Multiplier bootstrap of tail copulas - with applications

Axel Bücher, Holger Dette
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

January 26, 2011

Abstract

In the problem of estimating the lower and upper tail copula we propose two bootstrap procedures for approximating the distribution of the corresponding empirical tail copulas. The first method uses a multiplier bootstrap of the empirical tail copula process and requires estimation of the partial derivatives of the tail copula. The second method avoids this estimation problem and uses multipliers in the two-dimensional empirical distribution function and in the estimates of the marginal distributions. For both multiplier bootstrap procedures we prove consistency.

For these investigations we demonstrate that the common assumption of the existence of continuous partial derivatives in the the literature on tail copula estimation is so restrictive, such that the tail copula corresponding to asymptotic independence is the only tail copula with this property. Moreover, we are able to solve this problem and prove weak convergence of the empirical tail copula process under nonrestrictive smoothness assumptions which are satisfied for many commonly used models. These results are applied in several statistical problems including minimum distance estimation and goodness-of-fit testing.

Keywords and Phrases: tail copula, stable tail dependence function, multiplier bootstrap, minimum distance estimation, comparison of tail copulas, goodness-of-fit
AMS Subject Classification: Primary 62G32 ; secondary 62G20

1 Introduction

The stable tail dependence function appears naturally in multivariate extreme value theory as a function that characterizes extremal dependence: if a bivariate distribution function F lies in the max-domain of attraction of an extreme-value distribution G, then the copula of G is completely determined by the stable tail dependence function [see e.g. Einmahl et al. (2008)]. The function is closely related to tail copulas [see Schmidt and Stadtmüller (2006)] and represents the current standard to describe extremal dependence [see Embrechts et al. (2003) and Malevergne and Sornette (2004)]. Following Schmidt and

Stadtmüller (2006) the lower and the upper tail copulas are defined by

$$\Lambda_L(\mathbf{x}) = \lim_{t\to\infty} t\, C(x_1/t, x_2/t) \tag{1.1}$$

$$\Lambda_U(\mathbf{x}) = \lim_{t\to\infty} t\, \bar{C}(x_1/t, x_2/t), \tag{1.2}$$

provided that the limits exist. Here $\mathbf{x} = (x_1, x_2) \in \bar{\mathbb{R}}_+^2 := [0,\infty]^2 \setminus \{(\infty,\infty)\}$, C denotes the copula of the two-dimensional distribution function F, which relates F and its marginals F_1, F_2 by

$$F(\mathbf{x}) = C(F_1(x_1), F_2(x_2)) \tag{1.3}$$

[see Sklar (1959)], and $\bar{C}(\mathbf{u}) = u_1 + u_2 - 1 + C(1-u_1, 1-u_2)$ denotes the survival copula. The stable tail dependence function l and the upper tail copula Λ_U are associated by the relationship

$$l(\mathbf{x}) = x_1 + x_2 - \Lambda_U(\mathbf{x}) \quad \forall\, \mathbf{x} \in \mathbb{R}_+^2.$$

Since its introduction various parametric and nonparametric estimates of the tail copulas and of the stable tail dependence function have been proposed in the literature. Several authors assume that the dependence function belongs to some parametric family. Coles and Tawn (1994), Tiago de Oliveira (1980) or Einmahl et al. (1993) imposed restrictions on the marginal distributions to estimate multivariate extreme value distributions. Nonparametric estimates of the stable tail dependence function have been investigated by Huang (1992), Qi (1997) and Drees and Huang (1998), while corresponding estimates for tail copulas have been discussed by Schmidt and Stadtmüller (2006). More recent work on inference on the stable tail dependence function can be found in Einmahl et al. (2008) and Einmahl et al. (2006), who investigated moment estimators of tail dependence and weighted approximations of tail copula processes, respectively. The present paper has two main purposes. First we clarify some curiosities in the literature on tail copula estimation, which stem from the fact that most authors assume the existence of continuous partial derivatives of the tail copula [see e.g. Huang (1992), Drees and Huang (1998), Schmidt and Stadtmüller (2006), Einmahl et al. (2006), de Haan and Ferreira (2006), Peng and Qi (2008) or de Haan et al. (2008) among others]. However, the tail copula corresponding to asymptotic independence is the only tail copula with this property, because the partial derivatives of a tail copula satisfy

$$\partial_1 \Lambda_L(0, x) = \begin{cases} \lim_{t\to\infty} \Lambda_L(1, t) & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases} \tag{1.4}$$

As a consequence we provide a result regarding the weak convergence of the empirical tail copula process (and thus also of the empirical stable tail dependence function) under weak smoothness assumptions (see Theorem 2.2 in the following section). The smoothness conditions are nonrestrictive in the sense, that in the case where they are not satisfied, the candidate limiting process does not have continuous trajectories. The second objective of the paper is devoted to the approximation of the distribution of estimators for the tail copulas by new bootstrap methods. In contrast to the problem of estimation of the stable dependence function and tail copulas, this problem has found much less attention in the literature. Recently, Peng and Qi (2008) considered the tail empirical distribution function and showed the consistency of the bootstrap based on resampling (again under the assumption of continuous partial derivatives). These results were used to construct confidence bands for the tail dependence function. While these authors considered the naive bootstrap, the present paper is devoted to multiplier bootstrap procedures for tail copula estimation. On the one hand, our research is motivated by the observation that the parametric bootstrap, which is commonly applied in goodness-of-fit testing problems [see de Haan et al. (2008)], has very high computational costs, because it heavily relies on random number generation and estimation [see also Kojadinovic and Yan (2010) and Kojadinovic et al. (2010) for a more detailed discussion of the computational efficiency of the multiplier bootstrap]. On the other hand, it was pointed out by Bücher and Dette (2010) in the context of nonparametric copula estimation that some multiplier bootstrap procedures lead to more reliable approximations than the bootstrap based on resampling.

In Section 2 we briefly review the nonparametric estimates of the tail copula and discuss their main properties. In particular we establish weak convergence of the empirical tail copula process under nonrestrictive smoothness assumptions, which are satisfied for many commonly used models. In Section 3 we introduce the multiplier bootstrap for the empirical tail copula and prove its consistency. In particular, we discuss two ways of approximating the distribution of the empirical tail copula by a multiplier bootstrap. Our first method is called *partial derivatives multiplier bootstrap* and uses the limit distribution of the empirical tail copula process. As a consequence, this approach requires the estimation of the partial derivatives of the tail copula. The second method, which avoids this problem, is called *direct multiplier bootstrap* and uses multipliers in the two-dimensional empirical distribution function and in the estimates of the marginal distributions. Finally, in Section 4 we discuss several statistical applications of the multiplier bootstrap. In particular, we investigate the problem of testing for equality between two tail copulas and we discuss the bootstrap approximations in the context of testing parametric assumptions for the tail copula. Finally, the proofs and some of the technical details are deferred to an appendix.

