New estimators of the Pickands dependence function and a test for extreme-value dependence

Axel Bücher, Holger Dette, Stanislav Volgushev
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: axel.buecher@ruhr-uni-bochum.de
e-mail: holger.dette@ruhr-uni-bochum.de
e-mail: stanislav.volgushev@ruhr-uni-bochum.de

May 27, 2010

Abstract

We propose a new class of estimators for Pickands dependence function which is based on the best L^2-approximation of the logarithm of the copula by logarithms of extreme-value copulas. An explicit integral representation of the best approximation is derived and it is shown that this approximation satisfies the boundary conditions of a Pickands dependence function. The estimators $\hat{A}(t)$ are obtained by replacing the unknown copula by its empirical counterpart and weak convergence of the process $\sqrt{n}\{\hat{A}(t) - A(t)\}_{t\in[0,1]}$ is shown. A comparison with the commonly used estimators is performed from a theoretical point of view and by means of a simulation study. Our asymptotic and numerical results indicate that some of the new estimators outperform the rank-based versions of Pickands estimator and an estimator which was recently proposed by Genest and Segers (2009). As a by-product of our results we obtain a simple test for the hypothesis of an extreme-value copula, which is consistent against all alternatives with continuous partial derivatives of first order satisfying $C(u, v) \geq uv$.

Keywords and Phrases: Extreme-value copula, minimum distance estimation, Pickands dependence function, weak convergence, copula process, test for extreme-value dependence
AMS Subject Classification: Primary 62G05, 62G32; secondary 62G20

1 Introduction

The copula provides an elegant margin-free description of the dependence structure of a random variable. By the famous theorem of Sklar (1959) it follows that the distribution function H of a bivariate random variable (X, Y) can be represented in terms of the marginal distributions F and G of X and Y, that is

$$H(x, y) = C(F(x), G(y)),$$

where C denotes the copula, which characterizes the dependence between X and Y. Extreme-value copulas arise naturally as the possible limits of copulas of component-wise maxima of independent, identically distributed or strongly mixing stationary sequences [see Deheuvels (1984) and Hsing (1989)]. These copulas provide flexible tools for modelling joint extremes in risk management. An important application of extreme-value copulas appears in the modelling of data with positive dependence, and in contrast to the more popular class of Archimedean copulas they are not symmetric [see Tawn (1988) or Ghoudi et al. (1998)]. Further applications can be found in Coles et al. (1999) or Cebrian et al. (2003) among others. A copula C is an extreme-value copula if and only if it has a representation of the form

$$C(u, v) = \exp\Big(\log(uv) A\Big(\frac{\log v}{\log uv}\Big)\Big), \tag{1.1}$$

where $A : [0, 1] \to [1/2, 1]$ is a convex function satisfying $\max\{t, 1-t\} \leq A(t) \leq 1$, which is called Pickands dependence function [see Pickands (1981)]. The representation of (1.1) of the extreme-value copula C depends only on the one-dimensional function A and statistical inference on a bivariate extreme-value copula C may now be reduced to inference on its Pickands dependence function A.

The problem of estimating Pickands dependence function nonparametrically has found considerable attention in the literature. Roughly speaking, there exist two classes of estimators. The classical nonparametric estimator is that of Pickands (1981) [see Deheuvels (1991) for its asymptotic properties] and several variants have been discussed. Alternative estimators have been proposed and investigated in the papers by Capéraà et al. (1997), Jiménez et al. (2001), Hall and Tajvidi (2000), Segers (2007) and Zhang et al. (2008), where the last-named authors also discussed the multivariate case. In most references the estimators of Pickands dependence function are constructed assuming knowledge of the marginal distributions. Recently Genest and Segers (2009) proposed rank-based versions of the estimators of Pickands (1981) and Capéraà et al. (1997), which do not require knowledge of the marginal distributions. In general all of these estimators are neither convex nor do they satisfy the boundary restriction $\max\{t, 1-t\} \leq A(t) \leq 1$, in particular the condition $A(0) = A(1) = 1$. Consequently, the estimators are modified without changing their asymptotic properties, such that they satisfy the endpoint constraints (or the boundary condition), see e.g. Fils-Villetard et al. (2008).

Before a specific model of an extreme-value copula is selected it is necessary to check this assumption by a statistical test, that is a test for the hypothesis

$$H_0 : C \in \mathcal{C} \tag{1.2}$$

where $\mathcal{C}$ denotes the class of all copulas satisfying (1.1). Throughout this paper we call (1.2) the hypothesis of extreme-value dependence. The problem of testing this hypothesis has found much less attention in the literature. To our best knowledge, the only currently available test of extremeness was proposed by Ghoudi et al. (1998). It exploits the fact that for an extreme-value copula the random variable $W = H(X, Y)$ satisfies the identity

$$-1 + 8E[W] - 9E[W^2] = 0. \tag{1.3}$$

The properties of this test have recently been studied by Ben Ghorbal et al. (2009), who determined the finite- and large-sample variance of the test statistic. In particular, the test proposed by Ghoudi et al. (1998) is not consistent against alternatives satisfying (1.3).
The present paper has two purposes. The first is the development of some alternative estimators of Pickands dependence function, which are based on the concept of minimum distance estimation. More precisely, we propose to consider the best approximation of the logarithm of the empirical copula by functions of the form

$$\log(uv) A\left(\frac{\log v}{\log uv}\right) \tag{1.4}$$

with respect to a weighted L^2-distance. It turns out that the minimal distance and the corresponding optimal function can be determined explicitly. On the basis of this result, we derive new estimators which have a similar structure as the integral representations obtained in Lemma 3.1 of Genest and Segers (2009). By choosing various weight functions in the L^2-distance we obtain an infinite dimensional class of estimators.
The new estimators can be alternatively motivated observing that the identity (1.1) yields the representation $A(t) = \log C(y^{1-t}, y^t)/\log y$ for any $y \in (0,1)$. This leads to a simple class of estimators, i.e.

$$\hat{A}_{n,\delta_y}(t) = \frac{\log \tilde{C}_n(y^{1-t}, y^t)}{\log y} \ ; \quad y \in (0,1),$$

where δ_y is the Dirac measure at the point y and $\tilde{C}_n$ is the slightly modified version of the empirical copula, such that the logarithm is well defined (see Section 3 for details). By averaging these estimators with respect to a distribution, say π, we obtain an infinite dimensional class of estimators of the form

$$\hat{A}_{n,\pi}(t) = \int_0^1 \frac{\log \tilde{C}_n(y^{1-t}, y^t)}{\log y} \pi(dy),$$

which turn out to coincide with the estimators obtained by the concept of best L^2-approximation described in the previous paragraph.
The second purpose of the paper is to present a new test for the hypothesis of extreme-value dependence, which is consistent against all alternatives with continuous partial derivatives of

first order satisfying $C \geq \Pi$, where Π denotes the independence copula. Our approach is based on an estimator of a minimum L^2-distance between the true copula and the class of extreme-value copulas. To our best knowledge, this method provides the first test in this context which is consistent against such a general class of alternatives.

The remaining part of the paper is organized as follows. In Section 2 we consider the approximation problem from a theoretical point of view. In particular, we derive explicit representations for the minimal L^2-distance between the logarithm of the copula and its best approximation by a function of the form (1.4), which will be the basis for all statistical applications. The new estimators are defined in Section 3, where we also investigate their asymptotic properties. In particular, we prove weak convergence of the process $\{\sqrt{n}(\hat{A}_{n,\pi}(t) - A(t))\}_{t\in[0,1]}$ in the space of uniformly bounded functions on the interval $[0,1]$ under appropriate assumptions on the distribution π. Furthermore, we accomplish a theoretical and empirical comparison of the new estimators with the rank-based estimators proposed in Genest and Segers (2009). In particular, we demonstrate that some of the "new" estimators have a substantially smaller mean squared error than the estimators proposed by the last-named authors. In Section 4 we introduce and investigate the new test of extreme-value dependence. In particular, we derive the asymptotic distribution of the test statistic under the null hypothesis as well as under the alternative and study its finite sample properties by means of a simulation study. Finally some technical details are deferred to an Appendix.

2 A measure of extreme-value dependence

Let $\mathcal{A}$ denote the set of all functions $A : [0,1] \to [1/2, 1]$, and define Π as the copula corresponding to independent random variables, i.e. $\Pi(u,v) = uv$. Throughout this paper we assume that the copula C satisfies $C \geq \Pi$ which holds for any extreme-value copula due to the lower bound for the function A. As pointed out by Scaillet (2005), this property is equivalent to the concept of positive quadrant dependence, that is

$$P(X \leq x, Y \leq y) \geq P(X \leq x)P(Y \leq y) \;\; \forall \; (x,y) \in \mathbb{R}^2.$$

For a copula with this property we define the L^2-distance

$$M_C(A) = \int_{(0,1)^2} \left(\log C(u,v) - \log(uv) A\left(\frac{\log v}{\log uv}\right)\right)^2 h(-\log uv)\, d(u,v), \tag{2.1}$$

where $h : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous weight function with the following properties

$$\text{for all } K > 0 : \sup_{x\in[0,K]} |x^2 h(x)| < \infty \tag{2.2}$$

$$h^*(y) \in L^1(0,1) \text{ or equivalently } x^3 e^{-x} h(x) \in L^1(0,\infty) \tag{2.3}$$

$$\int_0^1 h^*(y)(-\log y)^{-1} y^{-\lambda}\, dy = \int_0^1 (\log y)^2 h(-\log y) y^{-\lambda}\, dy < \infty \text{ for some } \lambda > 1 \tag{2.4}$$

and the function h^* is defined by $h^*(y) = -\log^3(y)h(-\log y)$. These conditions are needed in Sections 3 and 4 to establish weak convergence of an appropriate estimator of $\arg\min_A M_C(A)$ and of an estimator for the minimal distance

$$\min\Big\{M_C(A) \mid A \in \mathcal{A}\Big\},$$

respectively. The following result is essential for our approach and provides an explicit expression for the best L^2-approximation of the logarithm of the copula by the logarithm of an extreme-value copula of the form (1.1) and as a by-product characterizes the function A^* minimizing $M_C(A)$.

Theorem 2.1. *Assume that the given copula satisfies $C \geq \Pi$ and that the weight function h satisfies conditions (2.2) and (2.3). Then the function*

$$A^* = \arg\min\{M_C(A) \mid A \in \mathcal{A}\}$$

is uniquely determined and given by

$$A^*(t) = B_h^{-1} \int_0^1 \frac{\log C(y^{1-t}, y^t)}{\log y} h^*(y)\, dy \ , \tag{2.5}$$

where h^* is defined by $h^*(y) = -(\log y)^3 h(-\log y)$ and

$$B_h = \int_0^\infty x^3 e^{-x} h(x)\, dx = -\int_0^1 (\log y)^3 h(-\log y)\, dy = \int_0^1 h^*(y)\, dy. \tag{2.6}$$

Moreover, the minimal L^2-distance between the logarithms of the given copula and the class of extreme-value copulas is given by

$$M_C(A^*) = \int_{(0,1)^2} \left(\frac{\log C(y^{1-t}, y^t)}{\log y}\right)^2 h^*(y)\, d(y,t) - B_h \int_0^1 (A^*(t))^2\, dt. \tag{2.7}$$

Proof. Substituting

$$(s,t) = g(u,v) = (-\log uv, \frac{\log v}{\log uv})$$

in the integral (2.1) with inverse $g^{-1}(s,t) = (\exp(-s(1-t)), \exp(-st))$ and Jacobian determinant $\det Dg^{-1}(s,t) = se^{-s}$ yields for the functional M_C the representation

$$M_C(A) = \int_0^1 \int_0^\infty (\bar{C}(s,t) - sA(t))^2 se^{-s} h(s)\, ds\, dt,$$

where $\bar{C}(s,t) = -\log C(g^{-1}(s,t))$. With the notation

$$A^*(t) = \frac{\int_0^\infty \bar{C}(x,t) x^2 e^{-x} h(x)\, dx}{\int_0^\infty x^3 e^{-x} h(x)\, dx}$$

it follows by a straightforward calculation and Fubini's theorem that

$$\begin{aligned} M_C(A) &= \int_0^1 \int_0^\infty (\bar{C}(s,t) - sA^*(t))^2 s e^{-s} h(s)\, ds\, dt + \int_0^\infty s^3 e^{-s} h(s)\, ds \int_0^1 (A^*(t) - A(t))^2\, dt \\ &\quad + 2\int_0^1 \int_0^\infty (sA^*(t) - \bar{C}(s,t))(sA(t) - sA^*(t)) s e^{-s} h(s)\, ds\, dt \\ &= \int_0^1 \int_0^\infty (\bar{C}(s,t) - sA^*(t))^2 s e^{-s} h(s)\, ds\, dt + \int_0^\infty s^3 e^{-s} h(s)\, ds \int_0^1 (A^*(t) - A(t))^2\, dt, \end{aligned}$$

and we can conclude that $A^*(t)$ is the best approximation of functions A to the copula C with respect to the distance $M_C(A)$. Resubstituting $x = -\log y$ we obtain the alternative expression

$$A^*(t) = -B_h^{-1} \int_0^1 \log C(y^{1-t}, y^t)(\log y)^2\, h(-\log y)\, dy \tag{2.8}$$

where the constant B_h is defined in (2.6). Observing the definition of h^* this completes the proof of Theorem 2.1. □

Note that $A^*(t) = A(t)$ if and only if C is an extreme-value copula with Pickands dependence function A. Furthermore, the following Lemma shows that the minimizing function A^* defined in (2.5) satisfies the boundary conditions of Pickands dependence functions.

Lemma 2.2. *Assume that C is a copula satisfying $C \geq \Pi$. Then the function A^* defined in (2.5) has the following properties*

(i) $A^*(0) = A^*(1) = 1$

(ii) $A^*(t) \geq t \vee (1-t)$

(iii) $A^*(t) \leq 1$.

Proof. Assertion (i) is obvious. For a proof of (ii) one uses the Fréchet-Hoeffding bound $C(u,v) \leq u \wedge v$ and obtains the assertion by a direct calculation. Similarly assertion (iii) follows from the inequality $C \geq \Pi$. □

Example 2.3. In the following we discuss two interesting examples for the weight function h, which will be used for the construction of the new estimators of Pickands dependence function.

(a) For the choice $h_\alpha^{(1)}(x) = x^{-\alpha}$ $\alpha \in [0,2]$ we obtain $B_{h_\alpha}^{(1)} = \Gamma(4-\alpha)$ and

$$A^*(t) = -\Gamma(4-\alpha)^{-1} \int_0^1 \log C(y^{1-t}, y^t)(-\log)^{2-\alpha}(y)\, dy. \tag{2.9}$$

In particular we have for $\alpha = 2$

$$A^*(t) = -\int_0^1 \log\; C(y^{1-t}, y^t) dy \tag{2.10}$$

for the best approximation.

