\mathcal{Z}_{n,i}\|_{\mathcal{F}} > \lambda\}} \to 0, \text{ for every } \lambda > 0,$$

where $\|\mathcal{Z}_{n,i}\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}} |\mathcal{Z}_{n,i}(f)|$, and

$$\int_0^{\delta_n} \sqrt{\log N_{[\,]}(\varepsilon, \mathcal{F}, L_2^n)}\, d\varepsilon \to 0, \text{ for every } \delta_n \downarrow 0.$$

Then the sequence $\sum_{i=1}^n (\mathcal{Z}_{n,i} - E\mathcal{Z}_{n,i})$ is asymptotically tight in $\ell^\infty(\mathcal{F})$ and converges weakly, provided the finite-dimensional distributions converge weakly.

We briefly sketch the total boundedness of $(\mathcal{F}, d)$. We only consider the subclass $\mathcal{F}_2$ of $\mathcal{F}$ consisting of the bivariate $f_{x,y}$'s; moreover we restrict ourselves to the case $x \ge y$, $u \ge v$ and $x \ge u$, $y \ge v$. For any $\delta > 0$, assuming $|x - u| \le \delta$ and $|y - v| \le \delta$, one has

$$\begin{aligned} d^2(f_{x,y},\, f_{u,v}) &= E\left(\frac{W_R(x,y)}{(x\vee y)^\eta} - \frac{W_R(u,v)}{(u\vee v)^\eta}\right)^2 \\ &= E\left(\frac{u^\eta W_R(x,y) - x^\eta W_R(u,v)}{(xu)^\eta}\right)^2 \\ &= \frac{u^{2\eta} R(x,y) - 2x^\eta u^\eta R(u,v) + x^{2\eta} R(u,v)}{(xu)^{2\eta}}. \end{aligned}$$

If $u \le \delta$, then

$$\begin{aligned} d^2(f_{x,y}, f_{u,v}) &\le \frac{R(x,y)}{x^{2\eta}} + \frac{2R(u,v)}{u^{2\eta}} + \frac{R(u,v)}{u^{2\eta}} \\ &\le x^{1-2\eta} + 3u^{1-2\eta} \\ &\le (2\delta)^{1-2\eta} + 3\delta^{1-2\eta} \le 5\delta^{1-2\eta}. \end{aligned}$$

If $u > \delta$, then, since

$$\begin{aligned} R(x,y) &\le R(u,v) + \Lambda([u,x]\times[0,\infty]) + \Lambda([0,\infty]\times[v,y]) \\ &\le R(u,v) + 2\delta, \end{aligned}$$

we have

$$\begin{aligned} d^2(f_{x,y}, f_{u,v}) &\le \frac{R(u,v)(u^\eta - x^\eta)^2}{(xu)^{2\eta}} + \frac{2\delta u^{2\eta}}{(xu)^{2\eta}} \\ &\le u^{1-4\eta}(u^\eta - x^\eta)^2 + 2\delta^{1-2\eta} \\ &\le u^{1-4\eta}x^{2\eta-2}(x-u)^2 + 2\delta^{1-2\eta} \\ &\le u^{-1-2\eta}(x-u)^2 + 2\delta^{1-2\eta} \le 3\delta^{1-2\eta}. \end{aligned}$$

So, since $1-2\eta > 0$, we see that for every $\varepsilon > 0$ we can find a $\delta > 0$ such that for $|x-u| \le \delta$ and $|y-v| \le \delta$, $d^2(f_{x,y},\, f_{u,v}) < \varepsilon$. Hence, since $[0,T]^2$ is totally bounded with respect to the Euclidean metric, we obtain the total boundedness of $(\mathcal{F}, d)$.

Observe that

$$Z_{n,i}(f_{x,y}) = \frac{1}{\sqrt{k}} I_{\{U_i < \frac{k}{n}x,\, V_i < \frac{k}{n}y\}}/(x \vee y)^\eta,$$

$$\sum_{i=1}^n (Z_{n,i} - EZ_{n,i})(f_{x,y}) = v_{n,\eta}(x,y)$$

and similarly for the marginal processes. First we have to show that for every $\lambda > 0$

(3.5) $$\sum_{i=1}^n E||Z_{n,i}||_{\mathcal{F}} I_{\{||Z_{n,i}||_{\mathcal{F}} > \lambda\}} \to 0$$

as $n \to \infty$. Again we will restrict ourselves to the subclass $\mathcal{F}_2$. For the univariate $f_x^{(1)}$'s and $f_y^{(2)}$'s, it can be shown in a similar but easier way.

Note that

$$\sup_{f_{x,y}\in\mathcal{F}_2} \frac{1}{\sqrt{k}} I_{\{U_i < \frac{k}{n}x,\, V_i < \frac{k}{n}y\}}/(x \vee y)^\eta \le \frac{1}{\sqrt{k}} \frac{1}{(\frac{n}{k}(U_i \vee V_i))^\eta},$$

so for each $\lambda > 0$

$$\begin{aligned}
&\sum_{i=1}^{n} E||Z_{n,i}||_{\mathcal{F}_2} I_{\{||Z_{n,i}||_{\mathcal{F}_2}>\lambda\}} \\
&\leq \frac{n}{\sqrt{k}} E \frac{1}{(\frac{n}{k}(U_i \vee V_i))^\eta} I_{\{\frac{n}{k}(U_i \vee V_i) < (\sqrt{k}\lambda)^{-1/\eta}\}} \\
&= \frac{n}{\sqrt{k}} \int_0^{(\sqrt{k}\lambda)^{-1/\eta}} x^{-\eta} dC\big(\frac{k}{n}x, \frac{k}{n}x\big) \\
&= \frac{n}{\sqrt{k}} \Big(\sqrt{k}\lambda C\big(\frac{k}{n}(\sqrt{k}\lambda)^{-1/\eta}, \frac{k}{n}(\sqrt{k}\lambda)^{-1/\eta}\big) + \eta \int_0^{(\sqrt{k}\lambda)^{-1/\eta}} C\big(\frac{k}{n}x, \frac{k}{n}x\big) x^{-\eta-1} dx\Big) \\
&\leq \frac{n}{\sqrt{k}} \Big(\sqrt{k}\lambda \frac{k}{n} (\sqrt{k}\lambda)^{-1/\eta} + \eta \int_0^{(\sqrt{k}\lambda)^{-1/\eta}} \frac{k}{n} x^{-\eta} dx\Big) \\
&= \lambda^{1-1/\eta} k^{1-1/(2\eta)} + \sqrt{k} \frac{\eta}{1-\eta} (\sqrt{k}\lambda)^{1-1/\eta} \\
&= \frac{1}{1-\eta} \lambda^{1-1/\eta} k^{1-1/(2\eta)} \to 0, \quad (\eta < 1/2).
\end{aligned}$$

Next we want to show

$$\int_0^{\delta_n} \sqrt{\log\, N_{[\,]}(\varepsilon, \mathcal{F}, L_2^n)}\, d\varepsilon \to 0 \tag{3.6}$$

for every $\delta_n \downarrow 0$. We present the proof for $T = 1$ for notational convenience; for general $T > 0$ the proof is similar. Let $\varepsilon > 0$ be small, define $a = \varepsilon^{3/(1-2\eta)}$ and $\theta = 1 - \varepsilon^3$. We again consider only $\mathcal{F}_2$; the univariate f's are easier to handle. Define

$$\begin{aligned}
\mathcal{F}(a) &= \{f_{x,y} \in \mathcal{F}_2 :\ x \wedge y \leq a\}, \\
\mathcal{F}(l,m) &= \{f_{x,y} \in \mathcal{F}_2 :\ \theta^{l+1} \leq x \leq \theta^l,\ \theta^{m+1} \leq y \leq \theta^m\}.
\end{aligned}$$

Then

$$\mathcal{F}_2 = \mathcal{F}(a) \bigcup \left(\bigcup_{m=0}^{[\frac{\log a}{\log \theta}]} \bigcup_{l=0}^{[\frac{\log a}{\log \theta}]} \mathcal{F}(l,m) \right)$$

First check (3.4) for $\mathcal{F}(a)$:

$$\begin{aligned}
&\sum_{i=1}^{n} E \sup_{f,g \in \mathcal{F}(a)} (Z_{n,i}(f) - Z_{n,i}(g))^2 = nE \sup_{f,g \in \mathcal{F}(a)} (Z_{n,i}(f) - Z_{n,i}(g))^2 \\
\leq\ & 4nE \sup_{f \in \mathcal{F}(a)} Z_{n,i}^2(f) = \frac{4n}{k} E \sup_{\substack{x,y>0 \\ x \wedge y \leq a}} I_{\{U_i < \frac{kx}{n}, V_i < \frac{ky}{n}\}} / (x \vee y)^{2\eta} \\
\leq\ & \frac{4n}{k} E\, (\frac{n}{k} U_i)^{-2\eta} I_{\{\frac{n}{k} U_i < a\}} = \frac{4n}{k} \int_0^{ak/n} (\frac{n}{k} x)^{-2\eta}\, dx = \frac{4}{1-2\eta} a^{1-2\eta} \leq \varepsilon^2.
\end{aligned}$$