2 Empirical tail copulas

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ denote an i.i.d. sample of random variables distributed according to F and denote the empirical distribution functions of F, F_1 and F_2 by F_n, F_{n1} and F_{n2}, respectively. Following Schmidt and Stadtmüller (2006) we consider the estimators

$$\hat{\Lambda}_L(\mathbf{x}) = \frac{n}{k} C_n \left(\frac{kx_1}{n}, \frac{kx_2}{n} \right), \tag{2.1}$$

$$\hat{\Lambda}_U(\mathbf{x}) = \frac{n}{k} \bar{C}_n \left(\frac{kx_1}{n}, \frac{kx_2}{n} \right), \tag{2.2}$$

for the lower and upper tail copula, respectively, where $k \to \infty$ such that $k = o(n)$, and C_n (resp. $\bar{C}_n$) denotes the empirical copula (resp. empirical survival copula), that is

$$C_n(\mathbf{u}) = F_n(F_{n1}^{-}(u_1), F_{n2}^{-}(u_2))$$
$$\bar{C}_n(\mathbf{u}) = \bar{F}_n(\bar{F}_{n1}^{-}(u_1), \bar{F}_{n2}^{-}(u_2))$$

It is easy to see that these estimators are asymptotically equivalent to the estimates

$$\hat{\Lambda}_L(\mathbf{x}) \approx \frac{1}{k}\sum_{i=1}^{n} \mathbb{I}\{R(X_{i1}) \le kx_1, R(X_{i2}) \le kx_2\}, \tag{2.3}$$

$$\hat{\Lambda}_U(\mathbf{x}) \approx \frac{1}{k}\sum_{i=1}^{n} \mathbb{I}\{R(X_{i1}) > n - kx_1, R(X_{i2}) > n - kx_2\} \tag{2.4}$$

where $R(X_{ij}) = nF_{n1}(X_{j1})$ denotes the rank of X_{ij} among $X_{1j}, \dots X_{nj}$ $(j = 1, 2)$. Therefore we introduce analogs of (2.3) and (2.4) where the marginals F_1 and F_2 are assumed to be known, that is

$$\tilde{\Lambda}_L(\mathbf{x}) = \frac{1}{k}\sum_{i=1}^{n} \mathbb{I}\{F_1(X_{i1}) \le \frac{kx_1}{n}, F_2(X_{i2}) \le \frac{kx_2}{n}\}, \tag{2.5}$$

$$\tilde{\Lambda}_U(\mathbf{x}) = \frac{1}{k}\sum_{i=1}^{n} \mathbb{I}\{F_1(X_{i1}) > 1 - \frac{kx_1}{n}, F_2(X_{i2}) > 1 - \frac{kx_2}{n}\}. \tag{2.6}$$

For the sake of brevity we restrict our investigations to the case of lower tail copulas and we assume that this function is non-zero in a single point $\mathbf{x} \in \mathbb{R}^2_+$ [and as a consequence non-zero everywhere on $\mathbb{R}^2_+$, see Theorem 1 in Schmidt and Stadtmüller (2006)].
Let $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ denote the space of all functions $f : \bar{\mathbb{R}}^2_+ \to \mathbb{R}$, which are locally uniformly bounded on every compact subset of $\bar{\mathbb{R}}^2_+$, with metric

$$d(f_1, f_2) = \sum_{i=1}^{\infty} 2^{-i}(||f_1 - f_2||_{T_i} \wedge 1),$$

where the sets T_i are defined recursively by $T_{3i} = T_{3i-1} \cup [0, i]^2, T_{3i-1} = T_{3i-2} \cup ([0, i] \times \{\infty\}), T_{3i-2} = T_{3(i-1)} \cup (\{\infty\} \times [0, i])$ and $T_0 = \emptyset$. Note that with this metric the set $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ is a complete metric space and that a sequence f_n in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ converges with respect to d if and only if it converges uniformly on every T_i, see Van der Vaart and Wellner (1996). Throughout this paper $l^\infty(T)$ denotes the set of uniformly bounded functions on a set T and $\rightsquigarrow$ denotes weak convergence in the sense of Hoffmann-Jørgensen, see e.g. Van der Vaart and Wellner (1996).
Schmidt and Stadtmüller (2006) assumed that the lower tail copula Λ_L satisfies the second-order condition

$$\lim_{t\to\infty} \frac{\Lambda_L(\mathbf{x}) - tC(x_1/t, x_2/t)}{A(t)} = g(\mathbf{x}) \tag{2.7}$$

locally uniformly for $\mathbf{x} = (x_1, x_2) \in \bar{\mathbb{R}}^2_+$, where g is a non-constant function and the function $A : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\lim_{t\to\infty} A(t) = 0$. Under this and the additional assumptions $\Lambda_L \ne 0$, $\sqrt{k}A(n/k) \to 0$, $k = k(n) \to \infty$, $k = o(n)$, they showed that the lower tail copula process with known marginals defined in (2.5) converges weakly in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$, that is

$$\sqrt{k}\left(\tilde{\Lambda}_L(\mathbf{x}) - \Lambda_L(\mathbf{x})\right) \rightsquigarrow \mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x}), \tag{2.8}$$

where $\mathbb{G}_{\tilde{\Lambda}_L}$ is a centered Gaussian field with covariance structure given by

$$\mathbb{E}\,\mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x})\mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{y}) = \Lambda_L(x_1 \wedge y_1, x_2 \wedge y_2). \tag{2.9}$$

For the empirical tail copula $\hat{\Lambda}_L(\mathbf{x})$ they established the weak convergence

$$\alpha_n(\mathbf{x}) = \sqrt{k}\left(\hat{\Lambda}_L(\mathbf{x}) - \Lambda_L(\mathbf{x})\right) \rightsquigarrow \mathbb{G}_{\hat{\Lambda}_L}(\mathbf{x}) \tag{2.10}$$

in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$, provided that the tail copula has continuous partial derivatives. Here the limiting process $\mathbb{G}_{\hat{\Lambda}_L}$ has the representation

$$\mathbb{G}_{\hat{\Lambda}_L}(\mathbf{x}) = \mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x}) - \partial_1\Lambda_L(\mathbf{x})\mathbb{G}_{\tilde{\Lambda}_L}(x_1, \infty) - \partial_2\Lambda_L(\mathbf{x})\mathbb{G}_{\tilde{\Lambda}_L}(\infty, x_2). \tag{2.11}$$

The assumption of continuous partial derivatives is made in the whole literature on estimation of stable tail dependence functions and tail copulas. However, as demonstrated in (1.4) there does not exist any tail copula $\Lambda_L \neq 0$ with continuous partial derivatives at the origin $(0, 0)$. With our first result we will fill this gap and prove weak convergence of the empirical tail copula process under substantially weaker smoothness assumptions. For this purpose we will use a similar approach as in Schmidt and Stadtmüller (2006) since this turns out to be also useful for a proof of consistency of the multiplier bootstrap. First we consider the case of known marginals. Due to the second order condition (2.7) the proof of (2.8) can be given by showing weak convergence of the centered statistic

$$\tilde{\alpha}_n(\mathbf{x}) := \sqrt{k}\left(\tilde{\Lambda}_L(\mathbf{x}) - \frac{n}{k}C(x_1k/n, x_2k/n)\right). \tag{2.12}$$

Lemma 2.1. *If $\Lambda_L \neq 0$ and the second order condition (2.7) holds with $\sqrt{k}A(n/k) \to 0$, where $k = k(n) \to \infty$ and $k = o(n)$, then we have, as n tends to infinity*

$$\tilde{\alpha}_n(\mathbf{x}) = \sqrt{k}\left(\tilde{\Lambda}_L(\mathbf{x}) - \frac{n}{k}C(x_1k/n, x_2k/n)\right) \rightsquigarrow \mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x}) \tag{2.13}$$

in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$, where $\mathbb{G}_{\tilde{\Lambda}_L}$ is a tight centered Gaussian field concentrated on $\mathcal{C}_\rho(\bar{\mathbb{R}}^2_+)$ with covariance structure given in (2.9), where ρ is a pseudometric on the space $\bar{\mathbb{R}}^2_+$ defined by

$$\rho(\mathbf{x}, \mathbf{y}) = \mathbb{E}\left[(\mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x}) - \mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{y}))^2\right]^{1/2} = (\Lambda_L(\mathbf{x}) - 2\Lambda_L(\mathbf{x} \wedge \mathbf{y}) + \Lambda_L(\mathbf{x}))^{1/2},$$

$\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$, $\mathbf{x} \wedge \mathbf{y} = (x_1 \wedge y_1, x_2 \wedge y_2)$ and $\mathcal{C}_\rho(\bar{\mathbb{R}}^2_+) \subset \mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ denotes the subset of all functions that are uniformly ρ-continuous on every T_i.