(b) For the choice $h_k^{(2)}(x) = x^{-2}e^{-kx}$ with $k \geq 0$ it follows

$$B_{h_k}^{(2)} = \int_0^\infty xe^{-(k+1)x}\,dx = (k+1)^{-2}$$

and

$$A^*(t) = -(k+1)^2 \int_0^1 \log C(y^{1-t}, y^t)\, y^k\, dy. \tag{2.11}$$

Again we obtain for $k = 0$ the representation (2.10). Observing the identity ($A \geq 0$)

$$-\int_0^1 \log(y^A)dy = A = \Big(\int_0^1 y^{A-1}dy\Big)^{-1}$$

it follows for an extreme-value copula from the identities (1.1) and (2.10) that

$$A^*(t) = \Big(\int_0^1 \frac{C(y^{1-t}, y^t)}{y}\,dy\Big)^{-1}.$$

This is the integral formula, which motivates the representation of Pickands estimator in Lemma 3.1 of Genest and Segers (2009).

Example 2.4. In the following we exemplarily calculate the minimal distance $M_C(A^*)$ and its corresponding best approximation A^* for two copula families. The weight function is chosen as in Example 2.3 (b), that is $h_1^{(2)}(x) = x^{-2}e^{-x}$. First we investigate the Gaussian copula defined by

$$C_\rho(u, v) = \Phi_2(\Phi(u), \phi(v), \rho),$$

where Φ is the standard normal distribution function and $\Phi_2(\cdot, \cdot, \rho)$ is the distribution function of two bivariate standard normal distributed random variables with correlation $\rho \in [0, 1]$. For the limiting cases $\rho = 0$ and $\rho = 1$ we obtain the independence and perfect dependence copula, respectively, while for $\rho \in (0, 1)$ C_ρ is not an extreme-value copula. The minimal distances are plotted as a function of ρ in the left part of the first line of Figure 1. In the right part we show some functions A^* corresponding to the best approximation of the Gaussian copula by an extreme-value copula.

In the second example we consider a convex combination of a Gumbel copula with parameter $\log 2/\log 1.5$ (corresponding to a coefficient of tail dependence of 0.5) and a Clayton copula with parameter 2, i.e.

$$C_\alpha(u, v) = \alpha\, C_{\text{Clayton}}(u, v) + (1-\alpha)\, C_{\text{Gumbel}}(u, v), \quad \alpha \in [0, 1].$$

Note that only the Gumbel copula is an extreme-value copula and obtained for $\alpha = 0$. The minimal distances are depicted in the left part of the lower panel of Figure 1 as a function of α. In the right part we show the functions A^* corresponding to the best approximation of C_α by an extreme-value copula.

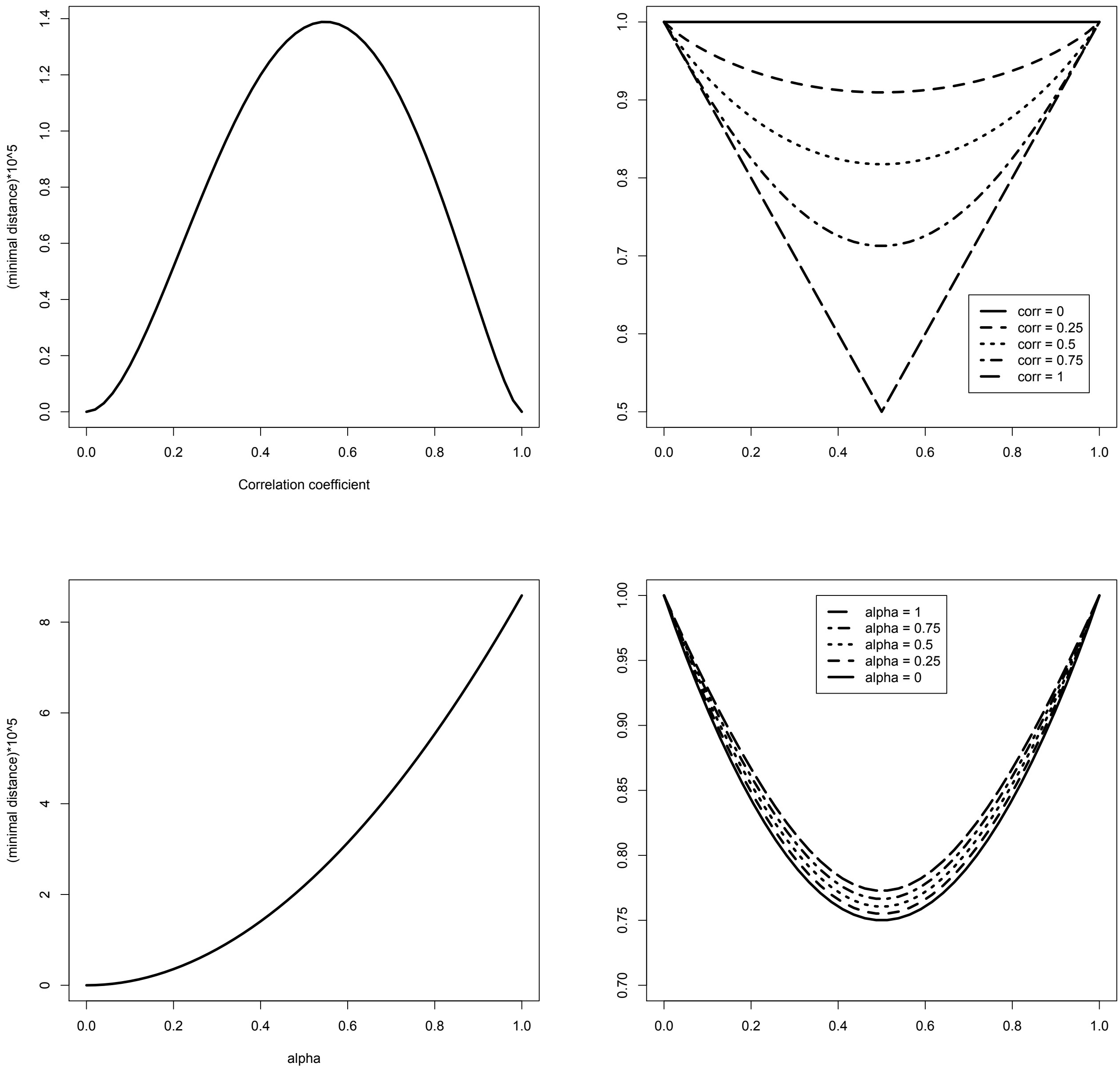

Figure 1: *Left: Minimal distances* $M_C(A^*) \times 10^5$ *for the Gaussian copula (as a function of its correlation coefficient) and for the convex combination of a Gumbel and a Clayton copula (as a function of the parameter* α *in the convex combination). Right: Pickands dependence functions corresponding to the best approximations by extreme-value copulas.*

3 A class of minimum distance estimators

3.1 New estimators and weak convergence

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ denote independent identically distributed bivariate random variables with copula C and marginals F and G. Observing Theorem 2.1 it is reasonable to define a class of estimators for Pickands dependence function by replacing the unknown copula in (2.5) through the empirical copula

$$\hat{C}_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} I\{\hat{U}_i \le u, \quad \hat{V}_i \le v\}, \tag{3.1}$$

where

$$\hat{U}_i = \frac{1}{n+1} \sum_{j=1}^{n} I\{X_j \le X_i\} \quad \text{and} \quad \hat{V}_i = \frac{1}{n+1} \sum_{j=1}^{n} I\{Y_j \le Y_i\}$$

denote the (slightly modified) empirical distribution functions of the samples $\{X_j\}_{j=1}^n$ and $\{Y_j\}_{j=1}^n$ at the points X_i and Y_i, respectively. The asymptotic properties of the corresponding estimators will be investigated in this section. For technical reasons we require that the argument in the logarithm in the representation (2.5) is positive and propose to use the estimator $\tilde{C}_n = \hat{C}_n \vee n^{-\gamma}$ where the constant γ satisfies $\gamma > 1/2$ and the empirical copula $\hat{C}_n$ is defined in (3.1). Assuming that the copula C has continuous partial derivatives of first order, it follows that the process $\sqrt{n}(\tilde{C}_n - C)$ shows the same limiting behavior as the empirical copula process $\sqrt{n}(\hat{C}_n - C)$, i.e.

$$\sqrt{n}(\tilde{C}_n - C) \overset{w}{\rightsquigarrow} \mathbb{G}_C, \tag{3.2}$$

where the symbol $\overset{w}{\rightsquigarrow}$ denotes weak convergence in $l^\infty[0,1]^2$. Here $\mathbb{G}_C$ is a Gaussian field on the square $[0,1]^2$ which admits the representation

$$\mathbb{G}_C(\mathbf{x}) = \mathbb{B}_C(\mathbf{x}) - \partial_1 C(\mathbf{x}) \mathbb{B}_C(x_1, 1) - \partial_2 C(\mathbf{x}) \mathbb{B}_C(1, x_2),$$

where $\mathbf{x} = (x_1, x_2), \mathbb{B}_C$ is a bivariate pinned C-Brownian sheet on the square $[0,1]^2$ with covariance kernel given by

$$\operatorname{Cov}(\mathbb{B}_C(\mathbf{x}), \mathbb{B}_C(\mathbf{y})) = C(\mathbf{x} \wedge \mathbf{y}) - C(\mathbf{x})C(\mathbf{y})$$

and the minimum $\mathbf{x} \wedge \mathbf{y}$ is understood component-wise [see Rüschendorf (1976), Fermanian et al. (2004) or Tsukahara (2005)]. Observing the representation (2.5) we obtain the estimator

$$\hat{A}_{n,h}(t) = B_h^{-1} \int_0^1 \frac{\log \tilde{C}_n(y^{1-t}, y^t)}{\log y} h^*(y) dy \tag{3.3}$$

for Pickands dependence function. Note that this relation specifies an infinite dimensional class of estimators indexed by the set of all functions h satisfying the conditions (2.2) - (2.4). The following results specify the asymptotic properties of these estimators.

Theorem 3.1. *If the copula $C \geq \Pi$ has continuous partial derivatives of first order and the weight function h satisfies conditions* (2.2) - (2.4) *for some $\lambda > 1$, then we have for any $\gamma \in (1/2, \lambda/2)$ as $n \to \infty$*

$$\mathbb{A}_{n,h} = \sqrt{n}(\hat{A}_{n,h} - A^*) \overset{w}{\rightsquigarrow} \mathbb{A}_{C,h} = B_h^{-1} \int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} \frac{h^*(y)}{\log y} dy \tag{3.4}$$

in $l^\infty[0,1]$.

Note that Theorem 3.1 is also correct if the given copula is not an extreme-value copula. In other words: it establishes the weak convergence of the process $\sqrt{n}(\hat{A}_{n,h} - A^*)$ to a centered Gaussian process, where A^* denotes the function corresponding to the best approximation of the copula C by an extreme-value copula. A^* coincides with Pickands dependence function if C is an extreme-value copula. Note also that Theorem 3.1 excludes the case $h(x) = x^{-2}$, which corresponds to the function $h^*(y) = -\log y$, because condition (2.4) is not satisfied for this weight function. Nevertheless, under the additional assumption that Pickands dependence function A is twice continuously differentiable, the assertion of the preceding theorem is still valid.

Theorem 3.2. *Assume that C is an extreme-value copula with twice continuously differentiable Pickands dependence function A. For the weight function $h(x) = x^{-2}$ we have for any $\gamma \in (1/2, 3/4)$ as $n \to \infty$*

$$\mathbb{A}_{n,1/x^2} = \sqrt{n}(\hat{A}_{n,1/x^2} - A) \overset{w}{\rightsquigarrow} \mathbb{A}_{C,1/x^2} = -\int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} dy$$

in $l^\infty[0,1]$, where $\hat{A}_{n,1/x^2}(t) = -\int_0^1 \log \tilde{C}_n(y^{1-t}, y^t)\, dy$.

Remark 3.3.
(a) If the marginals of (X, Y) are independent the distribution of the random variable $\mathbb{A}_{C,1/x^2}$ coincides with the distribution of the random variable $\mathbb{A}_r^P$, which appears as the weak limit of the appropriately standardized Pickands estimator, see Genest and Segers (2009).
(b) A careful inspection of the proof of Theorem 3.1 reveals that the condition $C \geq \Pi$ can be relaxed to $C \geq \Pi^\kappa$ for some $\kappa > 1$, if one imposes stronger conditions on the weight function.
(c) We note that the estimator depends on the parameter γ which is used for the construction of the statistic $\tilde{C}_n = \hat{C}_n \vee n^{-\gamma}$. This modification is only made for technical purposes and from a practical point of view the behavior of the estimators does not change substantially provided that γ is chosen larger than 2/3.

3.2 Examples

In this subsection we illustrate the results investigating the two examples discussed at the end of Section 2. With $h_\alpha^{(1)}(x) = x^{-\alpha}$ ($\alpha \in [0,2]$) as defined in Example 2.3(a) we obtain from (2.9)

$$\hat{A}_{n,h_\alpha^{(1)}}(t) = -\Gamma(4-\alpha)^{-1} \int_0^1 \log \tilde{C}_n(y^{1-t}, y^t)(-\log)^{2-\alpha}(y)\, dy, \tag{3.5}$$

as an estimator of Pickands dependence function. Secondly, considering the weight function $h_k^{(2)}(x) = x^{-2}e^{-kx}$ with $k \geq 0$ corresponding to Example 2.3(b), we obtain

$$\hat{A}_{n,h_k^{(2)}}(t) = -(k+1)^2 \int_0^1 \log \tilde{C}_n(y^{1-t}, y^t)\, y^k\, dy. \tag{3.6}$$

Note that for $\alpha = 2$ and $k = 0$ we have

$$\hat{A}_{n,1/x^2}(t) = -\int_0^1 \log \tilde{C}_n(y^{1-t}, y^t)\, dy.$$

The processes $\{\mathbb{A}_{n,h_\alpha^{(1)}}(t)\}_{t\in[0,1]}$ and $\{\mathbb{A}_{n,h_k^{(2)}}(t)\}_{t\in[0,1]}$ converge weakly in $l^\infty[0,1]$ to the process $\{\mathbb{A}_{C,h_\alpha^{(1)}}\}_{t\in[0,1]}$ and $\{\mathbb{A}_{C,h_k^{(2)}}\}_{t\in[0,1]}$, which are given by

$$\begin{aligned} \mathbb{A}_{C,h_\alpha^{(1)}}(t) &= -\Gamma(4-\alpha)^{-1} \int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} (-\log y)^{2-\alpha}\, dy, && (3.7)\\ \mathbb{A}_{C,h_k^{(2)}}(t) &= -(k+1)^2 \int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)}\, y^k\, dy, && (3.8) \end{aligned}$$

respectively. Consequently, for $C \in \mathcal{C}$, the asymptotic variances of $\hat{A}_{n,h_\alpha^{(1)}}$ and $\hat{A}_{n,h_k^{(2)}}$ are obtained as

$$\begin{aligned} \mathrm{Var}(\mathbb{A}_{C,h_\alpha^{(1)}}(t)) &= \Gamma(4-\alpha)^{-2} \int_0^1 \int_0^1 \sigma(u,v;t)(uv)^{-A(t)}(-\log u)^{2-\alpha}(-\log v)^{2-\alpha}\, du\, dv, && (3.9)\\ \mathrm{Var}(\mathbb{A}_{C,h_k^{(2)}}(t)) &= (k+1)^4 \int_0^1 \int_0^1 \sigma(u,v;t)(uv)^{k-A(t)}\, du\, dv, && (3.10) \end{aligned}$$

where the function σ is given by

$$\sigma(u,v;t) = \mathrm{Cov}\left(\mathbb{G}_C(u^{1-t}, u^t), \mathbb{G}_C(v^{1-t}, v^t)\right).$$

In order to find an explicit expression for these variances we assume that the function A is differentiable and use a similar argument as in Genest and Segers (2009). To be precise, we introduce the notation

$$\mu(t) = A(t) - tA'(t), \qquad \nu(t) = A(t) + (1-t)A'(t),$$

where A' denotes the derivative of A. Furthermore, the asymptotic variance of $\mathbb{A}_{C,h_\alpha^{(1)}}(t)$ involves the symmetric incomplete beta function, defined by

$$B_z(a) = \int_0^z x^{a-1}(1-x)^{a-1}\, dx$$

and $B(a) = B_1(a)$. The following results are proved in the Appendix.