Now we consider (3.4) for the $\mathcal{F}(l,m)$; w.l.o.g. we take $l \leq m$:

$$\begin{aligned}
&\sum_{i=1}^{n} E \sup_{f,g\in\mathcal{F}(l,m)} (Z_{n,i}(f) - Z_{n,i}(g))^2 \\
&\leq nE\left(\sup_{f\in\mathcal{F}(l,m)} Z_{n,i}(f) - \inf_{f\in\mathcal{F}(l,m)} Z_{n,i}(f)\right)^2 \\
&\leq \frac{n}{k}E\left(I_{\{U_i<\frac{k}{n}\theta^l,\,V_i<\frac{k}{n}\theta^m\}}/(\theta^{l+1}\vee\theta^{m+1})^\eta - I_{\{U_i<\frac{k}{n}\theta^{l+1},\,V_i<\frac{k}{n}\theta^{m+1}\}}/(\theta^l\vee\theta^m)^\eta\right)^2 \\
&= \frac{n}{k}E\left(I_{\{U_i<\frac{k}{n}\theta^l,\,V_i<\frac{k}{n}\theta^m\}}(\frac{1}{\theta^{\eta(l+1)}} - \frac{1}{\theta^{\eta l}}) + (I_{\{U_i<\frac{k}{n}\theta^l,\,V_i<\frac{k}{n}\theta^m\}} - I_{\{U_i<\frac{k}{n}\theta^{l+1},\,V_i<\frac{k}{n}\theta^{m+1}\}})\frac{1}{\theta^{\eta l}}\right)^2 \\
&\leq \frac{2n}{k}\left(C(\frac{k}{n}\theta^l,\,\frac{k}{n}\theta^m)\frac{1}{\theta^{2\eta l}}(\frac{1}{\theta^\eta}-1)^2 + \left[C(\frac{k}{n}\theta^l,\,\frac{k}{n}\theta^m) - C(\frac{k}{n}\theta^{l+1},\,\frac{k}{n}\theta^{m+1})\right]\frac{1}{\theta^{2\eta l}}\right) \\
&\leq \frac{2n}{k}\left(\frac{k}{n}\frac{\theta^l}{\theta^{2\eta l}}(\frac{1}{\theta^\eta}-1)^2 + \frac{2k}{n}\frac{\theta^l}{\theta^{2\eta l}}(1-\theta)\right) \\
&\leq 2\left(\frac{1}{\theta^{1/2}}-1\right)^2 + 4(1-\theta) \leq \varepsilon^6 + 4\varepsilon^3 \leq \varepsilon^2.
\end{aligned}$$

It is easy to see that the number of elements of the "partition" of $\mathcal{F}_2$ is bounded by ε^{-7}, which yields (3.6). Hence we proved the asymptotic tightness condition.

It remains to prove that the finite-dimensional distributions of our process converge weakly. This follows from the fact that multivariate weak convergence follows from weak convergence of linear combinations of the components and the univariate Lindeberg-Feller central limit theorem. It is easily seen that the Lindeberg condition is satisfied for these linear combinations since the elements of $\mathcal{F}$ are weighted indicators and hence bounded. □

Lemma 3.2. *For* $0 \leq \eta < 1/2$

$$P\left(\sup_{\substack{x\vee y\leq\varepsilon \\ x,y>0}} \frac{|W_R(x,y)|}{(x\vee y)^\eta} \geq \lambda\right) \leq 16\sum_{m=0}^{\infty}\exp\left(-\frac{\lambda^2}{2}\frac{2^{m(1-2\eta)}}{\varepsilon^{1-2\eta}}\right).$$

Proof. For $m = 0,1,2,...$ define

$$\mathcal{A}_m = \{(x,y) : \frac{\varepsilon}{2^{m+1}} \leq x \leq \frac{\varepsilon}{2^m},\ \frac{\varepsilon}{2^{m+1}} \leq y \leq \varepsilon\}.$$

Then, with Z a standard normal random variable,

$$
\begin{aligned}
& P\left(\sup_{\substack{x\vee y\le\varepsilon\\0<x\le y}}\frac{|W_R(x,y)|}{(x\vee y)^\eta}\ge\lambda\right)=P\left(\sup_{\substack{x\vee y\le\varepsilon\\0<x\le y}}\frac{|W_R(x,y)|}{y^\eta}\ge\lambda\right)\\
\le\ & P\left(\sup_{m\in\{0,1,2,\ldots\}}\sup_{(x,y)\in\mathcal{A}_m}\frac{|W_R(x,y)|}{y^\eta}\ge\lambda\right)\le\sum_{m=0}^{\infty}P\left(\sup_{(x,y)\in\mathcal{A}_m}|W_R(x,y)|\ge\lambda\left(\frac{\varepsilon}{2^{m+1}}\right)^\eta\right)\\
\le\ & 4\sum_{m=0}^{\infty}P\left(|W_R(\frac{\varepsilon}{2^m},\varepsilon)|\ge\lambda\left(\frac{\varepsilon}{2^{m+1}}\right)^\eta\right)\le 4\sum_{m=0}^{\infty}P\left(|Z|\ge\frac{\lambda}{2^\eta}\left(\frac{2^m}{\varepsilon}\right)^{1/2-\eta}\right)\\
\le\ & 8\sum_{m=0}^{\infty}\exp\left(-\frac{\lambda^2}{2}\frac{2^{m(1-2\eta)}}{\varepsilon^{1-2\eta}}\right),
\end{aligned}
$$

where the third inequality follows for instance from an adaptation of Lemma 1.2 in Orey and Pruitt (1973) and the last inequality from Mill's ratio. A symmetry argument completes the proof. □

By Theorem 2 in Einmahl *et al.* (2001) and Proposition 3.1 (and their proofs) it follows that

$$
\begin{aligned}
&\left(\sqrt{k}(\hat\Phi(\theta)-\Phi(\theta)),v_{n,\eta}(x,y),v_{n,\eta,1}(u),v_{n,\eta,2}(v)\right)\\
&\xrightarrow{d}\left(W_\Lambda(C_\theta)+Z(\theta)),\frac{W_R(x,y)}{(x\vee y)^\eta},\frac{W_1(u)}{u^\eta},\frac{W_2(v)}{v^\eta}\right),
\end{aligned}
$$

on $D[0,\,\pi/2]\times D[0,\,T]^2\times D[0,T]\times D[0,T]$. By the Skorohod construction, there exists now a probability space carrying $\hat\Phi^*$, v_n^*, $v_{n,1}^*$, $v_{n,2}^*$, $W_\Lambda^*(C_\cdot)$, Z^*, W_R^*, W_1^* and W_2^* such that

$$
\begin{aligned}
\left(\hat\Phi^*,v_n^*,v_{n,1}^*,v_{n,2}^*\right)&\stackrel{d}{=}\left(\hat\Phi,v_n,v_{n,1},v_{n,2}\right),\\
\left(W_\Lambda^*(C_\cdot),Z^*,W_R^*,W_1^*,W_2^*\right)&\stackrel{d}{=}\left(W_\Lambda(C_\cdot),Z,W_R,W_1,W_2\right)
\end{aligned}
$$

and for $0\le\eta<1/2$

$$D_n:=\sup_{0\le\theta\le\pi/2}\left|\sqrt{k}(\hat\Phi^*(\theta)-\Phi(\theta))-(W_\Lambda^*(C_\theta)+Z^*(\theta))\right|=o_P(1), \tag{3.7}$$

$$\sup_{0<x,y\le T}\frac{|v_n^*(x,y)-W_R^*(x,y)|}{(x\vee y)^\eta}=o_P(1), \tag{3.8}$$

$$\sup_{0<x\le T}\frac{|v_{n,1}^*(x)-W_1^*(x)|}{x^\eta}=o_P(1), \tag{3.9}$$

$$\sup_{0<x,y\le T} \frac{|v^*_{n,2}(y) - W^*_2(y)|}{y^\eta} = o_P(1), \tag{3.10}$$

as $n \to \infty$. Henceforth we will work on this probability space, but drop the * from the notation.

Proof of Theorem 2.1. By Lemma 3.1

$$\begin{aligned}
&\sqrt{k}(\hat{l}_1(x,y) - l(x,y)) \\
&= \begin{cases} x\sqrt{k}(\hat\Phi(\frac{\pi}{2}) - \Phi(\frac{\pi}{2})) + y\int_{\pi/4}^{\arctan\frac{y}{x}} \frac{1}{\sin^2\theta}\sqrt{k}(\hat\Phi(\theta) - \Phi(\theta))d\theta, & \text{if } \ y \ge x, \\ x\sqrt{k}(\hat\Phi(\frac{\pi}{2}) - \Phi(\frac{\pi}{2})) - x\int_{\arctan\frac{y}{x}}^{\pi/4} \frac{1}{\cos^2\theta}\sqrt{k}(\hat\Phi(\theta) - \Phi(\theta))d\theta, & \text{if } \ y < x. \end{cases}
\end{aligned}$$

Now, let's first consider the case $y \ge x$.

$$\begin{aligned}
&\sup_{0<x\le y\le 1}\left|\frac{\sqrt{k}(\hat{l}_1(x,y) - l(x,y)) - A(x,y)}{x\vee y}\right| \\
&= \frac{1}{x\vee y}\Bigg| x\left(\sqrt{k}(\hat\Phi^*(\frac{\pi}{2}) - \Phi(\frac{\pi}{2})) - (W^*_\Lambda(C_{\frac{\pi}{2}}) - Z^*(\frac{\pi}{2})\right) \\
&\qquad + y\int_{\pi/4}^{\arctan\frac{y}{x}} \frac{1}{\sin^2\theta}\left(\sqrt{k}(\hat\Phi^*(\theta) - \Phi(\theta)) - (W^*_\Lambda(C_\theta) - Z^*(\theta))\right)d\theta\Bigg| + o_P(1) \\
&\le \frac{xD_n}{x\vee y} + \frac{yD_n}{x\vee y}\int_{\pi/4}^{\pi/2}\frac{1}{\sin^2\theta}d\theta + o_P(1) \to 0,
\end{aligned}$$

in probability as $n \to \infty$. In case of $y < x$, the proof is similar. □

Let Q_{1n} and Q_{2n} be the empirical quantile functions of the $\{U_i\}_{i=1}^n$ and $\{V_i\}_{i=1}^n$, respectively. Define

$$\hat{R}(x,y) = \frac{1}{k}\sum_{i=1}^n I_{\{U_i<Q_{1n}(kx/n),\, V_i<Q_{2n}(ky/n)\}}.$$

Note that by (1.10)

$$\hat{l}_2(x,y) = \frac{1}{k}\sum_{i=1}^n I_{\{U_i<Q_{1n}(kx/n) \ \ or \ \ V_i<Q_{2n}(ky/n)\}}.$$