This assertion is proved in Schmidt and Stadtmüller (2006) by showing convergence of the finite dimensional distributions and tightness. The proof of consistency of the bootstrap procedures proposed in the following section follows in part by arguments from an alternative proof of (2.13) based on Donsker classes which will be accomplished in the appendix.
For a proof of a corresponding result for the empirical tail copula process with estimated marginals in (2.10) we will use the functional delta method in (2.8) with some suitable functional.

Theorem 2.2. *Let $\Lambda_L \neq 0$ be a lower tail copula whose partial derivatives satisfy the following first order properties*

$$\partial_p \Lambda_L \text{ exists on } \{\mathbf{x} \in \bar{\mathbb{R}}^2_+ \,|\, x_p < \infty\} \text{ and is continuous on } \{\mathbf{x} \in \bar{\mathbb{R}}^2_+ \,|\, 0 < x_p < \infty\} \tag{2.14}$$

for $p = 1, 2$. *If additionally the assumptions of Lemma 2.1 are satisfied then we have*

$$\alpha_n(\mathbf{x}) = \sqrt{k}\left(\hat{\Lambda}_L(\mathbf{x}) - \Lambda_L(\mathbf{x})\right) \rightsquigarrow \mathbb{G}_{\hat{\Lambda}_L}(\mathbf{x})$$

in $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$, *where the process* $\mathbb{G}_{\hat{\Lambda}_L}$ *is defined in* (2.11) *and* $\partial_p \Lambda_L$, $p = 1, 2$ *is defined as 0 on the set* $\{\mathbf{x} \in \bar{\mathbb{R}}_+^2 \mid x_p = \infty\}$.

Theorem 2.2 has been proved by Schmidt and Stadtmüller (2006) under the additional assumption that the tail copula has continuous partial derivatives. As pointed out in the previous paragraphs there does not exist any tail copula $\Lambda_L \neq 0$ with this property.

3 Multiplier bootstrap approximation

3.1 Asymptotic theory

In this section we will construct multiplier bootstrap approximations of the Gaussian limit distributions $\mathbb{G}_{\tilde{\Lambda}_L}$ and $\mathbb{G}_{\hat{\Lambda}_L}$ specified in (2.8) and (2.10), respectively. To this end let ξ_i be i.i.d. positive random variables, independent of the $\mathbf{X}_i$, with mean μ in $(0, \infty)$ and finite variance τ^2, which additionally satisfy $||\xi||_{2,1} := \int_0^\infty \sqrt{P(|\xi| > x)}\, dx < \infty$. We will first deal with the case of known marginals and define a multiplier bootstrap analogue of (2.5) by

$$\tilde{\Lambda}_L^\xi(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n} \mathbb{I}\{F_1(X_{i1}) \leq \frac{kx_1}{n}, F_2(X_{i2}) \leq \frac{kx_2}{n}\} \tag{3.1}$$

where $\bar{\xi}_n = n^{-1} \sum_{i=1}^n \xi_i$ denotes the mean of $\xi_1, \ldots, \xi_n$. We have

$$\tilde{\alpha}_n^m(\mathbf{x}) = \frac{\mu}{\tau} \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\frac{\xi_i}{\bar{\xi}_n} - 1\right) f_{n,\mathbf{x}}(U_i) = \frac{\mu}{\tau} \sqrt{k}\big(\tilde{\Lambda}_L^\xi - \tilde{\Lambda}_L\big), \tag{3.2}$$

where the function $f_{n,\mathbf{x}}(U_i)$ is defined by

$$f_{n,\mathbf{x}}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}\left\{U_{i1} \leq kx_1/n, U_{i2} \leq kx_2/n\right\}, \tag{3.3}$$

and

$$\mathbf{U}_i = (U_{i1}, U_{i2}); \qquad U_{ij} = F_j(X_{ij}), \text{ for } j = 1, 2.$$

Throughout this paper we use the notation

$$G_n \overset{\mathbb{P}}{\underset{\xi}{\rightsquigarrow}} G \text{ in } \mathbb{D} \tag{3.4}$$

for *conditional weak convergence in a metric space* $(\mathbb{D}, d)$ in the sense of Kosorok (2008), page 19. To be precise, (3.4) holds for some random variables $G_n = G_n(\mathbf{X}_1, \ldots, \mathbf{X}_n, \xi_1, \ldots \xi_n), G \in \mathbb{D}$ if and only if

$$\sup_{h \in BL_1(\mathbb{D})} |\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)| \xrightarrow{\mathbb{P}^*} 0 \tag{3.5}$$

and

$$\mathbb{E}_\xi h(G_n)^* - \mathbb{E}_\xi h(G_n)_* \xrightarrow{\mathbb{P}^*} 0 \quad \text{for every } h \in BL_1(\mathbb{D}), \tag{3.6}$$

where

$$BL_1(\mathbb{D}) = \{f : \mathbb{D} \to \mathbb{R} \mid ||f||_\infty \leq 1, |f(\beta) - f(\gamma)| \leq d(\beta, \gamma) \;\forall\, \gamma, \beta \in \mathbb{D}\}$$

denotes the set of all Lipschitz-continuous functions bounded by 1. The subscript ξ in the expectations indicates conditional expectation over the weights $\xi = (\xi_1, \ldots, \xi_n)$ given the data and $h(G_n)^*$ and $h(G_n)_*$ denote measurable majorants and minorants with respect to the joint data, including the weights ξ. The condition (3.5) is motivated by the metrization of weak convergence by the bounded Lipschitz-metric, see e.g. Theorem 1.12.4 in Van der Vaart (1998). The following result shows that the process (3.2) provides a valid bootstrap approximation of the process (2.12).

Theorem 3.1. *If $\Lambda_L \neq 0$ and the second order condition* (2.7) *holds with $\sqrt{k}A(n/k) \to 0$, $k = k(n) \to \infty$ and $k = o(n)$ we have, as n tends to infinity,*

$$\tilde{\alpha}_n^m = \frac{\mu}{\tau}\sqrt{k}\big(\tilde{\Lambda}_L^\xi - \tilde{\Lambda}_L\big) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\tilde{\Lambda}_L}$$

in the metric space $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$.

Since Theorem 3.1 states that we have weak convergence of $\tilde{\alpha}_n^m$ to $\mathbb{G}_{\tilde{\Lambda}_L}$ conditional on the data U_i, it provides a bootstrap approximation of the empirical tail copula in the case where the marginal distributions are known. To be precise, consider $B \in \mathbb{N}$ independent replications of the random variables $\xi_1, \ldots, \xi_n$ and denote them by $\xi_{1,b}, \ldots, \xi_{n,b}$. Compute the statistics $\tilde{\alpha}_{n,b}^m = \tilde{\alpha}_n^m(\xi_{1,b}, \ldots, \xi_{n,b})$ $(b = 1, \ldots, B)$ and use the empirical distribution of $\tilde{\alpha}_{n,1}^m, \ldots, \tilde{\alpha}_{n,B}^m$ as an approximation for the limiting distribution of $G_{\tilde{\Lambda}_L}$.
Because in most cases of practical interest there will be no information about the marginals one cannot use Theorem 3.1 in many statistical applications. We will now develop two consistent bootstrap approximation for the limiting distribution of the process (2.10) which do not require knowledge of the marginals. Intuitively, it is natural to replace the unknown marginal distributions in (3.1) by its empirical counterparts, that is

$$\hat{\Lambda}_L^{\xi,\cdot}(\mathbf{x}) = \frac{1}{k}\sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n}\mathbb{I}\{X_{i1} \leq F_{n1}^{-1}(kx_1/n), X_{i2} \leq F_{n2}^{-1}(kx_2/n)\} \tag{3.7}$$

which yields the process

$$\beta_n(\mathbf{x}) = \frac{\mu}{\tau}\sqrt{k}\left(\hat{\Lambda}_L^{\xi,\cdot} - \hat{\Lambda}_L\right) = \frac{\mu}{\tau}\frac{1}{\sqrt{k}}\sum_{i=1}^n\left(\frac{\xi_i}{\bar{\xi}_n} - 1\right)\mathbb{I}\{X_{i1} \leq F_{n1}^{-1}(kx_1/n), X_{i2} \leq F_{n2}^{-1}(kx_2/n)\}.$$

Unfortunately, this intuitive approach does not yield an approximation for the distribution of the process $\mathbb{G}_{\hat{\Lambda}_L}$, but of $\mathbb{G}_{\tilde{\Lambda}_L}$.