Proposition 3.4. *For $t \in [0,1]$ let $\bar{\mu}(t) = 1-\mu(t)$ and $\bar{\nu}(t) = 1-\nu(t)$. If C is an extreme-value copula with Pickands dependence function A and if $A(t) \neq 1$, then the variance of $\mathbb{A}_{C,h^{(1)}_\alpha}(t)$ is given by*

$$\frac{1}{(3-\alpha)^2 B(3-\alpha)} \Big\{ \frac{2B_{\frac{1-A(t)}{2-A(t)}}(3-\alpha)}{(1-A(t))^{3-\alpha}} - (\mu(t)+\nu(t)-1)^2 B(3-\alpha)$$
$$- \frac{2\mu(t)\bar{\mu}(t)B_{\frac{t}{t+1}}(3-\alpha)}{t^{3-\alpha}} - \frac{2\nu(t)\bar{\nu}(t)B_{\frac{1-t}{2-t}}(3-\alpha)}{(1-t)^{3-\alpha}}$$
$$+ \frac{2\mu(t)\nu(t)}{(1-t)t} \int_0^1 \Big(\frac{(1-s)s}{(1-t)t}\Big)^{2-\alpha} \Big(A(s) + \frac{1-s}{1-t} + \frac{s}{t} - 1\Big)^{2\alpha-6} ds$$
$$- \frac{2\mu(t)}{(1-t)t} \int_0^t \Big(\frac{(1-s)s}{(1-t)t}\Big)^{2-\alpha} \Big(A(s) + \frac{1-s}{1-t}t + \frac{s}{t}(1-A(t))\Big)^{2\alpha-6} ds$$
$$- \frac{2\nu(t)}{(1-t)t} \int_t^1 \Big(\frac{(1-s)s}{(1-t)t}\Big)^{2-\alpha} \Big(A(s) + \frac{1-s}{1-t}(1-A(t)) + \frac{s}{t}(1-t)\Big)^{2\alpha-6} ds \Big\},$$

while for $A(t) = 1$ the first summand inside the brackets, i.e.

$$\frac{2B_{\frac{1-A(t)}{2-A(t)}}(3-\alpha)}{(1-A(t))^{3-\alpha}},$$

has to replaced by its continuous extension $\frac{2}{3-\alpha}$.

Proposition 3.5. *For $t \in [0,1]$ let $\bar{\mu}(t) = 1-\mu(t)$ and $\bar{\nu}(t) = 1-\nu(t)$. If C is an extreme-value copula with Pickands dependence function A, then the variance of the random variable $\mathbb{A}_{C,h^{(2)}_k}(t)$ is given by*

$$(k+1)^2 \Big\{ \frac{2(k+1)}{2k+2-A(t)} - (\mu(t)+\nu(t)-1)^2 - \frac{2\mu(t)\bar{\mu}(t)(k+1)}{2k+1+t} - \frac{2\nu(t)\bar{\nu}(t)(k+1)}{2k+2-t}$$
$$+ 2\mu(t)\nu(t)\frac{(k+1)^2}{(1-t)t} \int_0^1 \Big(A(s) + (k+1)\Big(\frac{1-s}{1-t} + \frac{s}{t}\Big) - 1\Big)^{-2} ds$$
$$- 2\mu(t)\frac{(k+1)^2}{(1-t)t} \int_0^t \Big(A(s) + (k+t)\frac{1-s}{1-t} + (k+1-A(t))\frac{s}{t}\Big)^{-2} ds$$
$$- 2\nu(t)\frac{(k+1)^2}{(1-t)t} \int_t^1 \Big(A(s) + (k+1-A(t))\frac{1-s}{1-t} + (k+1-t)\frac{s}{t}\Big)^{-2} ds \Big\}.$$

It might be of interest to compare the behavior of the new estimators $\hat{A}_{n,h}(t)$ with estimators investigated by Genest and Segers (2009). Some finite sample results will be presented in the following section for various families of copulas. For a theoretical comparison we restrict ourselves to the weight functions $h^{(2)}_k$ and consider the independence copula Π, for which $A(t) =$

1. In the case $k = 0$ we obtain from Proposition 3.5 the same variance as for the rank based version of Pickands estimator, that is

$$\mathrm{Var}(\mathbb{A}_{\Pi,h_0^{(2)}}) = \frac{3t(1-t)}{(2-t)(1+t)},$$

while the case $k > 0$ yields

$$\mathrm{Var}(\mathbb{A}_{\Pi,h_k^{(2)}}) = \frac{(3+4k)(k+1)^2}{2k+1} \frac{t(1-t)}{(2k+2-t)(2k+1+t)}.$$

It is easy to see that $\mathrm{Var}(\mathbb{A}_{\Pi,h_k^{(2)}})$ is decreasing in k with

$$\lim_{k\to\infty} \mathrm{Var}(\mathbb{A}_{\Pi,h_k^{(2)}}) = \frac{t(1-t)}{2}.$$

Moreover, a straightforward calculation shows that

$$\mathrm{Var}(\mathbb{A}_{\Pi,h_0^{(2)}}) \geq \mathrm{Var}(\mathbb{A}_{\Pi,h_k^{(2)}})$$

for all $k \geq 0$ with strict inequality for all $k > 0$. This means that for the independence copula all estimators obtained by our approach with weight function $h_k^{(2)}(x) = x^{-2}e^{-kx}$, $k > 0$ have a smaller asymptotic variances than the rank based version of Pickands estimator. On the other hand a comparison with the rank based CFG estimator investigated by Genest and Segers (2009) does not provide a clear picture about the superiority of one estimator. In Figure 2 we show the asymptotic variances of the rank based CFG estimator and the new estimators for the choice $k = 1$, $k = 5$ and $k = 10$. We observe that Pickands estimator has the largest asymptotic variances (which are not displayed in the figure), while the CFG estimator yields smaller variances than the estimator $\hat{A}_{n,h_1^{(2)}}$, but larger asymptotic variances than the estimators $\hat{A}_{n,h_5^{(2)}}$ and $\hat{A}_{n,h_{10}^{(2)}}$. Note that for finite sample size an increase in k will decrease the variance but increase the bias and therefore the asymptotic results can not directly be transferred to applications. Nevertheless, in the finite sample study presented in the following paragraph the superiority of the new estimators $\hat{A}_{h_k^{(2)}}$ over the rank based Pickands estimator is also observed for other extreme-value copulas, provided that the parameter k is not chosen as too large. Moreover, the estimators $\hat{A}_{n,h_1^{(2)}}$ and $\hat{A}_{n,h_2^{(2)}}$ also yields a smaller mean squared error than the rank based CFG estimator.

3.3 Finite sample properties

In this subsection we investigate the small sample properties of the new estimators by means of a simulation study. Especially, we compare the new estimators with the rank-based estimators suggested by Genest and Segers (2009), which are most similar in spirit with the method proposed in this paper. All results presented here are based on 5000 simulation runs and the sample size is $n = 100$. As estimators we consider the statistics defined in (3.5) and (3.6). It

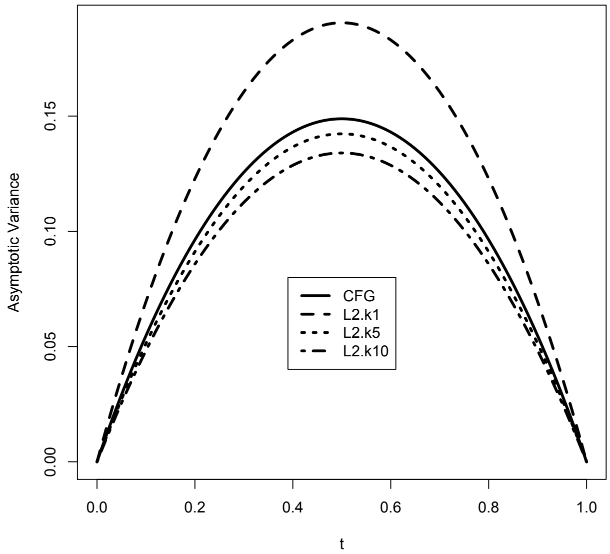

Figure 2: *Asymptotic variances of various estimators of the Pickands dependence function.*

turns out that the estimators obtained by the weight function $h_\alpha^{(1)}$ show a substantially worse behavior than the estimators $\hat{A}_{n,h_k^{(2)}}$, and for this reason we restrict the investigations to the latter class. Results for the estimator $\hat{A}_{n,h_\alpha^{(1)}}$ are available from the first author. An important question in the class $\{\hat{A}_{n,h_k^{(2)}} \mid k \geq 0\}$ is the choice of the parameter k in order to achieve a balance between bias and variance. For this purpose, we first study the performance of the estimator $\hat{A}_{n,h_k^{(2)}}$ with respect to different choices for the parameter k and consider the following four extreme-value copula models.

(i) The symmetric model of Gumbel [see Gumbel (1960)],

$$A(t) = \left(t^\theta + (1-t)^\theta\right)^{1/\theta}$$

with parameter $\theta \in [1, \infty)$. Complete dependence is obtained in the limit as θ approaches infinity. Independence is obtained when $\theta = 1$. The coefficient of tail dependence $\rho = 2(1 - A(0.5))$ is given by $\rho = 2 - 2^{1/\theta}$.

(ii) The model of Hüsler and Reiss [see Hüsler and Reiss (1989)]

$$A(t) = (1-t)\Phi\left(\lambda + \frac{1}{2\theta}\log\frac{1-t}{t}\right) + t\Phi\left(\theta + \frac{1}{2\theta}\log\frac{t}{1-t}\right),$$

where $\theta \in (0, \infty)$ and Φ is the standard normal distribution function. The coefficient of tail dependence is given by $\rho = 2(1 - \Phi(\theta))$, i.e. independence is obtained for $\theta \to \infty$ and complete dependence for $\theta \to 0$.

(iii) The asymmetric negative logistic model [see Joe (1990)]

$$A(t) = 1 - \left\{(\psi_1(1-t))^{-\theta} + (\psi_2 t)^{-\theta}\right\}^{-1/\theta}$$

with parameters $\theta \in (0,\infty), \psi_1, \psi_2 \in (0,1]$. For the simulations we set $\psi_1 = 2/3$ and $\psi_2 = 1$, then the coefficient of tail dependence is given by $\rho = 2\,(3^\theta + 2^\theta)^{-1/\theta}$ and varies in the interval $(0, 2/3)$.

(iv) The symmetric mixed model [see Tawn (1988)]

$$A(t) = 1 - \theta t + \theta t^2$$

with parameter $\theta \in [0,1]$ and $\rho = \theta/2 \in [0, 1/2]$.

In Figure 3 we display $n \times$ MISE of the estimator $\hat{A}_{n,h_k^{(2)}}$ as a function of the parameter k in the weight function $h_k^{(2)}(x) = x^{-2}e^{-kx}$ for the four copula models with different coefficients of tail dependence. For each estimator, the empirical version of the mean integrated squared error,

$$\text{MISE} = \mathbb{E}\left[\int_0^1 (\hat{A}_{n,h_k^{(2)}}(t) - A(t))^2\,dt\right],$$

was computed by an average over the 5000 samples. **The estimators turned out to be rather robust with respect to the choice of the parameter γ in the definition of the process $\tilde{C}_n = \hat{C}_n \vee n^{-\gamma}$ provided that $\gamma \geq 2/3$. For this reason we use $\gamma = 0.95$ throughout this section.** All cases yield a very similar picture and suggest that the "optimal" k is slightly larger than one for weak tail dependence and at approximately 1.25 in case of independence. For stronger tail dependence the optimal k turns out to decrease and is close to 0.5 for perfect positive dependence. Based on these observations and further results which are not shown for the sake of brevity, we recommend to use the value $k = 1$ or $k = 1.25$ in practical applications.

Next we compare the new estimators with rank-based versions of Pickands estimator and the CFG estimator investigated in Genest and Segers (2009). In Figure 4, the normalized MISE is plotted as a function of the tail dependence parameter ρ for the four copula models, where the parameter θ is chosen in such a way, that the coefficient of tail dependence $\rho = 2(1 - A(0.5))$ varies over the specific range of the corresponding model. For each sample we computed the rank-based versions of Pickands estimator, the CFG estimator [see Genest and Segers (2009)] and two of the new estimators $\hat{A}_{n,h_k^{(2)}}$ $(k = 1, 2)$. Summarizing all four pictures one can conclude that in general the best results are obtained for our new estimator based on the weight function $h_k^{(2)}$ with $k = 1$ and $k = 2$. A comparison of the two estimators $\hat{A}_{n,h_1^{(2)}}$ and $\hat{A}_{n,h_2^{(2)}}$ shows that the choice $k = 1$ performs globally better than the choice $k = 2$ in the Gumbel, asymmetric negative logistic and symmetric mixed model. On the other hand the estimator $\hat{A}_{n,h_2^{(2)}}$ yields