Proof of Theorem 2.2. It is easily seen that $\hat{l}_2(x,y) + \hat{R}(x,y) = (\lceil kx\rceil + \lceil ky\rceil - 2)/k \le$

$([kx]+[ky])/k$, for each $x, y \in (0,1]$, almost surely. So we have

$$\begin{aligned}
&\sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(\hat{l}_2(x,y)-l(x,y))+\sqrt{k}(\hat{R}(x,y)-R(x,y))|}{(x\vee y)^\eta}\\
&\stackrel{a.s.}{=} \sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{\left|\sqrt{k}(\frac{1}{k}(\lceil kx\rceil+\lceil ky\rceil-2)-(x+y))\right|}{(x\vee y)^\eta}\\
&\le k^{-\eta}\sup_{0<x,y\le 1}\sqrt{k}(x+y-([kx]+[ky])/k)\\
&\le 2\sqrt{k}\cdot k^{\eta-1}=2k^{\eta-1/2}\to 0.
\end{aligned}$$

Write $S_{jn}(x)=\frac{n}{k}Q_{jn}(\frac{k}{n}x)$, $j=1,2$. Then we have

$$\begin{aligned}
&\sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(\hat{l}_2(x,y)-l(x,y))+W_R(x,y)-R_1(x,y)W_1(x)-R_2(x,y)W_2(y)|}{(x\vee y)^\eta}\\
&\stackrel{a.s.}{=} \sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(\hat{R}(x,y)-R(x,y))-W_R(x,y)+R_1(x,y)W_1(x)+R_2(x,y)W_2(y)|}{(x\vee y)^\eta}+o(1)\\
&= \sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(\hat{R}(x,y)-R_n(S_{1n}(x),S_{2n}(y)))-W_R(x,y)|}{(x\vee y)^\eta}\\
&\quad+\sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(R_n(S_{1n}(x),S_{2n}(y)))-R(S_{1n}(x),S_{2n}(y))|}{(x\vee y)^\eta}\\
&\quad+\sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(R(S_{1n}(x),S_{2n}(y))-R(x,y))+R_1(x,y)W_1(x,y)+R_2(x,y)W_2(y)|}{(x\vee y)^\eta}+o(1)\\
&=: D_1+D_2+D_3+o(1).
\end{aligned}$$

We will show that $D_j\to 0$ in probability, $j=1,2,3$. We have

$$\begin{aligned}
D_1 &= \sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(T_n(S_{1n}(x),S_{2n}(y))-R_n(S_{1n}(x),S_{2n}(y)))-W_R(x,y)|}{(x\vee y)^\eta}\\
&\le \sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|\sqrt{k}(T_n(S_{1n}(x),S_{2n}(y))-R_n(S_{1n}(x),S_{2n}(y)))-W_R(S_{1n}(x),S_{2n}(y))|}{(S_{1n}(x)\vee S_{2n}(y))^\eta}\\
&\qquad\cdot\left(\frac{S_{1n}(x)\vee S_{2n}(y)}{x\vee y}\right)^\eta+\sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|W_R(S_{1n}(x),S_{2n}(y))-W_R(x,y)|}{(x\vee y)^\eta}
\end{aligned}$$

$$\leq \sup_{0<s,t\leq 2} \frac{|v_n(s,t) - W_R(s,t)|}{(s\vee t)^\eta} \cdot \sup_{\substack{0<s,t\leq k/n \\ s\vee t\geq 1/n}} \left(\frac{Q_{1n}(s)\vee Q_{2n}(t)}{s\vee t}\right)^\eta$$
$$+ \sup_{\substack{0<x,y\leq 1 \\ x\vee y\geq 1/k}} \frac{|W_R(S_{1n}(x), S_{2n}(y)) - W_R(x,y)|}{(x\vee y)^\eta}$$
$$=: D_{11}\cdot D_{12} + D_{13},$$

where the last inequality holds with arbitrarily high probability. Then $D_{11}\to 0$ in probability because of (3.8) with $T=2$. It is well known that

$$\sup_{s\geq 1/n} \frac{Q_{jn}(s)}{s} = O_P(1), \quad j=1,2 \tag{3.11}$$

(see Shorack and Wellner (1986), p. 419). Hence $D_{11}\cdot D_{12}\to 0$, in probability. Now consider for each $\varepsilon > 1/k$

$$D_{13} \leq \sup_{\substack{0<x,y\leq 1 \\ x\vee y\geq \varepsilon}} \frac{|W_R(S_{1n}(x), S_{2n}(y)) - W_R(x,y)|}{\varepsilon^\eta}$$
$$+ \sup_{\substack{0<x,y\leq 1 \\ 1/k\leq x\vee y\leq \varepsilon}} \frac{|W_R(S_{1n}(x), S_{2n}(y))|}{(S_{1n}(x)\vee S_{2n}(y))^\eta} \cdot \sup_{s,t\geq 1/n} \left(\frac{Q_{1n}(s)\vee Q_{2n}(y)}{s\vee t}\right)^\eta + \sup_{\substack{0<x,y\leq 1 \\ 1/k\leq x\vee y\leq \varepsilon}} \frac{|W_R(x,y)|}{(x\vee y)^\eta}$$
$$=: D_{14} + D_{15} + D_{16}.$$

By the (uniform) continuity of W_R and the fact that

$$\sup_{0<t\leq k/n} \frac{n}{k}|Q_{jn}(t) - t| \to 0, \quad a.s., \quad j=1,2, \tag{3.12}$$

$D_{14}\to 0$ in probability a.s. for any $\varepsilon > 0$. Let $\delta > 0$, by (3.11) and Lemma 3.2 we see that for large n, $P(D_{15}\geq \delta)\leq \delta$ for $\varepsilon > 0$ small enough. Again from Lemma 3.2 we have $P(D_{16}\geq \delta)\leq \delta$. Hence $D_{13}\to 0$ in probability and consequently $D_1\to 0$, in probability.

Consider D_2. Take (a,b) with $a\vee b = u$. Then according to (2.5)

$$\begin{aligned} \frac{1}{t}C(ta,tb) &= \frac{u}{ut}C(tu\frac{a}{u}, tu\frac{b}{u}) \\ &= uR(\frac{a}{u},\frac{b}{u}) + u^{1+\alpha}O(t^\alpha) \\ &= R(a,b) + (a\vee b)^{1+\alpha}O(t^\alpha). \end{aligned}$$

Now with arbitrarily high probability

$$D_2 \leq \sup_{0<x,y\leq 2} \frac{|\sqrt{k}(R_n(x,y) - R(x,y))|}{(x\vee y)^\eta} \cdot \sup_{s\vee t\geq 1/n} \left(\frac{Q_{1n}(s)\vee Q_{2n}(t)}{s\vee t}\right)^\eta.$$

We have seen before that second term of this product is $O_P(1)$. So it is suffices to show that the first term is $o(1)$:

$$\sup_{0<x,y\leq 2}\frac{|\sqrt{k}(R_n(x,y)-R(x,y))|}{(x\vee y)^\eta}=\left(\sup_{0<x,y\leq 2}\frac{\sqrt{k}(x\vee y)^{1+\alpha}}{(x\vee y)^\eta}\right)O\left(\left(\frac{k}{n}\right)^\alpha\right)$$
$$=O\left(\frac{k^{\alpha+1/2}}{n^\alpha}\right)=o(1),$$

by assumption. Hence $D_2\to 0$ in probability.

It remains to show that $D_3\to 0$ in probability. By two applications of the mean-value theorem we obtain

$$\begin{aligned}&R(S_{1n}(x),S_{2n}(y))-R(x,y)\\&=R(S_{1n}(x),S_{2n}(y))-R(x,S_{2n}(y))+R(x,S_{2n}(y))-R(x,y)\\&=R_1(\theta_{1n},S_{2n}(y))(S_{1n}(x)-x)+R_2(x,\theta_{2n})(S_{2n}(y)-y)\end{aligned}$$

with θ_{1n} between x and $S_{1n}(x)$ and θ_{2n} between y and $S_{2n}(y)$. So

$$\begin{aligned}D_3\leq&\sup_{\substack{0<x,y\leq 1\\ x\vee y\geq 1/k}}\frac{|R_1(\theta_{1n},S_{2n}(y))\sqrt{k}(S_{1n}(x)-x)+R_1(x,y)W_1(x)|}{(x\vee y)^\eta}\\&+\sup_{\substack{0<x,y\leq 1\\ x\vee y\geq 1/k}}\frac{|R_2(x,\theta_{2n})\sqrt{k}(S_{2n}(y)-y)+R_2(x,y)W_2(y)|}{(x\vee y)^\eta}.\end{aligned}$$

We consider only the first term in the right hand side of this expression; the second one can be dealt with similarly. Write $z_n(x)=\sqrt{k}(S_{1n}(x)-x)$. From (3.9) with $\eta=0$ it follows that

$$\sup_{0<x\leq 1}|z_n(x)+W_1(x)|\to 0$$

in probability. From this it can be shown that for $0\leq\eta<1/2$

$$\sup_{1/k\leq x\leq 1}\frac{|z_n(x)+W_1(x)|}{x^\eta}\to 0 \tag{3.13}$$

in probability (see, e.g., Einmahl (1992)). Now

$$\begin{aligned}
&\sup_{\substack{0<x,y\le 1\\ x\vee y\ge 1/k}} \frac{|R_1(\theta_{1n},S_{2n}(y))z_n(x)+R_1(x,y)W_1(x)|}{(x\vee y)^\eta}\\
&\le \sup_{0<x,y\le 1} R_1(\theta_{1n},S_{2n}(y))\cdot \sup_{1/k\le x\le 1}\frac{|z_n(x)+W_1(x)|}{x^\eta}\\
&\quad + \sup_{0<x,y\le 1}|R_1(x,y)-R_1(\theta_{1n},S_{2n}(y))|\cdot \sup_{0<x\le 1}\frac{|W_1(x)|}{x^\eta}\\
&=: D_{31}+D_{32}.
\end{aligned}$$

Since R_1 is continuous on $[0,2]^2$ it is uniformly continuous and bounded. This together with (3.13) yields $D_{31}\to 0$ in probability. The uniform continuity of R_1 together with (3.12) and the fact that

$$\sup_{0<x\le 1}\frac{|W_1(x)|}{x^\eta}<\infty \quad a.s.,$$

yields $D_{32}\to 0$ in probability and consequently $D_3\to 0$ in probability.