Theorem 3.2. *Suppose that the assumptions of Theorem 2.2 hold. Then we have, as n tends to infinity*

$$\beta_n = \frac{\mu}{\tau}\sqrt{k}\big(\hat{\Lambda}_L^{\xi,\cdot} - \hat{\Lambda}_L\big) \overset{\mathbb{P}}{\underset{\xi}{\rightsquigarrow}} \mathbb{G}_{\tilde{\Lambda}_L}$$

in the metric space $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$.

Although Theorem 3.2 provides a negative result and shows, that the distribution of β_n can not be used for approximating the limiting law $\mathbb{G}_{\hat{\Lambda}_L}$, it turns out to be essential for our first consistent multiplier bootstrap method. To be precise, we note that the distribution of β_n can be calculated from the data without knowing the marginal distributions. As a consequence, we obtain an approximation for the unknown distribution of the process $\mathbb{G}_{\tilde{\Lambda}_L}$. In order to get an approximation of $\mathbb{G}_{\hat{\Lambda}_L}$ we follow Rémillard and Scaillet (2009) and estimate the derivatives of the tail copula by

$$\widehat{\partial_p\Lambda_L}(\mathbf{x}) := \begin{cases} \frac{\hat{\Lambda}_L(\mathbf{x}+h\mathbf{e}_p)-\hat{\Lambda}_L(\mathbf{x}-h\mathbf{e}_p)}{2h} & ,\infty > x_p \geq h \\ \widehat{\partial_p\Lambda_L}(\mathbf{x}+(h-x_p)\mathbf{e}_p) = \frac{\hat{\Lambda}_L(\mathbf{x}+2h\mathbf{e}_p)-\hat{\Lambda}_L(\mathbf{x}-x_p\mathbf{e}_p)}{2h} & ,x_p < h \\ 0 & ,x_p = \infty \end{cases}$$

where $h \sim k^{-1/2}$ tends to 0 with increasing sample size. We will show in the Appendix (see the proof of the following Theorem in Appendix A) that these estimates are consistent, and consequently we define the process

$$\alpha_n^{pdm}(\mathbf{x}) = \beta_n(\mathbf{x}) - \widehat{\partial_1\Lambda_L}(\mathbf{x})\beta_n(x_1,\infty) - \widehat{\partial_2\Lambda_L}(\mathbf{x})\beta_n(\infty,x_2). \tag{3.8}$$

Note that α_n^{pdm} only depends on the data and the multipliers $\xi_1,\ldots,\xi_n$. As a consequence, a bootstrap sample can easily be generated as described in the previous paragraph and we call this method *partial derivatives multiplier bootstrap (pdm-bootstrap)* in the following discussion. Our next result shows that the *pdm*-bootstrap provides a valid approximation for the distribution of the process $\mathbb{G}_{\hat{\Lambda}_L}$.

Theorem 3.3. *Under the assumptions of Theorem 2.2 we have*

$$\alpha_n^{pdm} \overset{\mathbb{P}}{\underset{\xi}{\rightsquigarrow}} \mathbb{G}_{\hat{\Lambda}_L}$$

in the metric space $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$.

It turns out that there is an alternative valid multiplier bootstrap procedure in the case of unknown marginal distributions, which is attractive because it avoids the problem of estimating the partial derivatives of the lower tail copula. This method not only introduces multiplier random variables in the two-dimensional distribution function but also in the inner estimators of the marginals. To be precise define

$$F_n^\xi(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n}\mathbb{I}\{X_{i1} \leq x_1, X_{i2} \leq x_2\}$$

$$F_{nj}^\xi(x_j) = \frac{1}{n}\sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n}\mathbb{I}\{X_{ij} \leq x_j\}, \quad j = 1,2$$

$$C_n^{\xi,\xi}(\mathbf{u}) = F_n^{\xi}(F_{n1}^{\xi-}(u_1), F_{n2}^{\xi-}(u_2)).$$

and consider the process

$$\hat{\Lambda}_L^{\xi,\xi}(\mathbf{x}) := \frac{n}{k} C_n^{\xi,\xi}\left(\frac{k}{n}\mathbf{x}\right) = \frac{1}{k}\sum_{i=1}^{n} \frac{\xi_i}{\bar{\xi}_n} \mathbb{I}\{X_{i1} \leq F_{n1}^{\xi-}(kx_1/n), X_{i2} \leq F_{n2}^{\xi-}(kx_2/n)\} \tag{3.9}$$

$$\approx \frac{1}{k}\sum_{i=1}^{n} \frac{\xi_i}{\bar{\xi}_n} \mathbb{I}\{F_{n1}^{\xi}(X_{i1}) \leq kx_1/n, F_{n2}^{\xi}(X_{i2}) \leq kx_2/n\} \tag{3.10}$$

Throughout this paper we will call this bootstrap method the *direct multiplier bootstrap (dm-bootstrap).*

Theorem 3.4. *Under the assumptions of Theorem 2.2 we have*

$$\alpha_n^{dm}(\mathbf{x}) = \frac{\mu}{\tau}\sqrt{k}\left(\hat{\Lambda}_L^{\xi,\xi}(\mathbf{x}) - \hat{\Lambda}_L(\mathbf{x})\right) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\hat{\Lambda}_L} \quad in\ \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)). \tag{3.11}$$

3.2 Finite sample results

In this section we will present a small comparison of the finite sample properties of the two bootstrap approximations given in this section. For the sake of brevity we only consider data generated form the Clayton copula with a coefficient of lower tail dependence $\lambda_L = 0.25$. The Clayton copula,

$$C(\mathbf{u}; \theta) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}, \quad \theta > 0, \tag{3.12}$$

is a widely used copula family for the modeling of negative tail dependent data. Its lower tail copula is given by

$$\Lambda_L(\mathbf{x}) = \left(x_1^{-\theta} + x_2^{-\theta}\right)^{-1/\theta}.$$

In Tables 1 - 3 we investigate the accuracy of the bootstrap approximation of the covariances of the limiting variable $\mathbb{G}_{\hat{\Lambda}_L}$. More precisely, we chose three points on the unit circle $\{e^{i\varphi}, \varphi = k\pi/8 \text{ with } k = 1, 2, 3\}$ and show in the first four columns of Table 1 the true covariances of the limiting process $\mathbb{G}_{\hat{\Lambda}_L}$. The remaining columns show the simulated covariances of the process α_n on the basis of $5 \cdot 10^5$ simulation runs, where the sample size is $n = 1000$ and the parameter k is chosen as 50. This table is the benchmark for the bootstrap approximations of the covariances stated in Table 2. For the sake of completeness we also investigated the resampling bootstrap considered in Peng and Qi (2008) (which is hereafter denoted by α_n^{res}). The covariances are based on the average of 1000 simulation runs, where in each run the covariance is estimated on the basis of $B = 500$ bootstrap replications. In Table 3 we present the corresponding mean squared error.

As one can see all bootstrap procedures yield approximations of quite comparable magnitude. Considering only the bias in Table 2 the *pdm*-bootstrap has slight advantages in all cases, while there are basically no differences between the *dm*- and the resampling bootstrap. A comparison of the mean squared error in Table 3 shows that the *pdm*-bootstrap has the best performance on the diagonal. On the other hand, it yields a less accurate approximation in case of approximating off-diagonal covariances. In this case, the *dm*-bootstrap yields the best results.