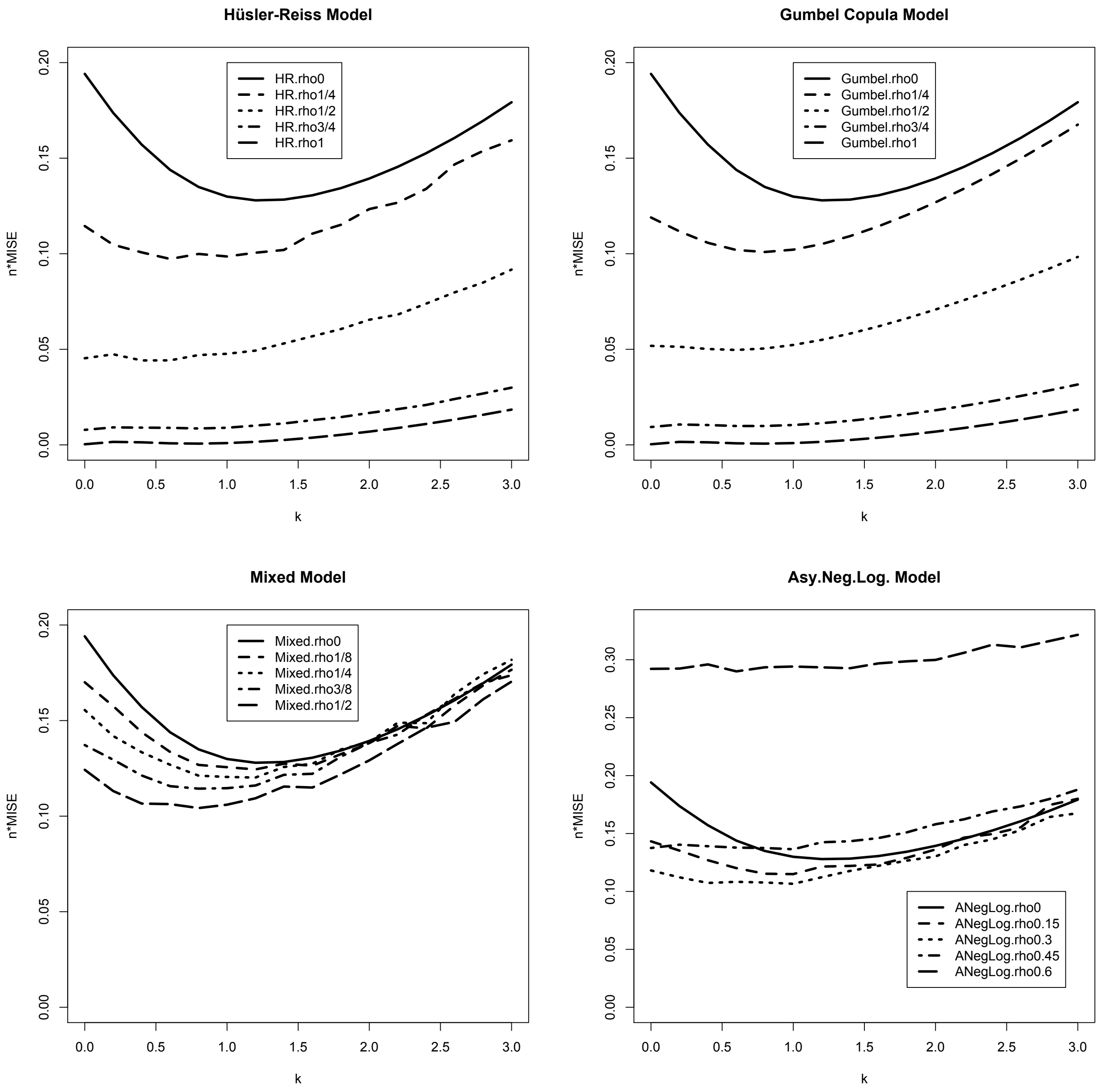

Figure 3: $100\times MISE$ *of the estimators* $\hat{A}_{n,h_k^{(2)}}$ *defined in* (3.6) *as a function of* k *for various models and coefficients of tail dependence. The sample is size* $n = 100$ *and the MISE is calculated by* 5000 *simulation runs.*

some advantages in the Hüsler and Reiss model if the coefficient of tail dependence is small. In all settings, the MISE obtained by $\hat{A}_{n,h_1^{(2)}}$ and $\hat{A}_{n,h_2^{(2)}}$ is smaller than the MISE of the CFG and Pickands estimator. On the other hand the latter estimators yield better results than the estimator $\hat{A}_{n,h_1^{(1)}}$ which corresponds to the weight function $h_1^{(1)}(x) = 1/x$ (these results are not depicted). Other scenarios yield similar results which are not displayed for the sake of brevity.

4 A test for an extreme-value dependence

4.1 The test statistic and its weak convergence

From the definition of the functional $M_C(A)$ in (2.1) it is easy to see that, for a strictly positive weight function h with $h^* \in L^1(0,1)$, a copula function C is an extreme-value copula if and only if

$$\min_{A\in\mathcal{A}} M_C(A) = M_C(A^*) = 0,$$

where A^* denotes the best approximation defined in (2.5). This suggests to use $M_{\tilde{C}_n}(\hat{A}_{n,h})$ as a test statistic for the hypothesis (1.2), i.e.

$$H_0 : C \quad \text{is an extreme-value copula.}$$

Recalling the representation (2.7)

$$M_C(A^*) = \int_0^1 \int_0^\infty \bar{C}^2(s,t) s e^{-s} h(s)\, ds\, dt - B_h \int_0^1 \big(A^*(t)\big)^2\, dt$$

with $\bar{C}(s,t) = -\log C(g^{-1}(s,t))$ and defining $\bar{C}_n(s,t) := -\log \tilde{C}_n(g^{-1}(s,t))$ we obtain the decomposition

$$\begin{aligned}
& M_{\tilde{C}_n}(\hat{A}_{n,h}) - M_C(A^*) \\
= & \int_0^1 \int_0^\infty \Big(\bar{C}_n^2(s,t) - \bar{C}^2(s,t)\Big) s e^{-s} h(s)\, ds\, dt - B_h \int_0^1 \hat{A}_{n,h}^2(t) - \big(A^*(t)\big)^2\, dt \\
= & \; 2\int_0^1 \int_0^\infty \Big(\bar{C}_n(s,t) - \bar{C}(s,t)\Big)\bar{C}(s,t) s e^{-s} h(s)\, ds\, dt - 2B_h \int_0^1 (\hat{A}_{n,h}(t) - A^*(t))A^*(t)\, dt \\
& + \int_0^1 \int_0^\infty \Big(\bar{C}_n(s,t) - \bar{C}(s,t)\Big)^2 s e^{-s} h(s)\, ds\, dt - B_h \int_0^1 (\hat{A}_{n,h}(t) - A^*(t))^2\, dt \\
= & \; 2\int_0^1 \int_0^\infty \Big(\bar{C}_n(s,t) - \bar{C}(s,t)\Big)\Big(\bar{C}(s,t) - sA^*(t)\Big) s e^{-s} h(s)\, ds\, dt \\
& + \int_0^1 \int_0^\infty \Big(\bar{C}_n(s,t) - \bar{C}(s,t)\Big)^2 s e^{-s} h(s)\, ds\, dt - B_h \int_0^1 (\hat{A}_{n,h}(t) - A^*(t))^2\, dt \\
=: & \; S_1 + S_2 + S_3,
\end{aligned}$$

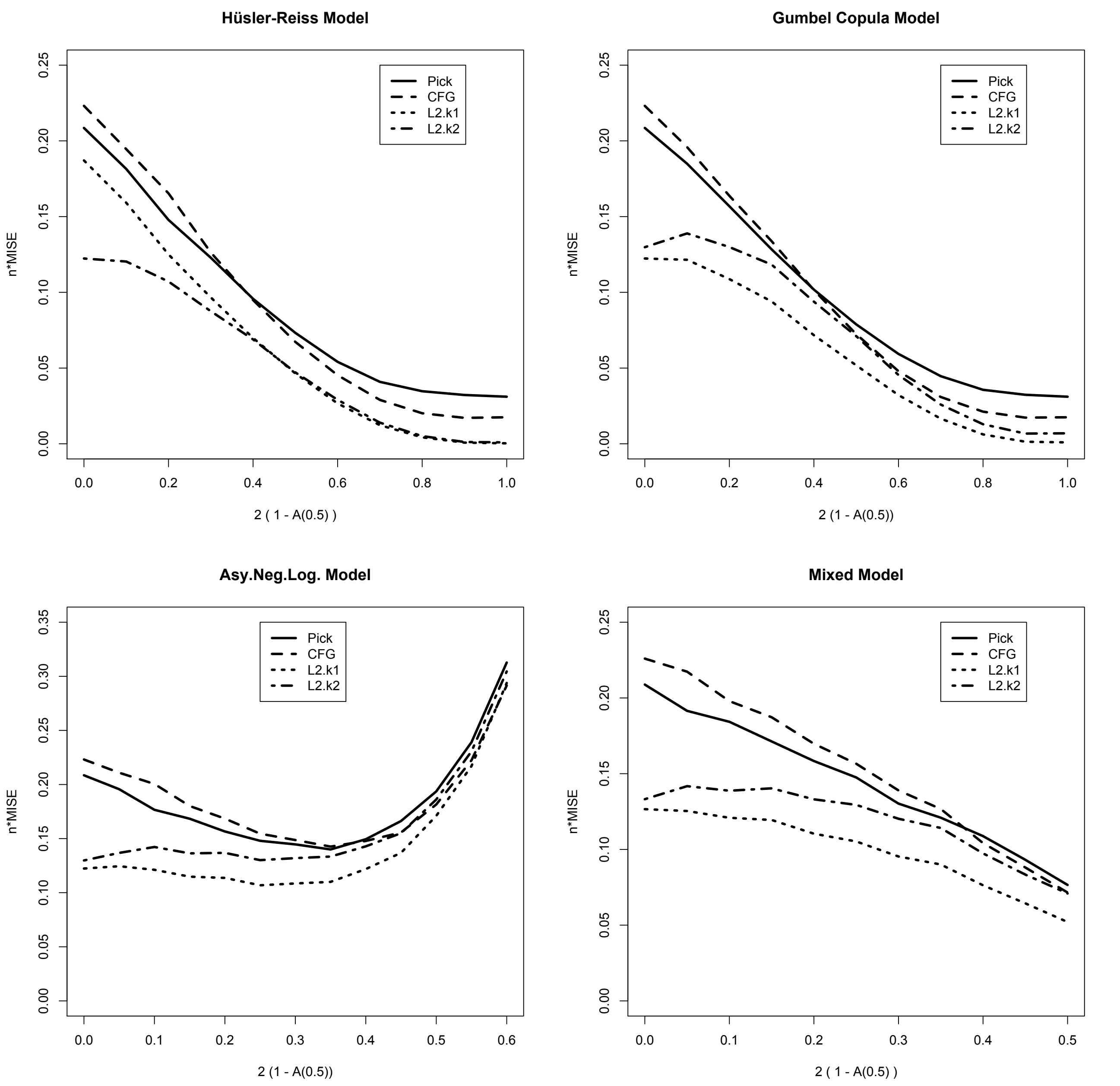

Figure 4: 100×*MISE for various estimators, models and coefficients of tail dependence, based on* 5000 *samples of size* $n = 100$.

where the last identity defines the terms S_1, S_2 and S_3 in an obvious manner. Note that under the null hypothesis of extreme-value dependence we have $A^* = A$ and thus $\bar{C}(s,t) = sA^*(t)$. This means that under H_0 the term S_1 will vanish and the asymptotic distribution will be determined by the large sample properties of the random variable S_2+S_3. Under the alternative the equality $\bar{C}(s,t) = sA^*(t)$ will not hold anymore and it turns out that in this case the statistic is asymptotically dominated by the random variable S_1. In order to derive the limiting distribution of the proposed test statistic under the null hypothesis and the alternative, we will need the following conditions on the corresponding weight functions. For some function $f : [0,1]^2 \to \bar{\mathbb{R}}$ assume that there exists a function $\bar{f} : [0,1] \to \bar{\mathbb{R}}_0^+$ such that

$$\forall\ \varepsilon > 0 : \sup_{y\in[\varepsilon,1]} \bar{f}(y) < \infty \tag{4.1}$$

$$\int_0^1 \bar{f}(y) y^{-\lambda} dy < \infty \tag{4.2}$$

$$\forall\ (y,t) \in [0,1]^2 : |f(y,t)| \leq \bar{f}(y) \tag{4.3}$$

where λ denotes some positive constant which will be specified later.

Theorem 4.1. *Assume that the given copula C is an extreme-value copula with continuous partial derivatives of first order and Pickands dependence function A^*. If the function $w(y) := (-\log y)h(-\log y)$ fulfills conditions* (4.1) - (4.3) *for some $\lambda > 2$ and the weight function h is strictly positive and satisfies assumptions (2.2) - (2.4) for $\tilde{\lambda} := \lambda/2 > 1$, then we have for any $\gamma \in (1/2, \lambda/4)$ and $n \to \infty$*

$$nM_{\tilde{C}_n}(\hat{A}_{n,h}) \overset{w}{\rightsquigarrow} Z_0,$$

where the random variable Z_0 is defined by

$$Z_0 := \int_0^1 \int_0^1 \Big(\frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} \Big)^2 w(y)\, dy\, dt - B_h \int_0^1 \mathbb{A}_{C,h}^2(t)\, dt$$

with $B_h = \int_0^\infty x^3 e^{-x} h(x)\, dx$ and the process $\{\mathbb{A}_{C,h}(t)\}_{t\in[0,1]}$ as defined in Theorem 3.1.

The next theorem gives the distribution of the test statistic $M_{\tilde{C}_n}(\hat{A}_{n,h})$ under the alternative. Note that in this case we have $M_C(A^*) > 0$.

Theorem 4.2. *Assume that the given copula C has continuous partial derivatives of first order and satisfies $C \geq \Pi$ and $M_C(A^*) > 0$. If additionally the weight function h is strictly positive and h and the function $w(y) := (-\log y)h(-\log y)$ satisfy the assumptions* (2.2) - (2.4) *and* (4.1) - (4.3) *for some $\lambda > 1$, respectively, then we have for any $\gamma \in (1/2, (1+\lambda)/4 \wedge \lambda/2)$ and $n \to \infty$*

$$\sqrt{n}(M_{\tilde{C}_n}(\hat{A}) - M_C(A^*)) \overset{w}{\rightsquigarrow} Z_1,$$

where the random variable Z_1 is defined as

$$Z_1 = 2 \int_0^1 \int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} v(y,t)\, dy\, dt,$$

with

$$v(y,t) = (\log C(y^{1-t}, y^t) - \log(y)A^*(t))(-\log y)h(-\log y).$$

Remark 4.3.
(a) Note that the weight functions $h_k^{(2)}(x) = x^{-2}e^{-kx}$ satisfy the assumptions of Theorem 4.1 and 4.2 for $k > 1$ and $k > 0$, respectively.
(b) The preceding two theorems yield a consistent asymptotic level α test for the hypothesis of extreme-value dependence by rejecting the null hypothesis H_0 if

$$n\, M_{\tilde{C}_n}(\hat{A}_{n,h}) > z_{1-\alpha}, \tag{4.4}$$

where $z_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the distribution of the random variable Z_0.
(c) By Theorem 4.2 the power of the test (4.4) is approximately given by

$$\mathbb{P}\left(n\, M_{\tilde{C}_n}(\hat{A}_{n,h}) > z_{1-\alpha}\right) \approx 1 - \Phi\left(\frac{z_{1-\alpha}}{\sqrt{n}\,\sigma} - \sqrt{n}\,\frac{M_C(A^*)}{\sigma}\right) \approx \Phi\left(\sqrt{n}\,\frac{M_C(A^*)}{\sigma}\right),$$

where the function A^* is defined in (2.5) corresponding to the best approximation of the copula C by an extreme-value copula, σ is the standard deviation of the distribution of the random variable Z_1 and Φ is the standard normal distribution function. Thus the power of the test (4.4) is an increasing function of the quantity $M_C(A^*)\,\sigma^{-1}$.
(d) In general the distribution of the random variable of the Z_0 can not be determined explicitly, because of its complicated dependence on the (unknown) copula C. For this purpose we propose to determine the quantiles by the multiplier bootstrap approach as described in Bücher and Dette (2009). To be precise let $\xi_1, \ldots, \xi_n$ denote independent identically distributed random variables with $P(\xi_1 = 0) = P(\xi_1 = 2) = 1/2$, We define $\bar{\xi}_n = n^{-1}\sum_{i=1}^n \xi_i$ as the mean of $\xi_1, \ldots, \xi_n$ and consider the multiplier statistics

$$\hat{C}_n^*(u_1, u_2) \;=\; \hat{F}_n^*(\hat{F}_{n1}^-(u_1), \hat{F}_{n2}^-(u_2)),$$

where $\hat{F}_n^*(x_1, x_2) = \frac{1}{n}\sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n}\mathbb{I}\{X_{i1} \le x_1, X_{i2} \le x_2\}$ and $\hat{F}_{nj}$ denotes the marginal empirical distribution functions. If we estimate the partial derivatives of the copula C by

$$\widehat{\partial_1 C}(u,v) \;:=\; \frac{\hat{C}_n(u+h, v) - \hat{C}_n(u-h, v)}{2h},$$
$$\widehat{\partial_2 C}(u,v) \;:=\; \frac{\hat{C}_n(u, v+h) - \hat{C}_n(u, v-h)}{2h},$$

where $h = n^{-1/2} \to 0$, then we can approximate the distribution of $\mathbb{G}_C$ by the distribution of the process

$$\hat{\alpha}_n^{pdm}(u_1, u_2) := \hat{\beta}_n(u_1, u_2) - \widehat{\partial_1 C}(u_1, u_2)\hat{\beta}_n(u_1, 1) - \widehat{\partial_2 C}(u_1, u_2)\hat{\beta}_n(1, u_2), \tag{4.5}$$

where $\hat{\beta}_n(u_1, u_2) = \sqrt{n}\,(\hat{C}_n^*(u_1, u_2) - \hat{C}_n(u_1, u_2))$. More precisely, it was shown by Bücher and Dette (2009) that we have weak convergence conditional on the data in probability towards $\mathbb{G}_C$, i.e.