Finally we show that

$$\sup_{0<x,y<1/k}\frac{|\sqrt{k}(\hat{l}_2(x,y)-l(x,y))+B(x,y)|}{(x\vee y)^\eta}=o_P(1).$$

Observing that $\sup_{0<x,y<1/k}\hat{l}_2(x,y)=0$ a.s., this follows easily. □

Proof of Theorem 2.3. For each $0\le\beta<3$, there exist $\alpha\in[0,2)$ and $\eta\in[0,1/2)$ such that $\beta=\alpha+2\eta$. By Theorem 2.1 and Theorem 2.2, and

$$\int_0^1\int_0^1\frac{1}{(x\vee y)^\alpha}\,dxdy<\infty,$$

it follows that as $n\to\infty$

$$\begin{aligned}
&\iint_{0<x,y\le 1}\frac{k\left(\hat{l}_1(x,y)-\hat{l}_2(x,y)\right)^2}{(x\vee y)^\beta}\,dxdy\\
&=o_P(1)\iint_{0<x,y\le 1}\frac{1}{(x\vee y)^\alpha}\,dxdy+\iint_{0<x,y\le 1}\frac{\left(A(x,y)+B(x,y)\right)^2}{(x\vee y)^\beta}\,dxdy\\
&\xrightarrow{d}\iint_{0<x,y\le 1}\frac{\left(A(x,y)+B(x,y)\right)^2}{(x\vee y)^\beta}\,dxdy\,.
\end{aligned}$$

□

4 Approximating the limit

For testing purposes, we have to find the probability distribution of the limiting random variable in Theorem 2.3. This can be done by simulating the processes A and B, but unfortunately their distributions depend on the unknown measure Λ. Therefore, we generate approximations A_n and B_n, respectively, of the processes A and B, not with parameter Λ but with approximated parameter Λ_n. In this section, we consider the convergence of the sequence of these approximated limiting random variables. Until further notice, we take $\{\Lambda_n\}_{n\geq 1}$ to be a sequence of *deterministic* measures.

Define

$$\begin{aligned}
R_{1n}(x,y) &:= \frac{1}{2}k^{1/5}\Lambda_n([x-k^{-1/5}, x+k^{-1/5}]\times[0,y)),\\
R_{2n}(x,y) &:= \frac{1}{2}k^{1/5}\Lambda_n([0,x)\times[y-k^{-1/5}, y+k^{-1/5}]),\\
W_{1n}(x) &:= W_{\Lambda_n}([0,x]\times[0,\infty]), \quad W_{2n}(y) := W_{\Lambda_n}([0,\infty]\times[0,y]),\\
W_{R_n}(x,y) &:= W_{\Lambda_n}([0,x]\times[0,y]),
\end{aligned}$$

and the process B_n by

$$B_n(x,y) := W_{R_n}(x,y) - R_{1n}(x,y)W_{1n}(x) - R_{2n}(x,y)W_{2n}(y).$$

Based on the definition of Z in (2.1) and the homogeneity property of λ (i.e., $\lambda(tx,ty) = \frac{1}{t}\lambda(x,y)$), we define the approximating process Z_n by

(4.1)

$$Z_n(\theta) = \begin{cases}
\lambda_n(1,\tan\theta)\tan\theta \displaystyle\int_0^{1/\tan\theta}\frac{W_{1n}(x)}{x}dx - \lambda_n(1,\tan\theta)\int_0^1\frac{W_{2n}(x)}{x}dx & \\
\qquad - W_{2n}(1)\displaystyle\int_{1/\tan\theta}^{\infty}\lambda_n(x,1)dx, & \theta\in[0,\pi/4]\\
\lambda_n(1/\tan\theta,1)\displaystyle\int_0^1\frac{W_{1n}(x)}{x}dx - \lambda_n(1/\tan\theta,1)\frac{1}{\tan\theta}\int_0^{\tan\theta}\frac{W_{2n}(x)}{x}dx & \\
\qquad - W_{2n}(1)\displaystyle\int_1^{\infty}\lambda_n(x,1)dx - W_{1n}(1)\int_1^{\tan\theta}\lambda_n(1,y)dy, & \theta\in(\pi/4,\pi/2)\\
-W_{2n}(1)\displaystyle\int_1^{\infty}\lambda_n(x,1)dx - W_{1n}(1)\int_1^{\infty}\lambda_n(1,y)dy, & \theta=\pi/2
\end{cases}$$

where λ_n is the approximation of λ defined by

$$\lambda_n(1,y) := \frac{1}{4}k^{1/3}\Lambda_n([1-k^{-1/6},1+k^{-1/6}]\times[y-k^{-1/6},y+k^{-1/6}]), \quad y>0,$$
$$\lambda_n(x,1) := \frac{1}{4}k^{1/3}\Lambda_n([x-k^{-1/6},x+k^{-1/6}]\times[1-k^{-1/6},1+k^{-1/6}]), \quad x>0.$$

Finally define the process A_n by

$$A_n(x,y) := \begin{cases} x(W_{\Lambda_n}(C_{\frac{\pi}{2}})+Z_n(\frac{\pi}{2}))+y\int_{\pi/4}^{\arctan\frac{y}{x}}\frac{1}{\sin^2\theta}(W_{\Lambda_n}(C_\theta)+Z_n(\theta))d\theta & \text{if } y\ge x, \\ x(W_{\Lambda_n}(C_{\frac{\pi}{2}})+Z_n(\frac{\pi}{2}))-x\int_{\arctan\frac{y}{x}}^{\pi/4}\frac{1}{\cos^2\theta}(W_{\Lambda_n}(C_\theta)+Z_n(\theta))d\theta & \text{if } y<x. \end{cases}$$

First we consider the weak convergence of the weighted approximating processes. We write $D_2 := D([0,1]^2)$ for the generalization of $D[0,1]$ to dimension 2, and $\mathcal{L}_d$ for the Borel σ-algebra on (D_2,d), where d is the metric on D_2 defined in Neuhaus (1971).

Proposition 4.1. *Let Λ be as in Theorem 2.3. Suppose that $\{\Lambda_n\}_{n\ge1}$ is a sequence of measures on $[0,\infty]^2\setminus\{(\infty,\infty)\}$ satisfying that for each $x,y\ge0$*

$$\Lambda_n([0,x]\times[0,\infty]) = [kx]/k, \quad \Lambda_n([0,\infty]\times[0,y]) = [ky]/k \tag{4.2}$$

and

$$\sup_{0<x,y\le1}|\Lambda_n([0,x]\times[0,y])-\Lambda([0,x]\times[0,y])|\to0 \tag{4.3}$$

as $n\to\infty$. Further suppose that

$$\sup_{0<x\le1}|\lambda_n(x,1)-\lambda(x,1)|\to0, \quad \sup_{0<y\le1}|\lambda_n(1,y)-\lambda(1,y)|\to0, \tag{4.4}$$

$$\sup_{0<x,y\le1}|R_{jn}(x,y)-R_j(x,y)|\to0, \quad j=1,2, \tag{4.5}$$

as $n\to\infty$. Then for each $0\le\eta<1/2$

$$\left\{\frac{A_n(x,y)+B_n(x,y)}{(x\vee y)^\eta},(x,y)\in[0,1]^2\right\} \to \left\{\frac{A(x,y)+B(x,y)}{(x\vee y)^\eta},(x,y)\in[0,1]^2\right\},$$

weakly in D_2.

Before proving this proposition, we present three corollaries. The last one is the main result of this section.

Corollary 4.1. *Under the conditions of Proposition 4.1 for each* $0 \le \beta < 3$

$$\iint_{0<x,y\le 1} \frac{(A_n(x,y)+B_n(x,y))^2}{(x\vee y)^\beta}\,dxdy \;\xrightarrow{d}\; \iint_{0<x,y\le 1} \frac{(A(x,y)+B(x,y))^2}{(x\vee y)^\beta}\,dxdy \tag{4.6}$$

as $n \to \infty$.

Let Q_{Λ_n} be the quantile function of the random variable on the left hand side of (4.6) and Q_Λ the quantile function of the random variable on the right hand side of (4.6).

Corollary 4.2. *Under the conditions of Proposition 4.1, for each* $0 \le \beta < 3$ *and for each continuity point* $1-\alpha$ *(*$0<\alpha<1$*) of* Q_Λ*,*

$$\lim_{n\to\infty} Q_{\Lambda_n}(1-\alpha) = Q_\Lambda(1-\alpha).$$

Next, with abuse of notation, we estimate Λ_n from the data, so it becomes random. In Einmahl *et al.* (2001), Λ_n is defined as

$$\begin{aligned} \Lambda_n(A) &:= \frac{1}{k}\sum_{i=1}^n I_{\frac{k}{n}A}\left(\frac{1}{n}\sum_{j=1}^n I_{(-\infty,U_i]}(U_j),\ \frac{1}{n}\sum_{j=1}^n I_{(-\infty,V_i]}(V_j)\right) \\ &= \frac{1}{k}\sum_{i=1}^n I_{kA}\left(n+1-R_i^X, n+1-R_i^Y\right) \end{aligned} \tag{4.7}$$

where $U_i := 1-F_1(X_i)$, $V_i := 1-F_2(Y_i)$ for $i=1,2,...,n$. Note that for $x,y>0$

$$\Lambda_n([0,x)\times[0,y)) = \frac{1}{k}\sum_{i=1}^n I_{\{U_i<Q_{1n}(kx/n),\,V_i<Q_{2n}(ky/n)\}}.$$

So $\Lambda_n([0,x)\times[0,\infty]) = (\lceil kx\rceil - 1)/k \le [kx]/k = \Lambda_n([0,x]\times[0,\infty])$ a.s. and $\Lambda_n([0,\infty]\times[0,y)) = (\lceil ky\rceil - 1)/k \le [ky]/k = \Lambda_n([0,\infty]\times[0,y])$ a.s.