True	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	α_n	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$
$\frac{\pi}{8}$	0.0874	0.0754	0.0516	$\frac{\pi}{8}$	0.0889	0.0737	0.0476
$2\frac{\pi}{8}$		0.1160	0.0754	$2\frac{\pi}{8}$		0.1218	0.0741
$3\frac{\pi}{8}$			0.0874	$3\frac{\pi}{8}$			0.0892

Table 1: *Left part: True covariances of* $\mathbb{G}_{\hat{\Lambda}_L}$ *for the Clayton Copula with* $\lambda_L = 0.25$*. Right part: sample covariances of the empirical tail copula process* α_n *with sample size* $n = 1000$ *and parameter* $k = 50$*.*

α_n^{pdm}	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	α_n^{dm}	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	α_n^{res}	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$
$\frac{\pi}{8}$	0.0948	0.0729	0.0468	$\frac{\pi}{8}$	0.1001	0.0714	0.0450	$\frac{\pi}{8}$	0.1004	0.0701	0.0439
$2\frac{\pi}{8}$		0.1300	0.0724	$2\frac{\pi}{8}$		0.1369	0.7073	$2\frac{\pi}{8}$		0.1363	0.0702
$3\frac{\pi}{8}$			0.0944	$3\frac{\pi}{8}$			0.0999	$3\frac{\pi}{8}$			0.0998

Table 2: *Averaged sample covariances of the Bootstrap approximations* α_n^{pdm}*,* α_n^{dm} *and* α_n^{res} *of* $\mathbb{G}_{\hat{\Lambda}_L}$ *under the conditions of Table 1.*

α_n^{pdm}	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	α_n^{dm}	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	α_n^{res}	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$
$\frac{\pi}{8}$	3.6767	4.6885	3.6509	$\frac{\pi}{8}$	3.8699	3.4967	2.7241	$\frac{\pi}{8}$	4.2179	3.8511	3.2198
$2\frac{\pi}{8}$		8.1104	4.8774	$2\frac{\pi}{8}$		8.8928	3.2598	$2\frac{\pi}{8}$		8.7310	3.6403
$3\frac{\pi}{8}$			3.7062	$3\frac{\pi}{8}$			3.7777	$3\frac{\pi}{8}$			3.9002

Table 3: *Mean squared error* $\times 10^4$ *of the different estimates for the covariances in Table 2.*

4 Statistical applications

In this section we investigate several statistical applications of the multiplier bootstrap. In particular we discuss the problem of comparing lower tail copulas from different samples, the problem of constructing confidence intervals and the problem of testing for a parametric form of the lower tail copula.

4.1 Testing for equality between two tail copulas

Let $\mathbf{X}_1, \ldots, \mathbf{X}_{n_1}$ and $\mathbf{Y}_1, \ldots, \mathbf{Y}_{n_2}$ denote two independent samples of i.i.d. random variables [we will relax the assumption of independence between the samples later on] with cumulative distribution function $F = C(F_1, F_2)$ and $H = D(H_1, H_2)$, respectively. We assume that the marginal distributions F_1, F_2 and H_1, H_2 of F and H are continuous and that for both distributions the corresponding lower tail copulas, say $\Lambda_{L,X}$ and $\Lambda_{L,Y}$, exist and do not vanish. We are interested in a test for the hypothesis

$$\mathcal{H}_0 : \Lambda_{L,X} \equiv \Lambda_{L,Y} \quad \text{vs.} \quad \mathcal{H}_1 : \Lambda_{L,X} \neq \Lambda_{L,Y} \tag{4.1}$$

Due to the homogeneity of tail copulas we have $\Lambda_L(t\mathbf{x}) = t\Lambda_L(\mathbf{x}) \quad \forall t > 0, \mathbf{x} \in \mathbb{R}^2_+$, and the hypotheses are equivalent to

$$\mathcal{H}_0 : \varrho(\Lambda_{L,X}, \Lambda_{L,Y}) = 0 \quad \text{vs} \quad \mathcal{H}_1 : \varrho(\Lambda_{L,X}, \Lambda_{L,Y}) > 0,$$

where the distance ϱ is defined by

$$\begin{aligned}\varrho(\Lambda_{L,X}, \Lambda_{L,Y}) &:= \int_0^{\pi/2} \big(\Lambda_{L,X}(\cos\varphi, \sin\varphi) - \Lambda_{L,Y}(\cos\varphi, \sin\varphi)\big)^2 \, d\varphi \qquad (4.2)\\ &= \int_0^{\pi/2} \big(\Lambda^{\angle}_{L,X}(\varphi) - \Lambda^{\angle}_{L,Y}(\varphi)\big)^2 \, d\varphi\end{aligned}$$

and we have used the notation

$$\Lambda^{\angle}_{L,X}(\varphi) = \Lambda_{L,X}(\cos\varphi, \sin\varphi), \;\; \Lambda^{\angle}_{L,Y}(\varphi) = \Lambda_{L,Y}(\cos\varphi, \sin\varphi).$$

We propose to base the test for the hypothesis (4.1) on the distance between the empirical tail copulas and define

$$\mathcal{S}_n = \frac{k_1 k_2}{k_1 + k_2} \varrho(\hat{\Lambda}_{L,X}, \hat{\Lambda}_{L,Y}) = \frac{k_1 k_2}{k_1 + k_2} \int_0^{\pi/2} (\hat{\Lambda}^{\angle}_{L,X}(\varphi) - \hat{\Lambda}^{\angle}_{L,Y}(\varphi))^2 \, d\varphi,$$

where $\hat{\Lambda}^{\angle}_{L,X}(\varphi) = \hat{\Lambda}_{L,X}(\cos(\varphi), \sin(\varphi))$, $\hat{\Lambda}^{\angle}_{L,Y} = \hat{\Lambda}_{L,Y}(\cos(\varphi), \sin(\varphi))$ denote the empirical tail copulas $\hat{\Lambda}_{L,X}$ and $\hat{\Lambda}_{L,Y}$ with corresponding parameters k_1 and k_2, satisfying

$$k_i = o(n_i), k_i/\log(n_i) \to \infty \;\; (i = 1, 2) \quad \text{and} \quad k_1/(k_1 + k_2) \to \lambda \in (0, 1).$$

Note that $\mathcal{S}_n$ can easily be computed from the ranks of X_{il} and Y_{il} in their respective samples, $X_{1l}, \ldots, X_{n_1 l}$ and $Y_{1l}, \ldots, Y_{n_2 l}$ that is

$$\mathcal{S}_n = \frac{k_1 k_2}{k_1 + k_2} \Big\{ \frac{1}{k_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} [(\arccos(S_{i1}) \wedge \arccos(S_{j1})) - (\arcsin(S_{i2}) \vee \arcsin(S_{j2}))]^+$$

$$- \frac{2}{k_1 k_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} [(\arccos(S_{i1}) \wedge \arccos(T_{j1})) - (\arcsin(S_{i2}) \vee \arcsin(T_{j2}))]^+$$
$$+ \frac{1}{k_2^2} \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} [(\arccos(T_{i1}) \wedge \arccos(T_{j1})) - (\arcsin(T_{i2}) \vee \arcsin(T_{j2}))]^+ .$$

Here we used the notation $S_{il} = \frac{R(X_{il})}{k_1} \wedge 1$, $T_{il} = \frac{R(Y_{il})}{k_1} \wedge 1$ $(l = 1, 2)$ and $[f]^+$ denotes the positive part of the function f.