$$\hat{\alpha}_n^{pdm} \overset{\mathbb{P}}{\underset{\xi}{\rightsquigarrow}} \mathbb{G}_C \quad \text{in} \;\; l^\infty[0,1]^2.$$

Since Z_0 is a continuous function of $(\mathbb{G}_C, C)$ we obtain that the distribution of

$$\begin{aligned} \hat{Z}_0^* &= \int_0^1 \int_0^1 \Big(\frac{\hat{\alpha}_n^{pdm}(y^{1-t}, y^t)}{\tilde{C}_n(y^{1-t}, y^t)} \Big)^2 (-\log y) h(-\log y)\, dy\, dt \\ &\quad - B_h^{-1} \int_0^1 \Big(\int_0^1 \frac{\hat{\alpha}_n^{pdm}(y^{1-t}, y^t)}{\tilde{C}_n(y^{1-t}, y^t)} (\log y)^2 h(-\log y)\, dy \Big)^2 dt \end{aligned}$$

gives a valid approximation for the distribution of Z_0 in the sense that $\hat{Z}_0^* \overset{\mathbb{P}}{\underset{\xi}{\rightsquigarrow}} Z_0$. Repeating this procedure B times yields a sample $\hat{Z}_0^*(1), \ldots, \hat{Z}_0^*(B)$ that is approximately distributed according to Z_0 and we can use the empirical $(1-\alpha)$-quantile of this sample, say $z_{1-\alpha}^*$, as an approximation for $z_{1-\alpha}$. Therefore rejecting the null hypothesis if

$$n\, M_{\tilde{C}_n}(\hat{A}_{n,h}) > z_{1-\alpha}^* \qquad (4.6)$$

yields a consistent asymptotic level α test for extreme-value dependence.

4.2 Finite sample properties

In this subsection we investigate the finite sample properties of the test for extreme-value dependence. We generated 1000 random samples of sample size $n = 200$ from various copula models and calculated the probability of rejecting the null hypothesis. Under the null hypothesis we chose the model parameters in such a way that the coefficient of tail dependence ρ varies over the specific range of the corresponding model. Under the alternative the coefficient of tail dependence does not need to exist and we therefore chose the model parameters, such that Kendall's τ is an element of the set $\{1/4, 1/2, 3/4\}$. The weight function is chosen according to the suggestion in the previous section as $h_{1.25}^{(2)}(x) = x^{-2} e^{-1.25x}$ and the critical values are determined by the multiplier bootstrap approach as described in Remark 4.3 with $B = 200$ Bootstrap replications. The results are stated in Table 1.

We observe from the left part of Table 1 that the level of test is accurately approximated for most of the models, if the tail dependence is not too strong. For a large tail dependence coefficient the bootstrap test is conservative. This phenomenon can be explained by the fact that for the limiting case of random variables distributed according to the upper Fréchet-Hoeffding the empirical copula $\hat{C}_n$ does not converge weakly to a non-degenerate process at a rate $1/\sqrt{n}$, rather in this case it follows that $||\hat{C}_n - C|| = O(1/n)$. Consequently, the approximations

H_0-model	ρ	0.05	0.1
Independence	0	0.046	0.092
Gumbel	0.25	0.066	0.113
	0.5	0.044	0.098
	0.75	0.016	0.043
Mixed model	0.25	0.062	0.117
	0.5	0.04	0.099
Asy. Neg. Log.	0.25	0.06	0.123
	0.5	0.07	0.123
Hüsler-Reiß	0.25	0.069	0.127
	0.5	0.038	0.099
	0.75	0.008	0.03

H_1-model	τ	0.05	0.1
Clayton	0.25	0.693	0.798
	0.5	0.988	0.994
	0.75	0.992	1
Frank	0.25	0.325	0.427
	0.5	0.746	0.842
	0.75	0.694	0.831
Gaussian	0.25	0.111	0.189
	0.5	0.166	0.246
	0.75	0.036	0.072
t_4	0.25	0.058	0.119
	0.5	0.062	0.12
	0.75	0.015	0.039

Table 1: *Simulated rejection probabilities of the test* (4.6) *for the null hypothesis of an extreme-value copula for various models. The first four columns deal with models under the null hypothesis, while the last four are from the alternative.*

proposed in this paper, which are based on the weak convergence of $\sqrt{n}(\hat{C}_n - C)$ to a non-degenerate process, are not appropriate for small samples, if the tail dependence coefficient is large.

Considering the alternative we observe reasonably good power for the Frank and Clayton copulas, while for the Gaussian or t-copula deviations from an extreme-value copula can not be detected with a sample size $n = 200$. In some cases the test (4.6) even underestimates the nominal level. This observation can be explained by the closeness to the upper Fréchet-Hoeffding bound again. Indeed, we can use the minimal distance $M_C(A^*)$ as a measure of deviation from an extreme-value copula. Calculating the minimal distance $M_C(A^*)$ [with Kendall's $\tau = 0.5$ and $h = h^{(2)}_{1.25}$] we observe that the minimal distances are about ten times smaller for the Gaussian and t_4 than for the Frank and Clayton copula, i.e.

$$M_C(A^*_{\text{Clayton}}) = 6.52 \times 10^{-4}, \qquad M_C(A^*_{\text{Frank}}) = 3.53 \times 10^{-4},$$
$$M_C(A^*_{\text{Gaussian}}) = 9.00 \times 10^{-5}, \qquad M_C(A^*_{t_4}) = 3.56 \times 10^{-5}.$$

Moreover, as explained in Remark 4.3 (b) the power of the tests (4.4) and (4.6) is an increasing function of the quantity $p(\text{copula}) = M_C(A^*)\,\sigma^{-1}$. For the four copulas considered in the simulation study [with $\tau = 0.5$] the corresponding ratios are approximately given by

$$p(\text{Clayton}) = 0.210, \quad p(\text{Frank}) = 0.147, \quad p(\text{Gaussian}) = 0.075, \quad p(t_4) = 0.050,$$

which provides some heuristic explanation of the findings presented in Table 1. Loosely speaking, if the value $M_C(A^*)\,\sigma^{-1}$ is very small a larger sample size is required to detect a deviation from an extreme-value copula. This statement is confirmed by further simulations results. For example, for the Gaussian and t_4 copula (with Kendall's $\tau = 0.75$) we obtain for the sample size $n = 500$ the rejection probabilities 0.339 (0.495) and 0.142 (0.244) for the bootstrap test with level 5% (10%), respectively.

Acknowledgements The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG). The authors would also like to thank Christian Genest for pointing out important references on the subject.

5 Appendix: Proofs

Many of the proofs that follow are based on a general result which shows weak convergence for the weigthed integrated process $\log \tilde{C}_n = \log(C_n \vee n^{-\gamma})$ where the weight function depends on y and t.

Theorem 5.1. *Denote by $w : [0,1]^2 \to \bar{\mathbb{R}}$ some weight function. Assume that the copula C has continuous partial derivatives of first order, that $C \geq \Pi$ and that the function w satisfies conditions* (4.1)-(4.3) *for some $\lambda > 1$. Then we have for any $\gamma \in (1/2, \lambda/2)$ as $n \to \infty$*

$$\sqrt{n}\mathbb{W}_{n,w} = \sqrt{n} \int_0^1 \log \frac{\tilde{C}_n(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} w(y,t)\, dy \stackrel{w}{\rightsquigarrow} \mathbb{W}_{C,w} = \int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} w(y,t)\, dy \qquad (5.1)$$

in $l^\infty[0,1]$.

Proof of Theorem 5.1. Fix $\lambda > 1$ as in (4.2) and $\gamma \in (1/2, \lambda/2)$, then choose some some $\alpha \in (0, 1/2)$ such that $\lambda\alpha > \gamma$. Due to Lemma 1.10.2 (i) in Van der Vaart and Wellner (1996), the process $\sqrt{n}(\tilde{C}_n - C)$ will have the same weak limit (with respect to the $\stackrel{w}{\rightsquigarrow}$ convergence) as $\sqrt{n}(\hat{C}_n - C)$.

For $i = 2, 3, \ldots$ we consider the following random functions in $l^\infty[0,1]$

$$\begin{aligned}
X_n(t) &= \int_0^1 \sqrt{n}\big(\log \tilde{C}_n(y^{1-t}, y^t) - \log C(y^{1-t}, y^t)\big) w(y,t)\, dy, \\
X_{i,n}(t) &= \int_{1/i}^1 \sqrt{n}\big(\log \tilde{C}_n(y^{1-t}, y^t) - \log C(y^{1-t}, y^t)\big) w(y,t)\, dy, \\
X(t) &= \int_0^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} w(y,t)\, dy, \\
X_i(t) &= \int_{1/i}^1 \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} w(y,t)\, dy,
\end{aligned}$$

We prove the theorem by an application of Theorem 4.2 in Billingsley (1968), adapted to the concept of weak convergence in the sense of Hoffmann-Jørgensen, see e.g. Van der Vaart and Wellner (1996). More precisely, we will show in Lemma 6.1 in Section 6 that the weak convergence $X_n \overset{w}{\rightsquigarrow} X$ in $l^\infty[0,1]$ follows from the following three assertions

$$\begin{aligned} &\text{(i)} \quad \text{For every } i \geq 2: \;\; X_{i,n} \overset{w}{\rightsquigarrow} X_i \;\; \text{for } n \to \infty \;\; \text{in } l^\infty[0,1], \\ &\text{(ii)} \quad X_i \overset{w}{\rightsquigarrow} X \;\; \text{for } i \to \infty \;\; \text{in } l^\infty[0,1], \\ &\text{(iii)} \quad \text{For every } \varepsilon > 0: \;\; \lim_{i\to\infty} \limsup_{n\to\infty} \mathbb{P}^*\big(\sup_{t\in[0,1]} |X_{i,n}(t) - X_n(t)| > \varepsilon \big) = 0. \end{aligned} \tag{5.2}$$

We begin by proving assertion (i). For this purpose set $T_i = [1/i, 1]^2$ and consider the space $\mathbb{D}_{\Phi_1}$ defined by $\mathbb{D}_{\Phi_1} = \{f \in l^\infty(T_i) : \inf_{x\in T_i} |f(x)| > 0\} \subset l^\infty(T_i)$. By Lemma 12.2 in Kosorok (2008) it follows that the mapping

$$\Phi_1 : \begin{cases} \mathbb{D}_{\Phi_1} \to l^\infty(T_i) \\ f \mapsto \log \circ f \end{cases}$$

is Hadamard-differentiable at C, tangentially to $l^\infty(T_i)$, with derivative $\Phi'_{1,C}(f) = f/C$. Since $\tilde{C}_n \geq n^{-\gamma}$ and $C \geq \Pi$ we have $\tilde{C}_n, C \in \mathbb{D}_{\Phi_1}$ and the functional delta method [see Theorem 2.8 in Kosorok (2008)] yields

$$\sqrt{n}(\log \tilde{C}_n - \log C) \overset{w}{\rightsquigarrow} \mathbb{G}_C/C$$

in $l^\infty(T_i)$. Next we consider the operator

$$\Phi_2 : \begin{cases} l^\infty(T_i) \to l^\infty([1/i, 1] \times [0,1]) \\ f \mapsto f \circ \varphi, \end{cases} \tag{5.3}$$

where the mapping $\varphi : [1/i, 1] \times [0,1] \to T_i$ is defined by $\varphi(y,t) = (y^{1-t}, y^t)$. Observing

$$\sup_{(y,t)\in[1/i,1]\times[0,1]} |f \circ \varphi(y,t) - g \circ \varphi(y,t)| \leq \sup_{\mathbf{x}\in T_i} |f(\mathbf{x}) - g(\mathbf{x})|$$

we can conclude that Φ_2 is Lipschitz-continuous. By the continuous mapping theorem [see e.g. Theorem 7.7 in Kosorok (2008)] and conditions (4.1) and (4.3) we immediately obtain

$$\sqrt{n}\big(\log \tilde{C}_n(y^{1-t}, y^t) - \log C(y^{1-t}, y^t)\big) w(y,t) \overset{w}{\rightsquigarrow} \frac{\mathbb{G}_C(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} w(y,t)$$

in $l^\infty([1/i, 1] \times [0,1])$. The assertion in (i) now follows by continuity of integration with respect to the variable y.

For the proof of assertion (ii) we simply note that $\mathbb{G}_C$ is bounded on $[0,1]^2$ and that

$$K(y,t) = \frac{w(y,t)}{C(y^{1-t}, y^t)}$$

is uniformly bounded with respect to $t \in [0,1]$ by the integrable function $\bar{K}(y) = \bar{w}(y)\, y^{-1}$.