The final and main corollary deals with the *random* measures Λ_n, where the functions derived from Λ_n, like λ_n, are defined as before. In particular, we define Q_{Λ_n}, as the quantile function of the random variable on the left hand side of (4.6), *conditional* on Λ_n, so it is also random.

Corollary 4.3. *Let* Λ_n *be as in (4.7). Under the conditions of Theorem 2.3, we have for each* $0\le\beta<3$ *and each continuity point* $1-\alpha$ *(*$0<\alpha<1$*) of* Q_Λ*, that*

$$Q_{\Lambda_n}(1-\alpha) \xrightarrow{P} Q_\Lambda(1-\alpha), \quad \text{as } n\to\infty.$$

For testing purposes, Corollary 4.3 shows that simulation of the limiting random variable in Theorem 2.3 with Λ replaced with the estimated Λ_n is asymptotically correct.

Now we turn to the proofs. In order to prove Proposition 4.1, by Prohorov's theorem it is necessary and sufficient to prove that
(i) The finite-dimensional distributions of $\{(A_n(x,y)+B_n(x,y))/(x\vee y)^\eta,\ (x,y)\in[0,1]^2\}_{n\geq 1}$ converge to those of $\{(A(x,y)+B(x,y))/(x\vee y)^\eta,\ (x,y)\in[0,1]^2\}$,
(ii) $\{(A_n(x,y)+B_n(x,y))/(x\vee y)^\eta,\ (x,y)\in[0,1]^2\}_{n\geq 1}$ is relatively compact.

For the relative compactness, we need several lemmas. First we present in Lemma 4.1 sufficient conditions for relative compactness ; the proof is similar to that of Theorem 15.5 in Billingsley (1968), see also Neuhaus (1971).

Lemma 4.1. *Let P_n be probability measures on $(D_2,\mathcal{L}_d)$. Suppose that, for each positive η, there exists an $M>0$ such that*

$$P_n(x\in D_2 : |x(0,0)|>M)\leq\eta,\quad n\geq 1.$$

Suppose further that, for each positive ε and η, there exist a δ, $0<\delta<1$, and an integer n_0 such that

$$P_n(x\in D_2 : \sup_{|u_1-u_2|\leq\delta,|v_1-v_2|\leq\delta}|x(u_1,v_1)-x(u_2,v_2)|>\varepsilon)\leq\eta,\quad n\geq n_0.$$

Then $\{P_n\}_{n\geq 1}$ is relatively compact.

Lemma 4.2. *Under the conditions of Proposition 4.1, for each c, $a>0$*

(i) $\int_0^c \frac{W_{jn}(t)}{t}dt\sim N(0,\sigma_n^2)$, with $\sigma_n^2\leq 2c$, $\quad j=1,2$,

(ii) $P(\sup_{t\geq c}|\frac{W_{jn}(t)}{t}|\geq a)\leq 2P(|W(2/c)|\geq a)$, $j=1,2$, where W is a standard Wiener process.

Proof. (i) This follows from Proposition 1, page 42, in Shorack and Wellner (1986).

(ii) Let W be a standard Wiener process. Since $\{W(t)/t, t\geq c\} =^d \{W(1/t), t\geq c\}$, then

$$P(\sup_{t\geq c}|W(t)/t|\geq a) = P(\sup_{0<s\leq 1/c}|W(s)|\geq a)\leq 2P(|W(1/c)|\geq a).$$

Write $\Lambda_{1n}(t)$ for $\Lambda_n([0,t]\times[0,\infty])$. Since $\{W_{1n}(t), t>0\} \stackrel{d}{=} \{W(\Lambda_{1n}(t)), t>0\}$, then

$$
\begin{aligned}
P(\sup_{t\geq c}|W_{1n}(t)/t|\geq a) &= P(\sup_{t\geq c}\left|\frac{W(\Lambda_{1n}(t))\cdot\Lambda_{1n}(t)}{\Lambda_{1n}(t)\cdot t}\right|\geq a) \\
\leq & P(\sup_{\Lambda_{1n}(t)\geq c/2}|W(\Lambda_{1n}(t))/\Lambda_{1n}(t)|\geq a)\leq 2P(|W(2/c)|\geq a),
\end{aligned}
$$

eventually (since $t-1/k\leq\Lambda_{1n}(t)\leq t$). For $j=2$ the proof is the same. □

Lemma 4.3. *Define*

$$H_n := \sup_{\theta\in[0,\pi/2]}|W_{\Lambda_n}(C_\theta)+Z_n(\theta)|.$$

Then under the conditions of Proposition 4.1, there exists an $n_0\in\mathbb{N}$ such that

$$\sup_{n\geq n_0}P(H_n\geq a)=O(e^{-a}) \text{ as } a\to\infty.$$

Proof. Define $H_{1n}:=\sup_{\theta\in[0,\pi/4]}|W_{\Lambda_n}(C_\theta)+Z_n(\theta)|$, $H_{2n}:=\sup_{\theta\in(\pi/4,\pi/2)}|W_{\Lambda_n}(C_\theta)+Z_n(\theta)|$, and $H_{3n}:=|W_{\Lambda_n}(C_{\pi/2})+Z_n(\pi/2)|$. It suffices to verify that there exists an $n_0\in\mathbb{N}$ such that

$$\sup_{n\geq n_0}P(H_{jn}\geq a)=O(e^{-a}),\quad j=1,2,3$$

as $a\to\infty$. Here we only check it in case of $j=1$. For the other two cases, the proofs are similar.

Since for all $n\geq 1$

$$\{W_{\Lambda_n}(C_\theta),\theta\in[0,\pi/2]\} \stackrel{d}{=} \{W(\Lambda_n(C_\theta)),\theta\in[0,\pi/2]\},$$

with W a standard Wiener process, we have

$$
\begin{aligned}
P(H_{1n}\geq a) &\leq P(\sup_{\theta\in[0,\pi/4]}|W(\Lambda_n(C_\theta))|\geq a/2)+P(\sup_{\theta\in[0,\pi/4]}|Z_n(\theta)|\geq a/2) \\
&\leq 2P(|W(\Lambda_n(C_{\pi/4}))|\geq a/2)+P(\sup_{\theta\in[0,\pi/4]}|Z_n(\theta)|\geq a/2).
\end{aligned}
$$

Clearly $\Lambda_n(C_{\pi/4})\leq 1$ for all $n\geq 1$, and hence $\sup_{n\geq 1}P(|W(\Lambda_n(C_{\pi/4}))|\geq a/2)=O(e^{-a})$, as $a\to\infty$.

From Einmahl *et al.* (2001), one has $\sup_{x>0}\lambda(x,1)<\infty$ and $\sup_{y>0}\lambda(1,y)<\infty$. Then by (4.4) there exists a constant $\lambda_0>0$ such that $\sup_{0<x\leq 1}\lambda_n(x,1)<\lambda_0$ and $\sup_{0<y\leq 1}\lambda_n(1,y)<\lambda_0$ for large n. Using (4.2) and the fact that Λ_n is a step function, one can prove with some effort that $\int_1^\infty \lambda_n(x,1)dx \le 2$ and $\int_1^\infty \lambda_n(1,y)dy \le 2$ for sufficiently large n, hence by the definition of $Z_n(\theta)$, one has

$$\begin{aligned}
&\sup_{\theta\in[0,\pi/4]} |Z_n(\theta)| \\
&\le \lambda_0 \left|\int_0^1 \frac{W_{1n}(x)}{x}dx\right| + \lambda_0 \sup_{\theta\in[0,\pi/4]} \left|\tan\theta \int_1^{1/\tan\theta} \frac{W_{1n}(x)}{x}dx\right| + \lambda_0 \left|\int_0^1 \frac{W_{2n}(x)}{x}dx\right| + 2\left|W_{2n}(1)\right| \\
&\le \lambda_0 \left|\int_0^1 \frac{W_{1n}(x)}{x}dx\right| + \lambda_0 \sup_{x\ge 1} \left|\frac{W_{1n}(x)}{x}\right| + \lambda_0 \left|\int_0^1 \frac{W_{2n}(x)}{x}dx\right| + 2\left|W_{2n}(1)\right|,
\end{aligned}$$

for sufficiently large n. By Lemma 4.2(i), $\int_0^1 \frac{W_{1n}(x)}{x}dx$ and $\int_0^1 \frac{W_{2n}(x)}{x}dx$ have centered normal distributions with uniformly bounded variances for all $n \ge 1$. By Lemma 4.2(ii) there exist an $n_0 \in \mathbb{N}$ such that

$$\sup_{n\ge n_0} P(\lambda_0 \sup_{x\ge 1} |W_{1n}(x)|/x \ge a/8) \le 2P(W(2) \ge a/(8\lambda_0)) = O(e^{-a})$$

as $a \to \infty$. Hence

$$\sup_{n\ge n_0} P(\sup_{\theta\in[0,\pi/4]} |Z_n(\theta)| \ge a/2) = O(e^{-a})$$

as $a \to \infty$. So $\sup_{n\ge n_0} P(H_{1n} \ge a) = O(e^{-a})$ as $a \to \infty$. □

Lemma 4.4. *Under the conditions of Proposition 4.1, for each* $0 \le \eta < 1/2$

$$\left\{\frac{B_n(x,y)}{(x\vee y)^\eta}, (x,y) \in [0,1]^2\right\}_{n\ge 1}$$

is relatively compact.