Under the null hypothesis (4.1) of equality between the tail copulas we have $\mathcal{S}_n = \mathcal{T}_n$ with

$$\mathcal{T}_n = \int_0^{\pi/2} \mathcal{E}_n^2(\cos\varphi, \sin\varphi)\, d\varphi,$$

where

$$\mathcal{E}_n(\mathbf{x}) = \sqrt{\frac{k_2}{k_1 + k_2}} \sqrt{k_1} (\hat{\Lambda}_{L,X}(\mathbf{x}) - \Lambda_{L,X}(\mathbf{x})) - \sqrt{\frac{k_1}{k_1 + k_2}} \sqrt{k_2} (\hat{\Lambda}_{L,Y}(\mathbf{x}) - \Lambda_{L,Y}(\mathbf{x})).$$

Since the two samples X and Y are independent we obtain independently of the hypotheses that

$$\mathcal{E}_n \rightsquigarrow \sqrt{1-\lambda}\, \mathbb{G}_{\hat{\Lambda}_{L,X}} - \sqrt{\lambda}\, \mathbb{G}_{\hat{\Lambda}_{L,Y}} =: \mathcal{E}.$$

in the metric space $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$, where the two-dimensional centered Gaussian fields $\mathbb{G}_{\hat{\Lambda}_{L,X}}$ and $\mathbb{G}_{\hat{\Lambda}_{L,Y}}$ are defined in (2.11). This yields by the continuous mapping theorem

$$\mathcal{T}_n \rightsquigarrow \int_0^{\pi/2} \mathcal{E}^2(\cos\varphi, \sin\varphi)\, d\varphi =: \mathcal{T}$$

under both the null hypothesis and the alternative. Note that $\varrho(\hat{\Lambda}_{L,X}, \hat{\Lambda}_{L,Y}) \xrightarrow{\mathbb{P}} \varrho(\Lambda_{L,X}, \Lambda_{L,Y})$, which vanishes if and only if the null hypothesis (4.1) is satisfied. Therefore we can conclude that

$$\mathcal{S}_n \rightsquigarrow_{\mathcal{H}_0} \mathcal{T}, \qquad \mathcal{S}_n \xrightarrow{\mathbb{P}}_{\mathcal{H}_1} \infty, \tag{4.3}$$

which shows that a test, which rejects the null hypothesis (4.1) for large values of $\mathcal{T}_n$ is consistent.

In order to determine critical values for the test we approximate the limiting distribution $\mathcal{T}$ by the multiplier bootstrap proposed in Section 3. For this purpose we exemplarily consider the *pdm*-bootstrap (the extension to the *dm*-bootstrap is straightforward) using the definition in equation (3.11) and denote for any $b \in \{1, \ldots, B\}$ by $\xi_{1,b}, \ldots, \xi_{n_1,b}, \zeta_{1,b}, \ldots, \zeta_{n_2,b}$ independent and identically distributed non-negative random variables with mean μ_1 (resp. μ_2) and variance τ_1^2 (resp. τ_2^2). We compute for each b and both samples the bootstrap statistics as given in (3.8) or (3.11), i.e.

$$\alpha_{X,n_1,b}^{pdm}(\mathbf{x}) = \beta_{X,n_1,b}(\mathbf{x}) - \widehat{\partial_1 \Lambda_{L,X}}(\mathbf{x}) \beta_{X,n_1,b}(x_1, \infty) - \widehat{\partial_2 \Lambda_{L,X}}(\mathbf{x}) \beta_{X,n_1,b}(\infty, x_2),$$
$$\alpha_{Y,n_2,b}^{pdm}(\mathbf{x}) = \beta_{Y,n_2,b}(\mathbf{x}) - \widehat{\partial_1 \Lambda_{L,Y}}(\mathbf{x}) \beta_{Y,n_2,b}(x_1, \infty) - \widehat{\partial_2 \Lambda_{L,Y}}(\mathbf{x}) \beta_{Y,n_2,b}(\infty, x_2),$$

where

$$\beta_{X,n_1,b}(\mathbf{x}) = \frac{\mu_1}{\tau_1} \frac{1}{\sqrt{k_1}} \sum_{i=1}^{n_1} \left(\frac{\xi_{i,b}}{\bar{\xi}_{\cdot,b_{n_1}}} - 1 \right) \mathbb{I}\{F_{n_1 1}(X_{i1}) \le k_1 x_1/n_1, F_{n_1 2}(X_{i2}) \le k_1 x_2/n_1\},$$

$$\beta_{Y,n_2,b}(\mathbf{x}) = \frac{\mu_2}{\tau_2} \frac{1}{\sqrt{k_2}} \sum_{i=1}^{n_2} \left(\frac{\zeta_{i,b}}{\bar{\zeta}_{\cdot,b_{n_2}}} - 1 \right) \mathbb{I}\{H_{n_2 1}(Y_{i1}) \le k_2 x_1/n_2, H_{n_2 2}(Y_{i2}) \le k_2 x_2/n_2\},$$

and $\widehat{\partial_j \Lambda_{L,X}}$ and $\widehat{\partial_j \Lambda_{L,Y}}$ are the corresponding estimates of the partial derivatives ($j = 1, 2$, $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$). For all $\mathbf{x} \in \bar{\mathbb{R}}_+^2$ and all $b \in \{1, \dots, B\}$ define

$$\hat{\mathcal{E}}_n^{(pdm,b)}(\mathbf{x}) := \sqrt{\frac{k_2}{k_1 + k_2}} \alpha_{X,n_1,b}^{pdm}(\mathbf{x}) - \sqrt{\frac{k_1}{k_1 + k_2}} \alpha_{Y,n_2,b}^{pdm}(\mathbf{x}),$$
$$\hat{\mathcal{T}}_n^{(pdm,b)} := \int_0^{\pi/2} \left\{ \hat{\mathcal{E}}_n^{(pdm,b)}(\cos\varphi, \sin\varphi) \right\}^2 d\varphi,$$

By Theorem 3.3 and Theorem 10.8 in Kosorok (2008), it follows for every $b \in \{1, \dots, B\}$

$$\hat{\mathcal{T}}_n^{(pdm,b)} \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathcal{T}^{(b)},$$

where $\mathcal{T}^{(b)}$ is an independent copy of $\mathcal{T}$ (note that we consider the processes $\hat{\mathcal{E}}_n^{(\beta,b)}$ and $\hat{\mathcal{E}}_n^{(\gamma,b)}$ in the Banachspace $l^\infty([0,1]^2)$). Similarly, we have $\hat{\mathcal{T}}_n^{(dm,b)} \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathcal{T}^{(b)}$. From (4.3) we therefore obtain a consistent asymptotic level α test for the null hypothesis (4.1) by rejecting $\mathcal{H}_0$ for large values of $\mathcal{S}_n$, that is

$$\mathcal{S}_n > q_{1-\alpha}^m \tag{4.4}$$

where $q_{1-\alpha}^m$ denotes the $(1-\alpha)$ quantile of the empirical distribution function

$$K_n^m(s) = \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\{\hat{\mathcal{T}}_n^{(m,b)} \le s\}$$

(here m is either pdm or dm corresponding to partial derivative or direct multiplier bootstrap).
The discussion up till now holds true for two independent populations $\mathbf{X}_i$ and $\mathbf{Y}_i$. Nevertheless it is easy to check that the methodology of the previous sections also applies if we are faced with paired observations, i.e. $\mathbf{X}_i$ is not independent of $\mathbf{Y}_i$, but $n_1 = n_2 = n$. In that case we have to set $\zeta_{i,b} = \xi_{i,b}$ for all $i = 1, \dots, n$ and $b = 1, \dots, B$. To see this, set $\mathbf{Z}_i = (\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{Y}_{i1}, \mathbf{Y}_{i2})$ and denote the (empirical) copula of $\mathbf{Z}_i$ by $(\mathcal{C}_n)$ $\mathcal{C}$. Clearly,

$$C(u_1, u_2) = \mathcal{C}(u_1, u_2, 1, 1), \qquad D(v_1, v_2) = \mathcal{C}(1, 1, v_1, v_2),$$
$$C_n(u_1, u_2) = \mathcal{C}_n(u_1, u_2, 1, 1), \qquad D_n(v_1, v_2) = \mathcal{C}_n(1, 1, v_1, v_2).$$

If we set

$$\Lambda_{L,Z}(\mathbf{x}, \mathbf{y}) = \lim_{t\to\infty} t\,\mathcal{C}(\mathbf{x}/t, \mathbf{y}/t),$$
$$\hat{\Lambda}_{L,Z}(\mathbf{x}, \mathbf{y}) = \frac{n}{k} \mathcal{C}_n(\frac{n\mathbf{x}}{k}, \frac{n\mathbf{y}}{k}),$$

we obtain

$$\Lambda_{L,X}(\mathbf{x}) = \Lambda_{L,Z}(\mathbf{x}, \infty, \infty), \quad \Lambda_{L,Y}(\mathbf{y}) = \Lambda_{L,Z}(\infty, \infty, \mathbf{y}),$$

$\lambda_{L,X}$	$\lambda_{L,Y}$	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$
0.25	0.25	0.143	0.098	0.054
0.5	0.5	0.140	0.099	0.047
0.75	0.75	0.117	0.078	0.029
0.25	0.5	0.764	0.706	0.605
0.5	0.75	0.896	0.856	0.783
0.25	0.75	1	1	1

Table 4: *Simulated rejection probabilities of the pdm bootstrap test* (4.4) *for the hypothesis* (4.1) .