For the proof of assertion (iii) recall that we fixed some $\alpha < 1/2$ at the beginning of the proof, and consider the decomposition

$$X_n(t) - X_{i,n}(t) = \int_0^{1/i} \sqrt{n}\big(\log \tilde{C}_n(y^{1-t}, y^t) - \log C(y^{1-t}, y^t)\big) w(y,t)\, dy = B_i^{(1)}(t) + B_i^{(2)}(t), \quad (5.4)$$

where

$$B_i^{(j)}(t) \;=\; \int_{I_{B_i^{(j)}(t)}} \sqrt{n} \log \frac{\tilde{C}_n}{C}(y^{1-t}, y^t) w(y,t)\, dy, \;\; j = 1,2 \quad (5.5)$$

and

$$I_{B_i^{(1)}(t)} = \left\{0 < y < 1/i \,|\, C(y^{1-t}, y^t) > n^{-\alpha}\right\}, \;\; I_{B_i^{(2)}(t)} = I^C_{B_i^{(1)}(t)} \cap (0, 1/i). \quad (5.6)$$

The usual estimate

$$\mathbb{P}^*(\sup_{t\in[0,1]} |X_{i,n}(t) - X_n(t)| > \varepsilon) \le \mathbb{P}^*(\sup_{t\in[0,1]} |B_i^{(1)} t)| > \varepsilon/2) + \mathbb{P}^*(\sup_{t\in[0,1]} |B_i^{(2)}(t)| > \varepsilon/2) \quad (5.7)$$

allows for individual investigation of both terms, and we begin with $\sup_{t\in[0,1]} |B_i^{(1)}(t)|$. By the mean value theorem we have

$$\log \frac{\tilde{C}_n}{C}(y^{1-t}, y^t) = (\tilde{C}_n - C)(y^{1-t}, y^t)\, \frac{1}{C^*(y,t)}, \quad (5.8)$$

where $|C^*(y,t) - C(y^{1-t}, y^t)| \le |\tilde{C}_n(y^{1-t}, y^t) - C(y^{1-t}, y^t)|$. Especially, observing $C \ge \Pi$ we have

$$C^*(y,t) \ge (C \wedge \tilde{C}_n)(y^{1-t}, y^t) \ge y \wedge \left(y \, \frac{\tilde{C}_n}{C}(y^{1-t}, y^t) \right) \quad (5.9)$$

and therefore

$$\begin{aligned} \sup_{t\in[0,1]} |B_i^{(1)}(t)| \;&\le\; \sup_{t\in[0,1]} \int_{I_{B_i^{(1)}(t)}} \sqrt{n}\left|(\tilde{C}_n - C)(y^{1-t}, y^t)\right| \times \left|1 \vee \frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\right| w(y,t)\, y^{-1}\, dy \\ &\le\; \sup_{\mathbf{x}\in[0,1]^2} \sqrt{n}|\tilde{C}_n(\mathbf{x}) - C(\mathbf{x})| \times \Big(1 \vee \sup_{\mathbf{x}\in[0,1]^2 \,:\, C(\mathbf{x})>n^{-\alpha}} \Big|\frac{C}{\tilde{C}_n}(\mathbf{x})\Big|\Big) \times \psi(i), \end{aligned}$$

with $\psi(i) = \int_0^{1/i} \bar{w}(y)\, y^{-1}\, dy = o(1)$ for $i \to \infty$. This yields for the first term on the right hand side of (5.7)

$$\begin{aligned} &\mathbb{P}^*(\sup_{t\in[0,1]} |B_i^{(1)}(t)| > \varepsilon) \\ &\le\; \mathbb{P}^*\Big(\sup_{\mathbf{x}\in[0,1]^2} \sqrt{n}|\tilde{C}_n(\mathbf{x}) - C(\mathbf{x})| > \sqrt{\frac{\varepsilon}{\psi(i)}}\Big) + \mathbb{P}^*\Big(1 \vee \sup_{C(\mathbf{x})>n^{-\alpha}} \Big|\frac{C}{\tilde{C}_n}(\mathbf{x})\Big| > \sqrt{\frac{\varepsilon}{\psi(i)}}\Big). \end{aligned} \quad (5.10)$$

Since $\sup_{\mathbf{x}\in[0,1]^2}\sqrt{n}|\tilde{C}_n(\mathbf{x})-C(\mathbf{x})|$ is asymptotically tight we immediately obtain

$$\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*\Big(\sup_{\mathbf{x}\in[0,1]^2}\sqrt{n}|\tilde{C}_n(\mathbf{x})-C(\mathbf{x})|>\sqrt{\frac{\varepsilon}{\psi(i)}}\Big)=0. \tag{5.11}$$

For the estimation of the second term in equation (5.10) we note that

$$\sup_{\mathbf{x}\in[0,1]^2\,:\,C(\mathbf{x})>n^{-\alpha}}\Big|\frac{\tilde{C}_n(\mathbf{x})-C(\mathbf{x})}{C(\mathbf{x})}\Big|<n^{\alpha}\sup_{\mathbf{x}\in[0,1]^2}|\tilde{C}_n(\mathbf{x})-C(\mathbf{x})|\xrightarrow{\mathbb{P}^*}0 \tag{5.12}$$

which in turn implies

$$\begin{aligned}\sup_{C(\mathbf{x})>n^{-\alpha}}\Big|\frac{C}{\tilde{C}_n}(\mathbf{x})\Big| &= \sup_{C(\mathbf{x})>n^{-\alpha}}\Big|1+\frac{\tilde{C}_n-C}{C}(\mathbf{x})\Big|^{-1}\\ &\le \Big(1-\sup_{C(\mathbf{x})>n^{-\alpha}}\Big|\frac{\tilde{C}_n-C}{C}(\mathbf{x})\Big|\Big)^{-1}\mathbb{I}_{A_n}+\Big(\sup_{C(\mathbf{x})>n^{-\alpha}}\Big|1+\frac{\tilde{C}_n-C}{C}(\mathbf{x})\Big|^{-1}\Big)\,\mathbb{I}_{A_n^C}\xrightarrow{\mathbb{P}^*}1,\end{aligned} \tag{5.13}$$

where $A_n=\{\sup_{C(\mathbf{x})>n^{-\alpha}}\Big|\frac{\tilde{C}_n-C}{C}(\mathbf{x})\Big|<1/2\}$. Thus the function $\max\{1,\sup_{C(\mathbf{x})>n^{-\alpha}}\Big|\frac{C}{\tilde{C}_n}(\mathbf{x})\Big|\}$ can be bounded by a function that converges to one in outer probability, which implies

$$\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*\Big(1\vee\sup_{C(\mathbf{x})>n^{-\alpha}}\Big|\frac{C}{\tilde{C}_n}(\mathbf{x})\Big|>\sqrt{\frac{\varepsilon}{\psi(i)}}\Big)=0.$$

Observing (5.10) and (5.11) it remains to estimate the second term on the right hand side of (5.7). We make use of the mean value theorem again, see equation (5.8), but use the estimate

$$C^*(y,t)\ge(C\wedge\tilde{C}_n)(y^{1-t},y^t)\ge y^{\lambda}\wedge y^{\lambda}\frac{\tilde{C}_n}{C^{\lambda}}(y^{1-t},y^t) \tag{5.14}$$

[recall that $\lambda>1$ by assumption (4.2)]. This yields

$$\begin{aligned}\sup_{t\in[0,1]}|B_i^{(2)}(t)| &\le \sup_{t\in[0,1]}\int_{I_{B_i^{(2)}(t)}}\sqrt{n}\Big|(\tilde{C}_n-C)(y^{1-t},y^t)\Big|\times\Big|1\vee\frac{C^{\lambda}}{\tilde{C}_n}(y^{1-t},y^t)\Big|\,w(y,t)\,y^{-\lambda}\,dy\\ &\le \sup_{\mathbf{x}\in[0,1]^2}\sqrt{n}|\tilde{C}_n(\mathbf{x})-C(\mathbf{x})|\times\Big(1\vee\sup_{\mathbf{x}\in[0,1]^2\,:\,C(\mathbf{x})\le n^{-\alpha}}\Big|\frac{C^{\lambda}}{\tilde{C}_n}(\mathbf{x})\Big|\Big)\times\phi(i),\end{aligned}$$

where $\phi(i)=\int_0^{1/i}\bar{w}(y)y^{-\lambda}\,dy=o(1)$ for $i\to\infty$ by condition (4.2). Using analogous arguments as for the estimation of $\sup_{t\in[0,1]}|B_i^{(1)}(t)|$ the assertion follows from

$$\sup_{\mathbf{x}\in[0,1]^2\,:\,C(\mathbf{x})\le n^{-\alpha}}\Big|\frac{C^{\lambda}}{\tilde{C}_n}(\mathbf{x})\Big|\le\sup_{\mathbf{x}\in[0,1]^2\,:\,C(\mathbf{x})\le n^{-\alpha}}\Big|n^{\gamma}C^{\lambda}(\mathbf{x})\Big|\le n^{\gamma-\lambda\alpha}=o(1)$$

due to the choice of γ and α. □

Proof of Theorem 3.1. This is a direct consequence of Theorem 5.1 using the weight function
$w(y,t):=-B_h^{-1}\log^2(y)h(-\log(y))$. □

Proof of Theorem 3.2. The proof will also be based on Lemma 6.1 in Section 6 verifying conditions (i) - (iii) in (5.2). A careful inspection of the previous proof shows that the verification of condition (i) in (5.2) remains valid. Regarding condition (ii) we have to show that the process $\frac{\mathbb{G}_C}{C}(y^{1-t}, y^t)$ is integrable on the interval $(0,1)$. For this purpose we write

$$\mathbb{G}_C(\mathbf{x}) = \mathbb{B}_C(\mathbf{x}) - \partial_1 C(\mathbf{x})\mathbb{B}_C(x_1, 1) - \partial_2 C(\mathbf{x})\mathbb{B}_C(1, x_2)$$

and consider each term separately. From Theorem G.1 in Genest and Segers (2009) we know that for any $\omega \in (0, 1/2)$ the process

$$\tilde{\mathbb{B}}_C(\mathbf{x}) = \begin{cases} \frac{\mathbb{B}_C(\mathbf{x})}{(x_1\wedge x_2)^\omega(1-x_1\wedge x_2)^\omega} & \text{, if } \quad x_1 \wedge x_2 \in (0,1) \\ 0 & \text{, if } \quad x_1 = 0 \text{ or } x_2 = 0 \text{ or } \mathbf{x} = (1,1), \end{cases}$$

has continuous sample paths on $[0,1]^2$. Considering $C(y^{1-t}, y^t) \geq y$ and using the notation

$$K_1(y,t) = q_\omega(y^{1-t} \wedge y^t)y^{-1} \tag{5.15}$$

$$K_2(y,t) = \partial_1 C(y^{1-t}, y^t) q_\omega(y^{1-t})y^{-1} \tag{5.16}$$

$$K_3(y,t) = \partial_2 C(y^{1-t}, y^t) q_\omega(y^t)y^{-1} \tag{5.17}$$

with $q_\omega(t) = t^\omega(1-t)^\omega$ it remains to show that there exist integrable functions $K_j^*(y)$ with $K_j(y,t) \leq K_j^*(y)$ for all $t \in [0,1]$ $(j = 1,2,3)$. For K_1 this is immediate because $K_1(y,t) \leq (y^{1-t} \wedge y^t)^\omega y^{-1} \leq y^{\omega/2-1}$. For K_2, note that $\partial_1 C(y^{1-t}, y^t) = \mu(t)\, y^{A(t)-(1-t)}$, with $\mu(t) = A(t) - tA'(t)$. Therefore

$$K_2(y,t) \leq \mu(t)\, y^{A(t)-(1-\omega)(1-t)-1} \leq \mu(t)\, y^{\omega/2-1} \leq 2\, y^{\omega/2-1}, \tag{5.18}$$

where the second estimate follows from the inequality $t \vee (1-t) \leq A(t) \leq 1$ and holds for $\omega \in (0,2)$. A similar argument works for the term K_3.

For the verification of condition (iii) we proceed along similar lines as in the previous proof. We begin by choosing some $\beta \in (1, 9/8), \omega \in (1/4, 1/2)$ and some $\alpha \in (4/9, \gamma \wedge (2-\omega)^{-1})$ in such a way that $\gamma < \beta\alpha$. First note that $y \leq 1/(n+2)^2$ implies $\tilde{C}_n(y^{1-t}, y^t) = n^{-\gamma}$ for all $t \in [0,1]$. This yields

$$\int_0^{(n+2)^{-2}} \sqrt{n}(\log \tilde{C}_n - \log C)(y^{1-t}, y^t)\, dy = O\Big(\frac{\log n}{n^{3/2}}\Big)$$

uniformly with respect to $t \in [0,1]$, and therefore it is sufficient to consider the decomposition in (5.4) with the sets

$$I_{B_i^{(1)}(t)} = \{1/(n+2)^2 < y < 1/i \,|\, C(y^{1-t}, y^t) > n^{-\alpha}\}, \quad I_{B_i^{(2)}(t)} = I^C_{B_i^{(1)}(t)} \cap (1/(n+2)^2, 1/i).$$

We can estimate the term $B_i^{(1)}(t)$ analogously to the previous proof by

$$|B_i^{(1)}(t)| \leq \int_{I_{B_i^{(1)}(t)}} \sqrt{n}\Big|(\tilde{C}_n - C)(y^{1-t}, y^t)\Big| \times \Big|1 \vee \frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\Big|\, y^{-1}\, dy.$$

Let F_n denote the empirical distribution function of $(F_1(X_{i1}), F_2(X_{i2})), \ldots, (F_1(X_{n1}), F_2(X_{n2}))$. By the results in Stute (1984) and Tsukahara (2005) we can decompose $\sqrt{n}(\tilde{C}_n - C) = \sqrt{n}(C_n \vee n^{-\gamma} - C)$ as follows

$$\begin{aligned}\sqrt{n}(\tilde{C}_n - C)(\mathbf{x}) &= \sqrt{n}(C_n - C)(\mathbf{x}) + \sqrt{n}(\tilde{C}_n - C_n)(\mathbf{x}) \\ &= \alpha_n(\mathbf{x}) - \partial_1 C(\mathbf{x})\alpha_n(x_1, 1) - \partial_2 C(\mathbf{x})\alpha_n(1, x_2) + \tilde{R}_n(\mathbf{x}), \qquad (5.19)\end{aligned}$$

where $\alpha_n(\mathbf{x}) = \sqrt{n}(F_n - F)(\mathbf{x})$ and the remainder satisfies

$$\sup_{\mathbf{x}\in[0,1]^2} |\tilde{R}_n(\mathbf{x})| = O(n^{1/2-\gamma} + n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}) \quad \text{a.s.} \qquad (5.20)$$

Note that the estimate of (5.20) requires continuity of all second order partial derivatives of the copula C. This condition is satisfied provided that the function A is assumed to be twice continuously differentiable. With (5.19) we can estimate the term $|B_i^{(1)}(t)|$ analogously to decomposition (5.4) by $B_{i,1}^{(1)}(t) + \cdots + B_{i,4}^{(1)}(t)$, where

$$\begin{aligned}B_{i,1}^{(1)}(t) &= \int_{I_{B_i^{(1)}(t)}} \left|\alpha_n(y^{1-t}, y^t)\right| \left|1 \vee \frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\right| y^{-1}\, dy, \\ B_{i,2}^{(1)}(t) &= \int_{I_{B_i^{(1)}(t)}} \partial_1 C(y^{1-t}, y^t) \left|\alpha_n(y^{1-t}, 1)\right| \left|1 \vee \frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\right| y^{-1}\, dy, \\ B_{i,3}^{(1)}(t) &= \int_{I_{B_i^{(1)}(t)}} \partial_2 C(y^{1-t}, y^t) \left|\alpha_n(1, y^t)\right| \left|1 \vee \frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\right| y^{-1}\, dy, \\ B_{i,4}^{(1)}(t) &= \int_{I_{B_i^{(1)}(t)}} \left|\tilde{R}_n(y^{1-t}, y^t)\right| \left|1 \vee \frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\right| y^{-1}\, dy.\end{aligned}$$