Proof. By the definition of R_{1n} and R_{2n}, one has

$$\begin{aligned}
R_{1n}(x,y) &= \frac{1}{2}k^{1/5}\Lambda_n([x-k^{-1/5}, x+k^{-1/5}] \times [0,\infty]) \\
&= \frac{1}{2}k^{1/5}\left(\frac{[k(x+k^{-1/5})]}{k} - \frac{[k(x-k^{-1/5})]}{k}\right) \\
&\le 1 + 1/k^{4/5} \le 2 \quad \text{ if } \; k \ge 1.
\end{aligned}$$

Also $R_{2n}(x,y) \le 2$ for $k \ge 1$. Hence it is sufficient to prove

$$\{W_{R_n}(x,y)/(x\vee y)^\eta, x,y \in [0,1]\}_{n\ge1}, \quad \{W_{1n}(x)/x^\eta, x \in [0,1]\}_{n\ge1}, \quad \{W_{2n}(y)/y^\eta, y \in [0,1]\}_{n\ge1}$$

are relatively compact. Here we only show the proof of the first one. The proofs of the others are similar.

Setting $0/0 = 0$, by Lemma 4.1 it suffices to prove that for each positive ε, there exist a $\delta\,(0 < \delta < 1)$ and $n_0 \in \mathbb{N}$ (n_0 may depend on δ) such that

$$(4.8)\qquad P\left(\sup_{\substack{x,y,u,v\in[0,1]\\|x-u|\le\delta,|y-v|\le\delta}}\left|\frac{W_{\Lambda_n}([0,x]\times[0,y])}{(x\vee y)^\eta}-\frac{W_{\Lambda_n}([0,u]\times[0,v])}{(u\vee v)^\eta}\right|>\varepsilon\right)\le\varepsilon,\quad n\ge n_0.$$

We partition the square $[0,1]\times[0,1]$ into m^2 ($m \in \mathbb{N}$) small squares, say $[0,1]\times[0,1] = \bigcup_{i=1}^m\bigcup_{j=1}^m \Delta_{ij}$, with $\Delta_{ij} := \{(x,y) : i\delta \le x \le (i+1)\delta, j\delta \le y \le (j+1)\delta\}$, $\delta := 1/m$ and $i,j = 0,1,...,m-1$. In order to prove (4.8), it suffices to prove that for each positive ε, there exist a $\delta\,(0<\delta<1)$ and $n_0 = n_0(\delta) \in \mathbb{N}$ such that

$$(4.9)\qquad \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} P\left(\sup_{\Delta_{ij}}\left|\frac{W_{\Lambda_n}([0,x]\times[0,y])}{(x\vee y)^\eta}-\frac{W_{\Lambda_n}([0,i\delta]\times[0,j\delta])}{\delta^\eta(i\vee j)^\eta}\right|>\varepsilon\right)\le\varepsilon,\quad n\ge n_0.$$

We consider the case $i\vee j\ge 1$ and the case $i=j=0$ separately. Let's first look at the case $i\vee j \ge 1$. Assume $i>j$. Let $S(x,y) := [0,x]\times[0,y]$. Note that for $(x,y)\in\Delta_{ij}$

$$\begin{aligned}
&\left|\frac{W_{\Lambda_n}([0,x]\times[0,y])}{(x\vee y)^\eta}-\frac{W_{\Lambda_n}([0,i\delta]\times[0,j\delta])}{\delta^\eta(i\vee j)^\eta}\right|\\
&=\left|\frac{W_{\Lambda_n}(S(x,y))}{x^\eta}-\frac{W_{\Lambda_n}(S(i\delta,j\delta))}{(i\delta)^\eta}\right|\\
&=\frac{|(i\delta)^\eta W_{\Lambda_n}(S(i\delta,j\delta))+(i\delta)^\eta W_{\Lambda_n}(S(x,y)\backslash S(i\delta,j\delta))-x^\eta W_{\Lambda_n}(S(i\delta,j\delta))|}{x^\eta(i\delta)^\eta}\\
&\le\frac{|(i\delta)^\eta W_{\Lambda_n}(S(x,y)\backslash S(i\delta,j\delta))-(x^\eta-(i\delta)^\eta)W_{\Lambda_n}(S(i\delta,j\delta))|}{(i\delta)^{2\eta}}
\end{aligned}$$

(since $x\ge i\delta\ge y$). Hence

$$\begin{aligned}
&P\left(\sup_{\Delta_{ij}}\left|\frac{W_{\Lambda_n}([0,x]\times[0,y])}{(x\vee y)^\eta}-\frac{W_{\Lambda_n}([0,i\delta]\times[0,j\delta])}{\delta^\eta(i\vee j)^\eta}\right|>\varepsilon\right)\\
&\le P(\sup_{\Delta_{ij}}|\frac{W_{\Lambda_n}(S(x,y)\backslash S(i\delta,j\delta))}{(i\delta)^\eta}|>\frac{\varepsilon}{2})+P(\sup_{\Delta_{ij}}|\frac{x^\eta-(i\delta)^\eta}{(i\delta)^{2\eta}}W_{\Lambda_n}(S(i\delta,j\delta))|>\frac{\varepsilon}{2})\\
&\le 4P(|\frac{W_{\Lambda_n}(S((i+1)\delta,(j+1)\delta)\backslash S(i\delta,j\delta))}{(i\delta)^\eta}|>\frac{\varepsilon}{4})+P(|\frac{(1+1/i)^\eta-1}{(i\delta)^\eta}W_{\Lambda_n}(S(i\delta,j\delta))|>\frac{\varepsilon}{2}).
\end{aligned}$$

Since $\Lambda_n(S((i+1)\delta,(j+1)\delta)\backslash S(i\delta,j\delta))\le 2\delta+4/k$ for all $i\vee j\ge 1$, there exist $n_* = n_*(\delta)\in\mathbb{N}$ such that $k_* = k(n_*)\ge 1/\delta$ and hence

$$\Lambda_n(S((i+1)\delta,(j+1)\delta)\backslash S(i\delta,j\delta))\le 6\delta,\quad n\ge n_*.$$

uniformly in $i \vee j \geq 1$. It follows that $(i\delta)^{-\eta} W_{\Lambda_n}(S((i+1)\delta,(j+1)\delta)\backslash S(i\delta,j\delta))$ has a normal distribution with mean zero and variance $\sigma_n^2(i,j)$ satisfying $\sigma_n^2(i,j) \leq 6\delta^{1-2\eta}$ for all $i > j, i \geq 1$, and $n \geq n_*$. Hence for all $\varepsilon > 0$

$$\sup_{n\geq n_*}\sup_{i>j,\, i\geq 1} P(|(i\delta)^{-\eta} W_{\Lambda_n}(S((i+1)\delta,(j+1)\delta)\backslash S(i\delta,j\delta))| > \varepsilon/4) = O(e^{-\delta^{\eta-1/2}})$$

as $\delta \to 0$. On the other hand, note that $\frac{(1+1/i)^\eta - 1}{(i\delta)^\eta} W_{\Lambda_n}(S(i\delta,j\delta))$ has a normal distribution with mean zero and variance $\tilde{\sigma}_n^2(i,j)$ satisfying $\tilde{\sigma}_n^2(i,j) \leq (i\delta)^{1-2\eta}((1+1/i)^\eta - 1)^2 \leq 4\delta^{1-2\eta}$. So

$$\sup_{n\geq n_*}\sup_{i>j,\, i\geq 1} P(|\frac{(1+1/i)^\eta - 1}{(i\delta)^\eta} W_{\Lambda_n}(S(i\delta,j\delta))| > \varepsilon/2) = O(e^{-\delta^{\eta-1/2}})$$

as $\delta \to 0$.

In case of $j > i, j \geq 1$ and case of $i = j \geq 1$, we can get similar results as above. Hence

(4.10)
$$\sup_{n\geq n_*}\sum_{i\vee j\geq 1}^{m-1} P\left(\sup_{\Delta_{ij}}\left|\frac{W_{\Lambda_n}([0,x]\times[0,y])}{(x\vee y)^\eta} - \frac{W_{\Lambda_n}([0,i\delta]\times[0,j\delta])}{\delta^\eta (i\vee j)^\eta}\right| > \varepsilon\right) = O(\delta^{-2} e^{-\delta^{\eta-1/2}})$$

as $\delta \to 0$.

Now let us look at the case $i = j = 0$. By Lemma 3.2 (in fact we can replace R by Λ_n in that lemma), one has

$$\sup_{n\geq 1} P\left(\sup_{x\vee y\leq\delta}\left|\frac{W_{\Lambda_n}([0,x]\times[0,y])}{(x\vee y)^\eta}\right| > \varepsilon\right) = O(e^{-\delta^{\eta-1/2}}) \tag{4.11}$$

as $\delta \to 0$.

Since (4.10) and (4.11) imply (4.9), the result follows. □

Lemma 4.5. *Under the conditions of Proposition 4.1, for each* $0 \leq \eta < 1$

$$\left\{\frac{A_n(x,y)}{(x\vee y)^\eta}, (x,y)\in[0,1]^2\right\}_{n\geq 1}$$

is relatively compact.