$\lambda_{L,X}$	$\lambda_{L,Y}$	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$
0.25	0.25	0.125	0.091	0.052
0.5	0.5	0.108	0.069	0.036
0.75	0.75	0.068	0.051	0.023
0.25	0.5	0.713	0.643	0.529
0.5	0.75	0.869	0.822	0.713
0.25	0.75	0.999	0.999	0.997

Table 5: *Simulated rejection probabilities of the dm bootstrap test* (4.4) *for the hypothesis* (4.1)

$$\hat{\Lambda}_{L,X}(\mathbf{x}) = \hat{\Lambda}_{L,Z}(\mathbf{x}, \infty, \infty), \quad \hat{\Lambda}_{L,Y}(\mathbf{y}) = \hat{\Lambda}_{L,Z}(\infty, \infty, \mathbf{y}).$$

Similar relations for the multiplier approximations are straightforward and the result follows along similar lines as in the previous sections.

For an investigation of the finite sample property we consider two independent samples of i.i.d. distributed random variables according to the Clayton copula, see (3.12), with a coefficient of lower tail dependence λ_L varying in the set $\{0.25, 0.5, 0.75\}$.

In Table 4 and 5 we show the simulated rejection probabilities of the *pdm* and *dm* bootstrap test defined in (4.4) for various nominal levels on the basis of 1000 simulation runs. The sample size is $n = 1000$, $k = 50$ and $B = 500$ bootstrap replications with $Laplacian(0, 2)$ multipliers have been used.

We observe that the nominal level is well approximated by the *pdm* bootstrap if the coefficient of tail dependence is not too large. For a larger coefficient the test is conservative. On the other hand, the *dm* bootstrap test is slightly more conservative and this effect is increasing with the coefficient of tail dependence. The alternative of different lower tail copulas is detected with reasonable power where both tests yield rather similar results with slight advantages for the *pdm*-bootstrap.

4.2 Bootstrap approximation of a minimum distance estimate and a computationally efficient goodness-of-fit test

In this section we are interested in estimating the tail copula of $\mathbf{X}$ under the additional assumption that the tail copula lies in some parametric class, say

$$\mathcal{L} = \{\Lambda_L(\cdot;\theta) \mid \theta \in \Theta\}. \tag{4.5}$$

Recently, estimates for parametric classes of tail copulas and stable tail dependence functions have been investigated by de Haan et al. (2008) and Einmahl et al. (2008) who proposed a censored likelihood and a moment based estimator, respectively. In the present section we investigate a further estimate, which is based on the minimum distance method. To be precise let Λ_L denote an arbitrary lower tail copula and $\Lambda_L(\,\cdot\,;\theta)$ an element in the parametric class $\mathcal{L}$ and define

$$\Lambda_L^{\angle}(\varphi) = \Lambda_L(\cos\varphi, \sin\varphi),\;\; \Lambda_L^{\angle}(\varphi;\theta) = \Lambda_L(\cos\varphi, \sin\varphi;\theta).$$

We consider the parameter corresponding to the best approximation by the distance ϱ defined in (4.2)

$$\theta_B = T(\Lambda_L) = \arg\min_{\theta\in\Theta} \varrho(\Lambda_L, \Lambda_L(\,\cdot\,;\theta)), \tag{4.6}$$

and call $\hat{\theta}_n^{MD} = T(\hat{\Lambda}_L)$ a minimum distance estimator for θ, where $\hat{\Lambda}_L$ is the empirical lower tail copula defined in (2.1) and $\hat{\Lambda}_L^{\angle} = \hat{\Lambda}_L(\cos\varphi, \sin\varphi)$. Note that θ_B is the "true" parameter if the null hypothesis is satisfied.

Throughout this subsection let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ denote i.i.d. bivariate random variables with cumulative distribution function $F = C(F_1, F_2)$ and existing lower tail copula Λ_L. Furthermore, we introduce the notations

$$\begin{aligned} Q(\theta) &= \varrho(\Lambda_L, \Lambda_L(\cdot;\theta)), \qquad \psi(\theta) = \partial_\theta\, Q(\theta), \\ Q_n(\theta) &= \varrho(\hat{\Lambda}_L, \Lambda_L(\cdot;\theta)), \qquad \psi_n(\theta) = \partial_\theta\, Q_n(\theta), \end{aligned}$$

and assume that the following regularity conditions are satisfied.

(B.1) The parameter space Θ has non-empty interior, say Θ^0, and the parameter $\theta_B = T(\Lambda_L) \in \Theta^0$ corresponding to the best approximation of the lower tail copula by the parametric class $\mathcal{L}$ exists and is unique.

(B.2) $\Lambda_L(\cdot;\theta)$ is continuously differentiable with respect to $\theta \in \Theta^0$ with $\delta_\theta(\mathbf{x}) = \partial_\theta\, \Lambda_L(\mathbf{x};\theta)$ and the mapping $\mathbf{x} \mapsto \sup_{\theta\in\Theta} ||\delta_\theta(\mathbf{x})||$ is integrable on $K_+ = \{(\cos\varphi, \sin\varphi) : \varphi \in [0, \pi/2]\}$.

(B.3) For every $\varepsilon > 0$

$$\inf_{\theta:||\theta-\theta_B||\geq\varepsilon} ||\psi(\theta)|| > 0 = ||\psi(\theta_B)||.$$

(B.4) $\partial_\theta\, \delta_\theta(\mathbf{x})$ exists for every $\mathbf{x} \in K_+$ and is continuous in θ_B.

(B.5) The matrix

$$A_{\theta_B} := \int \delta^{\angle}_{\theta_B}(\varphi)\delta^{\angle}_{\theta_B}(\varphi)^T + \partial_\theta \, \delta^{\angle}_{\theta_B}(\varphi)(\Lambda^{\angle}_L(\varphi;\theta_B) - \Lambda^{\angle}_L(\varphi))\, d\varphi$$

exists and is non-singular, where $\delta^{\angle}_{\theta}(\varphi) = \delta_\theta(\cos\varphi, \sin\varphi)$.