The decomposition in (5.19), Theorem G.1 in Genest and Segers (2009) and the inequality $\alpha < \gamma \wedge (2-\omega)^{-1}$ may be used to conclude

$$\sup_{(y,t):\, C(y^{1-t}, y^t) > n^{-\alpha}} \left|\frac{\tilde{C}_n - C}{C}(y^{1-t}, y^t)\right| = o_{\mathbb{P}^*}(1),$$

which in turn implies

$$1 \vee \sup_{(y,t):\, C(y^{1-t}, y^t) > n^{-\alpha}} \left|\frac{C}{\tilde{C}_n}(y^{1-t}, y^t)\right| = O_{\mathbb{P}^*}(1) \qquad (5.21)$$

analogously to (5.13). Together with (5.20) and the inequality $\int_{(n+2)^{-2}}^{1/i} y^{-1}\, dy \le 2\log(n+2)$ we obtain, for $n \to \infty$

$$\sup_{t\in[0,1]} B_{i,4}^{(1)}(t) = O_{\mathbb{P}^*}(n^{1/2-\gamma}\log n + n^{-1/4}(\log n)^{3/2}(\log\log n)^{1/4}) = o_{\mathbb{P}^*}(1),$$

which implies

$$\lim_{i\to\infty} \limsup_{n\to\infty} \mathbb{P}^*\Big(\sup_{t\in[0,1]} B_{i,4}^{(1)}(t) > \varepsilon/4\Big) = 0. \qquad (5.22)$$

Observing that $q_\omega(y^{1-t}\wedge y^t)\le y^{\omega/2}$ the first term $B_{i,1}^{(1)}(t)$ can be estimated by

$$\sup_{t\in[0,1]} B_{i,1}^{(1)}(t)\le \sup_{\mathbf{x}\in[0,1]^2}\frac{|\alpha_n(\mathbf{x})|}{q_\omega(x_1\wedge x_2)}\times\Big(1\vee \sup_{(y,t):\,C(y^{1-t},y^t)>n^{-\alpha}}\Big|\frac{C}{\tilde C_n}(y^{1-t},y^t)\Big|\Big)\times\psi(i),$$

where $\psi(i)=\int_0^{1/i}y^{-1+\omega/2}\,dy=o(1)$ for $i\to\infty$. Using analogous arguments as in the previous proof we can conclude, under consideration of (5.21) and Theorem G.1 in Genest and Segers (2009), that $\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*(\sup_{t\in[0,1]}B_{i,1}^{(1)}(t)>\varepsilon/4)=0$. For the second summand we note that

$$\sup_{t\in[0,1]} B_{i,2}^{(1)}(t)\le \sup_{x_1\in[0,1]}\frac{|\alpha_n(x_1,1)|}{q_\omega(x_1)}\times\Big(1\vee \sup_{(y,t):\,C(y^{1-t},y^t)>n^{-\alpha}}\Big|\frac{C}{\tilde C_n}(y^{1-t},y^t)\Big|\Big)\times\sup_{t\in[0,1]}\int_0^{1/i}K_2(y,t)\,dy,$$

where $K_2(y,t)$ is defined in (5.16). From (5.18), we have $\lim_{i\to\infty}\sup_{t\in[0,1]}\int_0^{1/i}K_2(y,t)\,dy=0$. Again, under consideration of (5.21) and Theorem G.1 in Genest and Segers (2009), we obtain $\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*\Big(\sup_{t\in[0,1]}B_{i,2}^{(1)}(t)>\varepsilon/4\Big)=0$. A similar argument works for $B_{i,3}^{(1)}$ and from the estimates for the different terms the assertion

$$\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*(\sup_{t\in[0,1]}|B_i^{(1)}(t)|>\varepsilon)=0$$

follows. Considering the term $\sup_{t\in[0,1]}|B_i^{(2)}(t)|$ we proceed along similar lines as in the proof of Theorem 5.1. For the sake of brevity we only state the important differences: in estimation (5.14) replace λ by β, then make use of decomposition (5.19), calculations similar to (5.18), and Theorem G.1 in Genest and Segers (2009) again and for the estimation of the remainder note that $\int_{1/(n+2)^2}^{1/i}y^{-\beta}=O(n^{2(\beta-1)})$. □

Proof of Proposition 3.4 and 3.5. The proof follows by similar lines as given in the proof of Proposition 3.3 in Genest and Segers (2009). We therefore only deal with the case $h_k(x)=x^{-2}e^{-kx}$ and note that the assertion for the weight function $h_\alpha(x)=x^{-\alpha}$ follows by similar arguments.
Observing that $\partial_1C(u^{1-t},u^t)=u^{A(t)+t-1}\mu(t)$ and $\partial_2C(u^{1-t},u^t)=u^{A(t)-t}\nu(t)$ we can decompose $\sigma(u,v;t)$ into

$$\sigma(u,v;t)=\sigma_0(u,v;t)+(uv)^{A(t)}\Big\{\sum_{l=1}^4\sigma_l(u,v;t)-\sum_{l=5}^8\sigma_l(u,v;t)\Big\},$$

where

$$\begin{aligned}\sigma_0(u,v;t)&=(u\wedge v)^{A(t)}-(uv)^{A(t)}\\ \sigma_1(u,v;t)&=(u^{t-1}\wedge v^{t-1}-1)\mu^2(t)\\ \sigma_2(u,v;t)&=(u^{-t}\wedge v^{-t}-1)\nu^2(t)\\ \sigma_3(u,v;t)&=(u^{t-1}v^{-t}C(u^{1-t},v^t)-1)\mu(t)\nu(t)\end{aligned}$$

$$
\begin{aligned}
\sigma_4(u,v;t) &= (u^{-t}v^{1-t}C(v^{1-t},u^t)-1)\mu(t)\nu(t)\\
\sigma_5(u,v;t) &= (u^{-A(t)}v^{t-1}C(u^{1-t}\wedge v^{1-t},u^t)-1)\mu(t)\\
\sigma_6(u,v;t) &= (u^{t-1}v^{-A(t)}C(u^{1-t}\wedge v^{1-t},v^t)-1)\mu(t)\\
\sigma_7(u,v;t) &= (u^{-A(t)}v^{-t}C(u^{1-t},u^t\wedge v^t)-1)\nu(t)\\
\sigma_8(u,v;t) &= (u^{-t}v^{-A(t)}C(v^{1-t},u^t\wedge v^t)-1)\nu(t)
\end{aligned}
$$

In view of (3.10) we need to evaluate $\int_0^1\int_0^1\sigma_0(u,v;t)(uv)^{k-A(t)}\,du\,dv$ and $\int_0^1\int_0^1\sigma_l(u,v;t)(uv)^k\,du\,dv$ for $l=1,\dots,8$. By symmetry, some of these integrals coincide, that is

$$
\int_0^1\int_0^1\sigma_l(u,v;t)(uv)^k\,du\,dv=\int_0^1\int_0^1\sigma_{l+1}(u,v;t)(uv)^k\,du\,dv\quad l=3,5,7.
$$

Considering the remaining integrals straightforward calculations yield

$$
\begin{aligned}
\int_0^1\int_0^1\sigma_0(u,v;t)(uv)^{k-A(t)}\,du\,dv &= \frac{2}{(k+1)(2k+2-A(t))}-\frac{1}{(k+1)^2},\\
\int_0^1\int_0^1\sigma_1(u,v;t)(uv)^{k}\,du\,dv &= \Big(\frac{2}{(k+1)(2k+1+t)}-\frac{1}{(k+1)^2}\Big)\mu^2(t),\\
\int_0^1\int_0^1\sigma_2(u,v;t)(uv)^{k}\,du\,dv &= \Big(\frac{2}{(k+1)(2k+2-t)}-\frac{1}{(k+1)^2}\Big)\nu^2(t).
\end{aligned}
$$

Regarding the integral with respect to σ_3 we need to evaluate

$$
H_1(t)=\int_0^1\int_0^1 u^{k+t-1}v^{k-t}C(u^{1-t},v^t)\,du\,dv=\frac{1}{t(1-t)}\int_0^1\int_0^1 C(x,y)x^{\frac{k+1}{1-t}-2}y^{\frac{k+1}{t}-2}\,dx\,dy,
$$

where we have used the substitution $u^{1-t}=x$ and $v^t=y$. Next substitute $x=w^{1-s}$ and $y=w^s$, then $w=xy\in(0,1]$ and $s=\frac{\log y}{\log xy}\in[0,1]$, while the Jacobian of the transformation is given by $-\log w$. One obtains

$$
H_1(t)=\frac{1}{t(1-t)}\int_0^1\Big(A(s)+(k+1)\Big(\frac{1-s}{1-t}+\frac{s}{t}\Big)-1\Big)^{-2}ds,
$$

where the last equality follows by integration by parts. In consequence,

$$
\begin{aligned}
&\int_0^1\int_0^1\sigma_3(u,v;t)(uv)^k\,du\,dv\\
&=\Big\{\frac{1}{t(1-t)}\int_0^1\Big(A(s)+(k+1)\Big(\frac{1-s}{1-t}+\frac{s}{t}\Big)-1\Big)^{-2}ds-\frac{1}{(k+1)^2}\Big\}\mu(t)\nu(t).
\end{aligned}
$$

Regarding the integral of σ_5 we decompose

$$
\begin{aligned}
&\int_0^1\int_0^1 u^{k-A(t)}v^{k+t-1}C(u^{1-t}\wedge v^{1-t},u^t)\,du\,dv\\
&=\int_0^1\int_0^v u^{k-A(t)}v^{k+t-1}C(u^{1-t},u^t)\,du\,dv+\int_0^1\int_v^1 u^{k-A(t)}v^{k+t-1}C(v^{1-t},u^t)\,du\,dv
\end{aligned}
$$

Straightforward calculations show that the first integral equals $\big((k+1)(2k+1+t)\big)^{-1}$. For the second integral we substitute $v^{1-t}=x$ and $u^t=y$ to obtain

$$\frac{1}{t(1-t)}\int_0^1\int_0^{y^{(1-t)/t}} y^{\frac{k+1-A(t)}{t}-1}x^{\frac{k+1}{1-t}-2}C(x,y)\,dx\,dy.$$

We proceed by the same transformation as for σ_3, namely $x=w^{1-s}$ and $y=w^s$. The inequality $x<y^{(1-t)/t}$ transforms to $t>s$ and in consequence the latter integral equals

$$\begin{aligned}
&-\frac{1}{t(1-t)}\int_0^t\int_0^1 w^{s\left(\frac{k+1-A(t)}{t}-1\right)+(1-s)\left(\frac{k+1}{1-t}-2\right)+A(s)}\log w\,dw\,ds\\
&=\frac{1}{t(1-t)}\int_0^t\left(A(s)+(k+t)\frac{1-s}{1-t}+(k+1-A(t))\,\frac{s}{t}\right)^{-2}ds,
\end{aligned}$$

where the last equality follows by integration by parts. Combining all terms for σ_5 we obtain

$$\begin{aligned}
\int_0^1\int_0^1 \sigma_5(u,v;t)(uv)^k\,du\,dv &= \mu(t)\bigg\{\frac{1}{(k+1)(2k+1+t)}\\
&\quad+\frac{1}{t(1-t)}\int_0^t\left(A(s)+(k+t)\frac{1-s}{1-t}+(k+1-A(t))\,\frac{s}{t}\right)^{-2}ds-\frac{1}{(k+1)^2}\bigg\}.
\end{aligned}$$

For the integrals with respect to σ_7 similar calculations yield

$$\begin{aligned}
\int_0^1\int_0^1 \sigma_7(u,v;t)(uv)^k\,du\,dv &= \nu(t)\bigg\{\frac{1}{(k+1)(2k+2-t)}\\
&\quad+\frac{1}{t(1-t)}\int_t^1\left(A(s)+(k+1-A(t))\,\frac{1-s}{1-t}+(k+1-t)\frac{s}{t}\right)^{-2}ds-\frac{1}{(k+1)^2}\bigg\}
\end{aligned}$$

and the conclusion finally follows by assembling all terms. □

Proof of Theorem 4.1. Since the integration mapping is continuous, it suffices to establish the weak convergence $X_n(t)\overset{w}{\rightsquigarrow}X(t)$ in $l^\infty[0,1]$ where we define

$$\begin{aligned}
X_n(t) &= \int_0^1 n\Big(\log\frac{\tilde C_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)^2 w(y)\,dy - nB_h(\hat A_{n,h}(t)-A^*(t))^2,\\
X(t) &= \int_0^1\Big(\frac{\mathbb{G}_C(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)^2 w(y)\,dy - B_h\mathbb{A}^2_{C,h}(t).
\end{aligned}$$

The proof of this assertion follows along the lines of the proof of Theorem 5.1. For $i\ge 2$ we recall the notation $w(y)=(-\log y)h(-\log y)$ and consider the following random functions in $l^\infty[0,1]$

$$\begin{aligned}
X_{i,n}(t) &= \int_{1/i}^1 n\Big(\log\frac{\tilde C_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)^2 w(y)\,dy - B_h^{-1}\Big(\int_{1/i}^1\sqrt{n}\Big(\log\frac{\tilde C_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)\frac{h^*(y)}{\log y}\,dy\Big)^2,\\
X_i(t) &= \int_{1/i}^1\Big(\frac{\mathbb{G}_C(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)^2 w(y)\,dy - B_h^{-1}\Big(\int_{1/i}^1\frac{\mathbb{G}_C(y^{1-t},y^t)}{C(y^{1-t},y^t)}\frac{h^*(y)}{\log y}\,dy\Big)^2.
\end{aligned}$$

By an application of Lemma 6.1 in Section 6, it suffices to show the conditions listed in (5.2). By arguments similar to those in the proof of Theorem 5.1 we obtain

$$\sqrt{n}\log\frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)} \overset{w}{\rightsquigarrow} \frac{\mathbb{G}_C(y^{1-t},y^t)}{C(y^{1-t},y^t)}$$

in $l^\infty([1/i,1]\times[0,1])$. Assertion (i) now follows immediately by the boundedness of the functions $w(y)$ and $h^*(y)(-\log y)^{-1}$ on $[1/i,1]$ [see conditions (4.1), (4.3) and (2.2)] and the continuous mapping theorem.
For the proof of assertion (ii) we simply note that $\mathbb{G}_C^2$ and $\mathbb{G}_C$ are bounded on $[0,1]^2$ and $K_1(y,t)=\frac{w(y)}{C^2(y^{1-t},y^t)}$ and $K_2(y,t)=\frac{h^*(y)}{C(y^{1-t},y^t)}$ are bounded uniformly with respect to $t\in[0,1]$ by the integrable functions $\bar{K}_1(y)=w(y)\,y^{-2}$ and $\bar{K}_2(y)=h^*(y)(-\log y)^{-1}\,y^{-1}$.