Proof. The proof is similar to that of Lemma 4.4. We use the same notation for Δ_{ij} and S. We only need to check that for each positive ε, there exist a δ $(0 < \delta < 1)$ and $n_0 = n_0(\delta) \in \mathbb{N}$ such that

$$\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} P\left(\sup_{\Delta_{ij}}\left|\frac{A_n(x,y)}{(x\vee y)^\eta} - \frac{A_n(i\delta,j\delta)}{\delta^\eta (i\vee j)^\eta}\right| > \varepsilon\right) \leq \varepsilon, \quad n \geq n_0. \tag{4.12}$$

We consider the case $i \vee j \geq 1$ and the case $i = j = 0$ separately. Let us first look at the case $i \vee j \geq 1$. In case of $i > j$, $i \geq 1$, note that for $(x, y) \in \Delta_{ij}$

$$
\begin{aligned}
&|A_n(x,y)/(x \vee y)^\eta - A_n(i\delta, j\delta)/((i\delta) \vee (j\delta))^\eta| \\
&= \Bigg|(x^{1-\eta} - (i\delta)^{1-\eta})(W_{\Lambda_n}(C_{\pi/2}) - Z_n(\pi/2)) - (x^{1-\eta} - (i\delta)^{1-\eta}) \int_{\arctan y/x}^{\pi/4} \frac{1}{\cos^2\theta}(W_{\Lambda_n}(C_\theta) + Z_n(\theta))d\theta \\
&\qquad + (i\delta)^{1-\eta} \int_{\arctan j/i}^{\arctan y/x} \frac{1}{\cos^2\theta}(W_{\Lambda_n}(C_\theta) + Z_n(\theta))d\theta\Bigg| \\
&\leq (i\delta)^{1-\eta}((1+1/i)^\eta - 1)(1+\pi/2)H_n + (i\delta)^{1-\eta}\left(\arctan\frac{j+1}{i} - \arctan\frac{j}{i}\right) 2\, H_n
\end{aligned}
$$

where H_n is defined in Lemma 4.3. Since $(i\delta)^{1-\eta}((1+1/i)^\eta - 1) = O(\delta^{1-\eta})$ and $(i\delta)^{1-\eta}(\arctan\frac{j+1}{i} - \arctan\frac{j}{i}) = O(\delta^{1-\eta})$ as $\delta \to 0$ and uniformly in i, j $(i > j, i \geq 1)$, then by Lemma 4.3 there exists $n_* = n_*(\delta) \in \mathbb{N}$ such that

$$
\sup_{n \geq n_*} \sup_{i > j,\, i \geq 1} P(|A_n(x,y)/(x \vee y)^\eta - A_n(i\delta, j\delta)/((i\delta) \vee (j\delta))^\eta| > \varepsilon/2) = O(e^{-\delta^{(\eta-1)/2}}) \tag{4.13}
$$

as $\delta \to 0$.

In case of $j > i, j \geq 1$ and case of $i = j \geq 1$ we can get a similar result as (4.13). Hence there exists $n_{01} = n_{01}(\delta) \in \mathbb{N}$ such that

$$
\sup_{n \geq n_{01}} \sum_{i \vee j \geq 1}^{m} P(|A_n(x,y)/(x \vee y)^\eta - A_n(i\delta, j\delta)/((i\delta) \vee (j\delta))^\eta| > \varepsilon) = O(\delta^{-2} e^{-\delta^{(\eta-1)/2}}) \tag{4.14}
$$

as $\delta \to 0$.

Now let's consider the case $i = j = 0$ and w.l.o.g. assume $y \geq x$. Then for $0 \leq x \leq y \leq \delta$

$$
\begin{aligned}
&|A_n(x,y)/(x \vee y)^\eta| \\
&= |xy^{-\eta} W_{\Lambda_n}(C_{\pi/2}) + Z_n(\pi/2)) + y^{1-\eta} \int_{\pi/4}^{\arctan y/x} \sin^{-2}\theta(W_{\Lambda_n}(C_\theta) + Z(\theta))d\theta| \\
&\leq \delta^{1-\eta}(1+\pi/2)H_n.
\end{aligned}
$$

Hence there exists $n_{02} = n_{02}(\delta) \in \mathbb{N}$ such that

$$
\sup_{n \geq n_{02}} P(\sup_{x \vee y \leq \delta} |A_n(x,y)/(x \vee y)^\eta| > \varepsilon) = O(e^{-\delta^{(\eta-1)/2}}) \tag{4.15}
$$

as $\delta \to 0$.

Now (4.14) and (4.15) imply (4.12). □

Proof of Proposition 4.1. By Lemmas 4.4 and 4.5,

$$\left\{ \frac{A_n(x,y) + B_n(x,y)}{(x \vee y)^\eta},\ (x,y) \in [0,1]^2 \right\}_{n \geq 1} \tag{4.16}$$

is relatively compact. It is easy to check that the finite-dimensional distributions of our estimated processes in (4.16) converge to those of the limiting process, which completes the proof. □

Proof of Corollary 4.1. After applying a Skorohod construction to the weak convergence statement of Proposition 4.1, the proof is similar to that of Theorem 2.3. □

Proof of Corollary 4.2. Proposition 4.1 implies the weak convergence of the distribution function of the left hand side of (4.6) to the distribution function of the right hand side of (4.6). This property carries over to the inverse functions Q_{Λ_n} and Q_Λ. □

Proof of Corollary 4.3. From another Skorohod construction we obtain an a.s. version of the statement of Theorem 2.2; without changing the notation we now work with this construction. Since for $0 < x, y \leq 1$

$$\Lambda([0,x] \times [0,y]) = x + y - l(x,y),$$
$$\Lambda_n([0,x] \times [0,y]) = \lceil kx \rceil / k + \lceil ky \rceil / k - \hat{l}_2(x,y) - \delta_n(x,y)/k$$

($\delta_n(x,y)$ takes values in $\{0,1,2\}$), it follows that for each $\varepsilon > 0$

$$\sup_{0<x,y\leq 1} k^{1/2-\varepsilon} \Big| \Lambda_n([0,x] \times [0,y]) - \Lambda([0,x] \times [0,y]) \Big| \to 0 \quad \text{a.s.} \tag{4.17}$$

as $n \to \infty$.

We now show that (4.2), (4.3), (4.4), (4.5) hold a.s. We already saw, below (4.7), that (4.2) holds a.s. and the a.s. version of (4.3) follows immediately from (4.17).

By (4.17) and (4.2), it is easily follows that

$$\sup_{E \in \mathcal{E}} k^{1/2-\varepsilon} \Big| \Lambda_n(E) - \Lambda(E) \Big| \to 0 \quad \text{a.s.} \tag{4.18}$$

as $n \to \infty$, where $\mathcal{E} := \{E |\ E = [x_1, x_2] \times [y_1, y_2], 0 < x_1 \leq x_2 \leq 2,\ 0 < y_1 \leq y_2 \leq 2\}$. Let

$E_n(x) = [x - k^{-1/6}, x + k^{-1/6}] \times [1 - k^{-1/6}, 1 + k^{-1/6}]$. Then

$$\begin{aligned}
&\sup_{0<x\leq 1} |\lambda_n(x,1) - \lambda(x,1)| \\
&= \sup_{0<x\leq 1} |\frac{1}{4}k^{1/3}\Lambda_n(E_n(x)) - \frac{1}{4}k^{1/3}\Lambda(E_n(x)) + \frac{1}{4}k^{1/3}\Lambda(E_n(x)) - \lambda(x,1)| \\
&\leq \sup_{0<x\leq 1} \frac{1}{4}k^{1/3}|\Lambda_n(E_n(x)) - \Lambda(E_n(x))| + \sup_{0<x\leq 1} |\frac{1}{4}k^{1/3}\Lambda(E_n(x)) - \lambda(x,1)| \\
&\to 0 \quad \text{a.s. as } n \to \infty,
\end{aligned}$$

as $n \to \infty$, by (4.18) and $\lambda(0,1) = 0$. The proofs of $\sup_{0<y\leq 1} |\lambda_n(1,y) - \lambda(1,y)| \to 0$ a.s. and $\sup_{0<x,y\leq 1} |R_{jn}(x,y) - R_j(x,y)| \to 0$, $j = 1, 2$, a.s. are similar. Hence (4.4) and (4.5) hold a.s.

According to Corollary 4.2 we have

$$Q_{\Lambda_n}(1-\alpha) \to Q_\Lambda(1-\alpha) \quad \text{a.s.}$$

as $n \to \infty$, hence also in probability. □

5 Simulation study and real data application

In this section we present a small simulation study, making use of the results of section 4. We will consider one distribution satisfying the domain of attraction condition and one that fails to satisfy it. At the end of the section, we will apply our procedure to financial data. *Throughout* we take $\beta = 2$ in the test statistic of (1.12).

Consider the bivariate Cauchy distribution restricted to the first quadrant, with density

$$f(x,y) = \frac{2}{\pi(1 + x^2 + y^2)^{\frac{3}{2}}}, \quad x, y > 0.$$

It readily follows that

$$\Lambda([0,x] \times [0,y]) = x + y - \sqrt{x^2 + y^2}, \quad \lambda(x,y) = \frac{xy}{(x^2+y^2)^{3/2}}, \quad x, y > 0.$$

This distribution satisfies the conditions of Theorem 2.3; in particular (2.5) holds with $\alpha = 2$ (see Einmahl *et al.* (2001), pp. 1409-1410). First we present in Table 1 the quantiles of the limiting random variable

$$\iint_{0<x,y\leq 1} \frac{(A(x,y) + B(x,y))^2}{(x \vee y)^2} dxdy,$$

using the approximation of section 4. We used 100,000 replications. With high probability these quantiles are accurate up to 0.01.

p	0.25	0.50	0.75	0.90	0.95
$Q(p)$	0.10	0.14	0.22	0.34	0.44

Table 1: Quantiles of the limiting r.v. for $\beta = 2$ for the Cauchy distribution.

Now for sample size $n = 2000$, we simulate 1000 times the test statistic

$$k \iint_{0<x,y\leq 1} \frac{(\hat{l}_1(x,y) - \hat{l}_2(x,y))^2}{(x \vee y)^2} dxdy,$$

for various values of k. Using the 0.95-th quantile above, we find the simulated type-I error probabilities; see Table 2. In the ideal situation the number of rejections is a binomial r.v.

k	20	40	60	80	100	125	150	175	200	250	300	400
$\hat{\alpha}$	.049	.048	.055	.039	.038	.049	.046	.055	.049	.060	.055	.082

Table 2: Simulated type-I error for the Cauchy distribution: $n = 2000$ and $\alpha = 0.05$.

with parameters 1000 and 0.05. So the numbers in the table are remarkebly close to 0.05. Only for $k = 400$, the bias seems to set in. In addition, in Figure 1 we see, for various k, on the left for one sample of size $n = 2000$ the values of the test statistic and on the right the median and 0.95-th quantile for the test statistic based on 800 samples. Note that the behavior of the test statistic fluctuates with k, but that for all k in the figure the value is far below 0.44, the 0.95-th quantile of the limiting random variable.