For the proof of assertion (iii) we fix some $\alpha\in(0,1/2)$ such that $\lambda\alpha>2\gamma$ and consider the decomposition

$$X_n(t)-X_{i,n}(t)=B_i^{(1)}(t)+B_i^{(2)}(t)+B_i^{(3)}(t), \tag{5.23}$$

where

$$B_i^{(1)}(t) = \int_{I_{B_i^{(1)}}(t)} n\Big(\log\frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)^2 w(y)\,dy, \tag{5.24}$$

$$B_i^{(2)}(t) = \int_{I_{B_i^{(2)}}(t)} n\Big(\log\frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)^2 w(y)\,dy, \tag{5.25}$$

$$B_i^{(3)}(t) = -B_h^{-1}I(t,1/i)(2I(t,1)-I(t,1/i)), \tag{5.26}$$

$I_{B_i^{(1)}}(t)$ and $I_{B_i^{(2)}}(t)$ are defined in (5.6) and

$$I(t,a) = \sqrt{n}\int_0^a \Big(\log\frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big)\frac{h^*(y)}{\log y}\,dy.$$

By the same arguments as in the proof of Theorem 5.1 we have for every $\varepsilon>0$

$$\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*(\sup_{t\in[0,1]}|I(t,1/i)|>\varepsilon)=0,$$

and $\sup_{t\in[0,1]}|I(t,1)|=O_{\mathbb{P}^*}(1)$, which yields $\lim_{i\to\infty}\limsup_{n\to\infty}\mathbb{P}^*(\sup_{t\in[0,1]}|B_i^{(3)}(t)|>\varepsilon)=0$. For $B_i^{(1)}(t)$ we obtain the estimate

$$\begin{aligned}\sup_{t\in[0,1]}|B_i^{(1)}(t)| &\le \sup_{t\in[0,1]}\int_{I_{B_i^{(1)}}(t)} n\Big|(\tilde{C}_n-C)(y^{1-t},y^t)\Big|^2\Big|1\vee\frac{C^2}{\tilde{C}_n^2}(y^{1-t},y^t)\Big|\,w(y)\,y^{-2}\,dy\\ &\le \sup_{\mathbf{x}\in[0,1]^2} n|\tilde{C}_n(\mathbf{x})-C(\mathbf{x})|^2\times\Big(1\vee\sup_{\mathbf{x}\in[0,1]^2:\,C(\mathbf{x})>n^{-\alpha}}\Big|\frac{C^2}{\tilde{C}_n^2}(\mathbf{x})\Big|\Big)\times\psi(i),\end{aligned}$$

where $\psi(i) := \int_0^{1/i} w(y) y^{-2} dy$, which can be handled by the same arguments as in the proof of Theorem 5.1. Finally, the term $B_i^{(2)}(t)$ can be estimated by

$$\sup_{t\in[0,1]} |B_i^{(2)}(t)| \le \sup_{t\in[0,1]} \int_{I_{B_i^{(2)}(t)}} n\Big|(\tilde{C}_n - C)(y^{1-t}, y^t)\Big|^2 \Big|1 \vee \frac{C^\lambda}{\tilde{C}_n^2}(y^{1-t}, y^t)\Big| \, w(y)\, y^{-\lambda}\, dy$$

$$\le \sup_{\mathbf{x}\in[0,1]^2} n|\tilde{C}_n(\mathbf{x}) - C(\mathbf{x})|^2 \times \Big(1 \vee \sup_{\mathbf{x}\in[0,1]^2 : C(\mathbf{x}) \le n^{-\alpha}} \Big|\frac{C^\lambda}{\tilde{C}_n^2}(\mathbf{x})\Big|\Big) \times \phi(i),$$

where $\phi(i) = \int_0^{1/i} w(y) y^{-\lambda}\, dy = o(1)$ for $i \to \infty$ by condition (4.2). Mimicking the arguments from the proof of Theorem 5.1 completes the proof. □

Proof of Theorem 4.2. With the notation $\bar{v}(y) := 2(\log y)^2 h(-\log y)$ it follows that $|v(y,t)| \le \bar{v}(y)$ and the assumptions on h yield the validity of (4.1)-(4.3) for $v(y,t)$. This allows for an application of Theorem 5.1 and together with the continuous mapping theorem we obtain $\sqrt{n} S_1 \overset{w}{\rightsquigarrow} Z_1$. Thus it remains to verify the negligibility of $S_2 + S_3$. For S_3 we note that by Theorem 3.1 and the continuous mapping theorem we have $S_3 = O_{\mathbb{P}^*}(1/n)$ and it remains to consider S_2. To this end we fix some $\alpha \in (0, 1/2)$ such that $(1 + (\lambda-1)/2)\alpha > \gamma$ and consider the decomposition

$$\begin{aligned}
&\int_0^\infty \Big(\bar{C}_n(s,t) - \bar{C}(s,t)\Big)^2 s e^{-s} h(s)\, ds \\
&= \int_0^1 \log^2 \frac{\tilde{C}_n(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} (-\log y) h(-\log(y))\, dy \\
&= \int_{I_{B_1^{(1)}(t)}} \log^2 \frac{\tilde{C}_n(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} (-\log y) h(-\log(y))\, dy + \int_{I_{B_1^{(2)}(t)}} \log^2 \frac{\tilde{C}_n(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} (-\log y) h(-\log(y))\, dy \\
&=: T_1(t,n) + T_2(t,n)
\end{aligned}$$

where the sets $I_{B_1^{(j)}(t)}, j = 1,2$ are defined in (5.6). On the set $I_{B_1^{(1)}(t)}$ we use the estimate

$$\log^2 \frac{\tilde{C}_n(y^{1-t}, y^t)}{C(y^{1-t}, y^t)} \le \frac{|\tilde{C}_n - C|^2}{(C^*)^2}(y^{1-t}, y^t) \le \frac{|\tilde{C}_n - C|^2}{C^*}(y^{1-t}, y^t) \frac{1}{n^{-\alpha}(1 \wedge \frac{\tilde{C}_n(y^{1-t}, y^t)}{C(y^{1-t}, y^t)})}$$

$$\le n^\alpha \frac{|\tilde{C}_n - C|^2}{C^*}(y^{1-t}, y^t)\Big(1 \vee \sup_{\mathbf{x}\in[0,1]^2 : C(\mathbf{x}) > n^{-\alpha}} \frac{C(\mathbf{x})}{\tilde{C}_n(\mathbf{x})}\Big)$$

where $|C^*(y,t) - C(y^{1-t}, y^t)| \le |\tilde{C}_n(y^{1-t}, y^t) - C(y^{1-t}, y^t)|$. By arguments similar to those used in the proof of Theorem 5.1, it is now easy to see that

$$\sqrt{n} \sup_t |T_1(t,n)| \le \sup_{\mathbf{x}\in[0,1]^2} n^{\alpha+1/2} |\tilde{C}_n(\mathbf{x}) - C(\mathbf{x})|^2 \times \Big(1 \vee \sup_{\mathbf{x}\in[0,1]^2 : C(\mathbf{x}) > n^{-\alpha}} \Big|\frac{C}{\tilde{C}_n}(\mathbf{x})\Big|\Big)^2 \times K = o_{\mathbb{P}^*}(1),$$

where $K := \int_0^1 (-\log y) h(-\log y)\, y^{-1}\, dy < \infty$ denotes a finite constant [see conditions (4.2) and (4.3)]. Now set $\beta := (\lambda - 1)/2 > 0$. From the estimate

$$C^*(y,t) \ge y^{1+\beta}\left(1 \wedge \frac{\tilde{C}_n}{C^{1+\beta}}(y^{1-t}, y^t)\right) = y^{-\beta} y^\lambda \left(1 \wedge \frac{\tilde{C}_n}{C^{1+\beta}}(y^{1-t}, y^t)\right)$$

we obtain by similar arguments as in the proof of the negligibility of $|B_i^{(2)}(t)|$ in the proof of Theorem 5.1 (note that on $I_{B_1^{(2)}(t)}$ we have $y \le C(y^{1-t}, y^t) \le n^{-\alpha}$)

$$\sup_{t\in[0,1]} |T_2(t,n)| \;\le\; \log(n)n^{-\beta\alpha} \sup_{\mathbf{x}\in[0,1]^2} \sqrt{n}|\tilde{C}_n(\mathbf{x}) - C(\mathbf{x})| \times \Big(1 \vee \sup_{\mathbf{x}\in[0,1]^2 : C(\mathbf{x})\le n^{-\alpha}} \Big|\frac{C^{1+\beta}}{\tilde{C}_n}(\mathbf{x})\Big|\Big) \times \tilde{K}$$

where $\tilde{K} := \gamma \int_0^1 (1-\log y)(-\log y)h(-\log y)y^{-\lambda}\,dy$ denotes a finite constant [see conditions (2.4) and (4.2)] and we used the estimate

$$\Big|\log \frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big|^2 \le (\gamma\log n - \log y)\Big|\log \frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big| \le \gamma\log(n)(1-\log y)\Big|\log \frac{\tilde{C}_n(y^{1-t},y^t)}{C(y^{1-t},y^t)}\Big|,$$

which holds for sufficiently large n. Finally, we observe that

$$\sup_{\mathbf{x}\in[0,1]^2 : C(\mathbf{x})\le n^{-\alpha}} \Big|\frac{C^{1+\beta}}{\tilde{C}_n}(\mathbf{x})\Big| \le \sup_{\mathbf{x}\in[0,1]^2 : C(\mathbf{x})\le n^{-\alpha}} \big|n^{\gamma}C^{1+\beta}(\mathbf{x})\big| \le n^{\gamma-(1+\beta)\alpha} = o(1).$$

Now the proof is complete. □

6 An auxiliary result

Lemma 6.1. *Let $X_n, X_{i,n} : \Omega \to \mathbb{D}$ for $i,n\in\mathbb{N}$ be arbitrary maps with values in the metric space $(\mathbb{D},d)$ and $X_i, X : \Omega \to \mathbb{D}$ be Borel-measurable. Suppose that*

(i) *For every $i\in\mathbb{N}$: $X_{i,n} \overset{w}{\rightsquigarrow} X_i$ for $n\to\infty$,*

(ii) *$X_i \overset{w}{\rightsquigarrow} X$ for $i\to\infty$*

(iii) *For every $\varepsilon>0$: $\lim_{i\to\infty}\limsup_{n\to\infty} \mathbb{P}^*(d(X_{i,n},X_n)>\varepsilon) = 0.$*

Then $X_n \overset{w}{\rightsquigarrow} X$ for $n\to\infty$.

Proof. Let $F \subset \mathbb{D}$ be closed and fix $\varepsilon > 0$. If $F^\varepsilon = \{x\in\mathbb{D} : d(x,F)\le\varepsilon)$ denotes the ε-enlargement of F we obtain

$$\mathbb{P}^*(X_n\in F) \le \mathbb{P}^*(X_{i,n}\in F^\varepsilon) + \mathbb{P}^*(d(X_{i,n},X_n)>\varepsilon).$$

By hypothesis (i) and the Portmanteau-Theorem [see Van der Vaart and Wellner (1996)]

$$\limsup_{n\to\infty} \mathbb{P}^*(X_n\in F) \le \mathbb{P}(X_i\in F^\varepsilon) + \limsup_{n\to\infty}\mathbb{P}^*(d(X_{i,n},X_n)>\varepsilon).$$

By conditions (ii) and (iii) $\limsup_{n\to\infty}\mathbb{P}^*(X_n\in F) \le P(X\in F^\varepsilon)$ and since $F^\varepsilon \downarrow F$ for $\varepsilon\downarrow 0$ and closed F the result follows by the Portmanteau-Theorem. □

References

Ben Ghorbal, N., Genest, C., and Nešlehová, J. (2009). On the Ghoudi, Khoudraji, and Rivest test for extreme-value dependence. *Canadian Journal of Statistics*, 37:534–552.

Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons Inc., New York.

Bücher, A. and Dette, H. (2009). A note on bootstrap approximations for the empirical copula process. *http://www.ruhr-uni-bochum.de/mathematik3/research/index.html*.

Capéraà, P., Fougères, A. L., and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84:567–577.

Cebrian, A., Denuit, M., and Lambert, P. (2003). Analysis of bivariate tail dependence using extreme values copulas: An application to the SOA medical large claims database. *Belgian Actuarial Journal*, 3:33–41.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2:339–365.

Deheuvels, P. (1984). Probabilistic aspects of multivariate extremes. In de Oliveira, J. T., editor, *Statistical Extremes and Applications*. Reidel, Dordrecht.

Deheuvels, P. (1991). On the limiting behavior of the pickands estimator for bivariate extreme-value distributions. *Statistics and Probability Letters*, 12:429–439.

Fermanian, J. D., Radulović, D., and Wegkamp, M. H. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10:847–860.

Fils-Villetard, A., Guillou, A., and Segers, J. (2008). Projection estimators of Pickands dependence functions. *Canad. J. Statist.*, 36(3):369–382.

Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *Annals of Statistics*, 37(5B):2990–3022.

Ghoudi, K., Khoudraji, A., and Rivest, L. (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Canadian Journal of Statistics*, 26:187–197.

Gumbel, É. J. (1960). Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris*, 9:171–173.

Hall, P. and Tajvidi, N. (2000). Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli*, 6:835–844.

Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *Journal of Multivariate Analysis*, 29:274–291.

Hüsler, J. and Reiss, R.-D. (1989). Maxima of normal random vectors: between independence and complete dependence. *Statist. Probab. Lett.*, 7(4):283–286.

Jiménez, J. R., Villa-Diharce, E., and Flores, M. (2001). Nonparametric estimation of the dependence function in bivariate extreme value distributions. *Journal of Multivariate Analysis*, 76:159–191.

Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statist. Probab. Lett.*, 9(1):75–81.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference.* Springer Series in Statistics, New York.

Pickands, J. (1981). Multivariate extreme value distributions (with a discussion). *Proceedings of the 43rd Session of the International Statistical Institute. Bull. Inst. Internat. Statist.*, 49:859–878,894–902.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Annals of Statistics*, 4:912–923.

Scaillet, O. (2005). A Kolmogorov-Smirnov type test for positive quadrant dependence. *Canadian Journal of Statistics*, 33:415–427.

Segers, J. (2007). Nonparametric inference for bivariate extreme-value copulas. In Ahsanullah, M. and Kirmani, S. N. U. A., editors, *Topics in Extreme Values.* Nova Science Publishers, New York.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Stute, W. (1984). The oscillation behavior of empirical processes: The multivariate case. *Annals of Probability*, 12:361–379.

Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika*, 75:397–415.

Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canadian Journal of Statistics*, 33:357–375.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer Verlag, New York.

Zhang, D., Wells, M. T., and Peng, L. (2008). Nonparametric estimation of the dependence function for a multivariate extreme value distribution. *Journal of Multivariate Analysis*, 99:577–588.