Next we consider a distribution with uniform-$(0,1)$ marginals (a copula), which does not satisfy the bivariate domain of attraction condition. Since both marginals are uniform, they are in the univariate domain of attraction of the reverse Weibull law. So it is the dependence structure that causes the failure. The distribution is an adaptation of a distribution in Schlather (2001): take a density of 3/2 on the following rectangles: $[2^{-(2m+1)}, 2^{-(2m)}] \times [2^{-(2r+1)}, 2^{-(2r)}]$, for $m = 0, 1, 2, \ldots$ and $r = 0, 1, 2, \ldots$; in this way a probability mass of 2/3 is assigned. The remaining 1/3 is assigned by taking the uniform distribution on the line segments from $(2^{-(2m+2)}, 2^{-(2m+2)})$ to $(2^{-(2m+1)}, 2^{-(2m+1)})$, $m = 0, 1, 2, \ldots$, such that

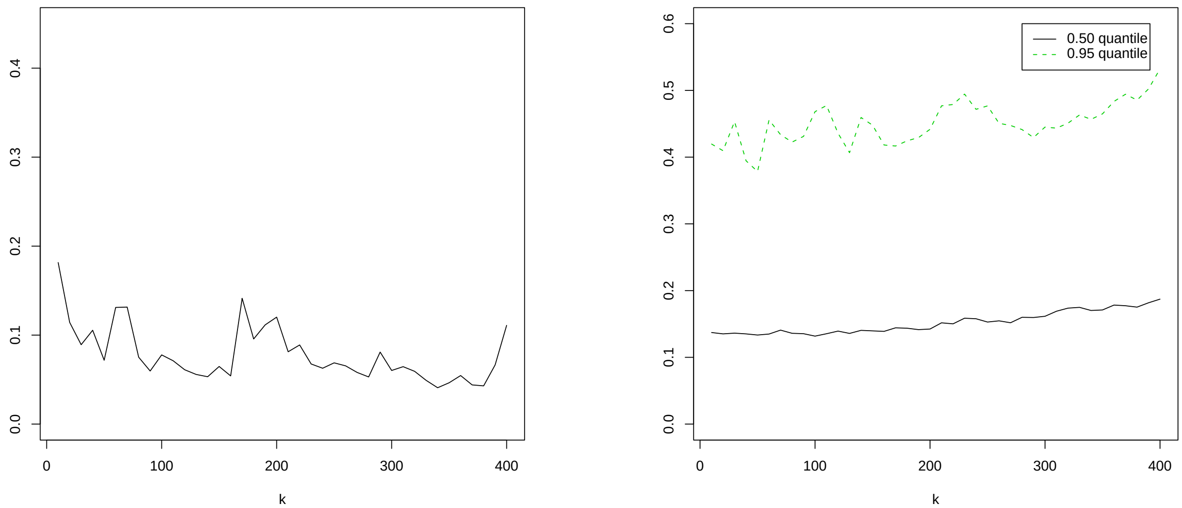

Figure 1: Cauchy distribution: test statistic (left) and quantiles of the test statistic (right).

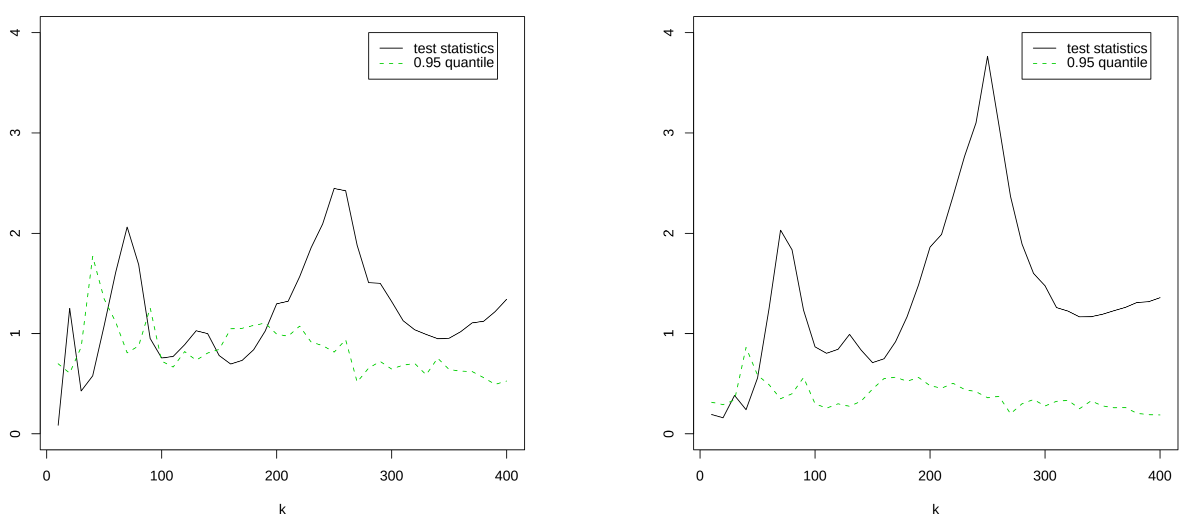

Figure 2: Alternative distribution: test statistics and 0.95-th quantiles of 2 samples.

the mass of the m-th segment is equal to $2^{-(2m+2)}$. In Figure 2, we see for varying k the test statistics and simulated 0.95-th quantiles of two samples of size $n = 2000$ from this distribution. Again the test statistics fluctuate with k, but from a certain k on (and for most values of k), the null hypothesis is clearly rejected.

Finally, we apply the test to real data, similarly as we just did for the simulated data sets in Figure 2. The data are 3283 daily logarithmic equity returns over the period 1991-2003 for two Dutch banks, ING and ABN AMRO bank. The bivariate, heavy-tailed data are shown in Figure 3 on the left; on the right we see again the test statistic and 0.95-th quantile. Since the test statistic is everywhere clearly below the quantile, we cannot reject

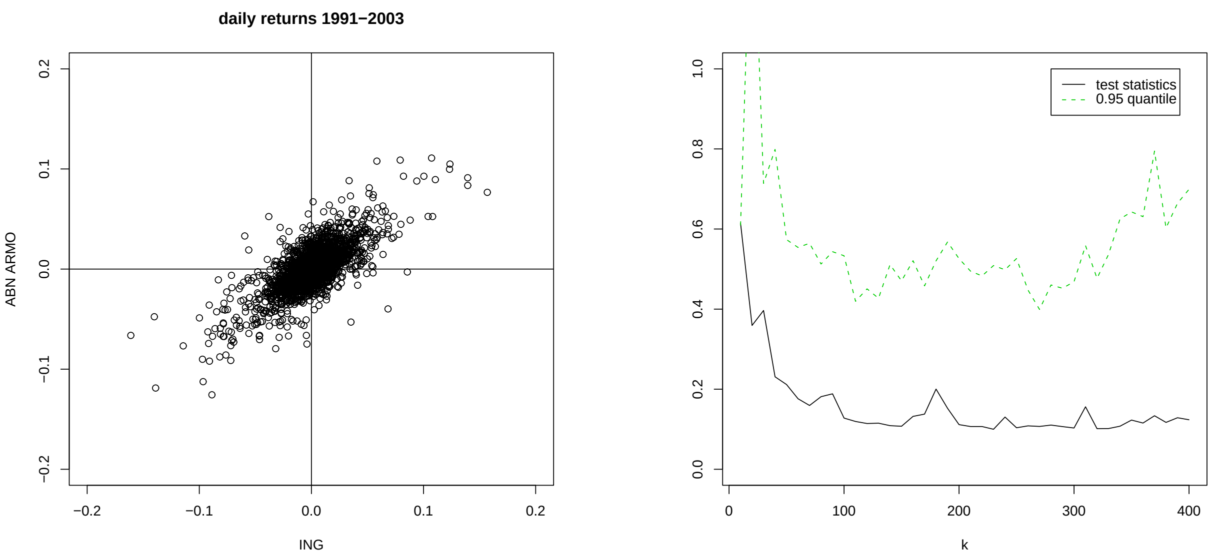

Figure 3: Daily equity returns of two Dutch banks (left) and test statistics and 0.95-th quantiles (right).

the null hypothesis. This is a satisfactory result, because it allows us to analyze these data further, using statistical theory of extremes.

References

[1] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[2] Dietrich, D., de Haan L. and Hüsler, J. (2002). Testing extreme value conditions. *Extremes*, **5**, 71-85.

[3] Drees, H., de Haan L. and Li, D. (2004). Approximations to the tail empirical distribution function with application to testing extreme value conditions. Submitted.

[4] Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *Journal of Multivariate Analysis*, **64**, 25-47.

[5] Einmahl, J. (1992). Limit theorems for tail processes with application to intermediate quantile estimation. *J. Statist. Plann. Inference*, **32**, 137-145.

[6] Einmahl, J., de Haan, L. and Piterbarg, V. (2001). Nonparametric estimation of the spectral measure of an extreme value distribution. *Ann. Statist.*, **29**, 1410-1423.

[7] de Haan, L. and Sinha, A.K. (1999). Estimating the probability of a rare event. *Ann. Statist.*, **27**, 732-759.

[8] Huang, X. (1992). *Statistics of bivariate extremes.* Thesis, Erasmus University Rotterdam. Tinbergen Institute Series no. 22.

[9] Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.*, **42**, 1285-1295.

[10] Orey, S. and Pruitt, W. (1973). Sample functions of the N-parameter Wiener process. *Ann. Probab.*, **1**, 138-163.

[11] Paulauskas, V. and Račkauskas, A. (1989). *Approximation Theory in the Central Limit Theorem. Exact Results in Banach Spaces.* Kluwer, Dordrecht.

[12] Schlather, M. (2001). Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution. *Statist. Probab. Lett.*, **53**, 325-329.

[13] Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

[14] van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer, New York.

JHJE
Dept. of Econometrics & OR
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
Email: j.h.j.einmahl@uvt.nl

LdH & DL
Econometric Institute
Erasmus University
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
Email: ldehaan@few.eur.nl & dli@few.eur.nl