Theorem 4.1. *If the assumptions* (B.1) - (B.3) *hold and the true tail copula* Λ_L *satisfies the first order condition* (2.14) *of Theorem 2.2, then the minimum distance estimator* $\hat\theta_n^{MD}$ *is consistent for the parameter* θ_B *corresponding to the best approximation with respect to the distance* ϱ*. If additionally the conditions* (B.4) *and* (B.5) *also hold, then the minimum distance estimator* $\hat\theta_n^{MD}$ *is asymptotically normally distributed, more precisely*

$$\begin{aligned}\Theta_n^{MD} &:= \sqrt{k}(\hat\theta_n^{MD} - \theta_B) = \sqrt{k}\int \gamma_{\theta_B}(\varphi)\left(\hat\Lambda^{\angle}_L(\varphi) - \Lambda^{\angle}_L(\varphi)\right) d\varphi + o_{\mathbb{P}^*}(1)\\ &\rightsquigarrow \int \gamma_{\theta_B}(\varphi)\mathbb{G}^{\angle}_{\hat\Lambda_L}(\varphi)\, d\varphi =: \Theta^{MD},\end{aligned}$$

where $\gamma_{\theta_B}(\varphi) = A^{-1}_{\theta_B}\delta^{\angle}_{\theta_B}(\varphi)$ *and* $\mathbb{G}^{\angle}_{\hat\Lambda_L}(\varphi) = \mathbb{G}_{\hat\Lambda_L}(\cos\varphi, \sin\varphi)$*. The limiting variable* Θ^{MD} *is centered normally distributed with variance*

$$\sigma^2 = \int_{[0,\pi/2]^2} \gamma_{\theta_B}(\varphi)\gamma_{\theta_B}(\varphi')r(\cos\varphi, \sin\varphi, \cos\varphi', \sin\varphi')\, d(\varphi,\varphi'),$$

where r *denotes the covariance functional of the process* $\mathbb{G}_{\hat\Lambda_L}$ *defined in (2.11).*

In order to make use of the latter result in statistical applications one needs the quantiles of the limiting distribution. We propose to use the multiplier bootstrap discussed in the previous section. The following theorem shows that the *pdm* and *dm* bootstrap yield a valid approximation of the distribution of the random variable Θ^{MD}.

Theorem 4.2. *If the assumptions of the Theorems 4.1, 3.3 and 3.4 hold and* Γ_n *denotes either the process* α_n^{pdm} *(Theorem 3.3) or* α_n^{dm} *(Theorem 3.4) obtained by the pdm- or dm-bootstrap, respectively, then*

$$\Theta_n^{MD,m} := \int \gamma_{\hat\theta_n^{MD}}(\varphi)\Gamma^{\angle}_n(\varphi)\, d\varphi \;\overset{\mathbb{P}}{\underset{\xi}{\rightsquigarrow}}\; \Theta^{MD},$$

where $\Gamma^{\angle}_n(\varphi) = \Gamma_n(\cos\varphi, \sin\varphi)$.

On the basis of this result it is possible to construct asymptotic confidence regions for the parameter θ as well as to test point hypotheses regarding the parameter. In Table 6 we present a small simulation study regarding the finite sample coverage probabilities of some confidence intervals for the parameter of a Clayton tail copula. The sample size is $n = 1000$ or $n = 3000$ and we used $B = 500$ bootstrap replications and 1000 simulation runs to calculate the coverage probabilities. The parameter of the Clayton tail copula is chosen in such a way that the tail dependence coefficient varies in the set $\{1/4, 2/4, 3/4\}$. As one can see we get accurate results in all cases. For small sample sizes the approximation works better for weak

n	λ_L	90%	95%	n	λ_L	90%	95%
1000	0.25	0.895	0.955	3000	0.25	0.833	0.914
	0.5	0.893	0.936		0.5	0.900	0.941
	0.75	0.838	0.887		0.75	0.890	0.935

Table 6: *Simulated coverage probability of the confidence intervals based on the pdm-bootstrap,* $n = 1000$ $(k = 50)$ *and* $n = 3000$ $(k = 100)$

tail dependence while for rather large sample sizes strong tail dependence yields slightly more accurate results.

It is also notable that the *dm-* and *pdm*-bootstrap can be used to construct a consistent approximation of the asymptotic distribution of the censored likelihood and moment estimator investigated in de Haan et al. (2008) and Einmahl et al. (2008).

In the following we will use the multiplier bootstrap to construct a computationally efficient goodness-of-fit test for the hypothesis that the lower tail copula has a specific parametric form, i.e.

$$\mathcal{H}_0 : \Lambda_L \in \mathcal{L}, \qquad \mathcal{H}_1 : \Lambda_L \notin \mathcal{L}. \tag{4.7}$$

This problem has also been discussed in de Haan et al. (2008) and Einmahl et al. (2008) who proposed a comparison between a nonparametric and a parametric estimate of the lower tail copula by an L^2-distance. In both cases the limiting distribution of the corresponding test statistic under the null hypothesis depends in a complicated way on the process $\mathbb{G}_{\hat{\Lambda}_L}$ and the unknown true parameter θ_B. While Einmahl et al. (2008) does not propose any bootstrap approximation, de Haan et al. (2008) proposed to use the parametric bootstrap. However, it was pointed out by Kojadinovic and Yan (2010) or Kojadinovic et al. (2010) that for copula models, approximations based on multiplier bootstraps are computationally more efficient, especially for large sample sizes. We will now illustrate how the multiplier bootstrap can be successfully applied in the problem of testing the hypothesis (4.7).

To be precise, we propose to compare a parametric (using the minimum distance estimate $\hat{\theta}_n^{MD}$) and a nonparametric estimate of the tail copula and to reject the null hypothesis (4.7) for large values of the statistic

$$GOF_n := k\, \varrho(\hat{\Lambda}_L, \Lambda_L(\cdot; \hat{\theta}_n^{MD})) = k \int \left(\hat{\Lambda}_L^{\angle}(\varphi) - \Lambda_L^{\angle}(\varphi; \hat{\theta}_n^{MD}) \right)^2 d\varphi,$$

where $\hat{\theta}_n^{MD}$ denotes the minimum distance estimate. If the assumptions (B.1) - (B.5) are satisfied we obtain for the process $H_n = \sqrt{k}\left(\hat{\Lambda}_L - \Lambda_L(\cdot; \hat{\theta}_n^{MD})\right)$ under the null hypothesis $\mathcal{H}_0 : \Lambda_L = \Lambda_L(\cdot; \theta_B)$

$$\begin{aligned} H_n &= \sqrt{k}\left(\hat{\Lambda}_L - \Lambda_L - (\hat{\Lambda}_L(\cdot, \hat{\theta}_n^{MD}) - \Lambda_L(\cdot, \theta))\right) \\ &= \sqrt{k}\left(\hat{\Lambda}_L - \Lambda_L - \delta_\theta(\hat{\theta}_n^{MD} - \theta)\right) + o_{\mathbb{P}^*}(1) \\ &= \sqrt{k}\left(\hat{\Lambda}_L - \Lambda_L - \delta_\theta \int \gamma_\theta(\varphi)(\hat{\Lambda}_L^{\angle}(\varphi) - \Lambda_L^{\angle}(\varphi)\, d\varphi\right) + o_{\mathbb{P}^*}(1) \end{aligned}$$

$$\rightsquigarrow \mathbb{G}_{\hat{\Lambda}_L} - \delta_\theta \int \gamma_\theta(\varphi)\mathbb{G}^{\angle}_{\hat{\Lambda}_L}(\varphi)\, d\varphi = \mathbb{G}_{\hat{\Lambda}_L} - \delta_\theta\, \Theta^{MD}.$$

Under the alternative hypothesis we get an additional summand

$$H_n = \sqrt{k}\left(\hat{\Lambda}_L - \Lambda_L - \delta_\theta(\hat{\theta}_n^{MD} - \theta) - (\Lambda_L(\cdot;\theta_B) - \Lambda_L)\right) + o_{\mathbb{P}^*}(1),$$

which converges to either plus or minus infinity whenever $\Lambda_L(\mathbf{x}, \theta_B) \neq \Lambda_L(\mathbf{x})$. The continuous mapping theorem yields the following result.

Theorem 4.3. *Assume that assumptions of Theorem 4.1 are satisfied. If the null hypothesis is valid then*

$$GOF_n = \int \{H_n^{\angle}(\varphi)\}^2\, d\varphi \rightsquigarrow \int \left(\mathbb{G}^{\angle}_{\hat{\Lambda}_L}(\varphi) - \delta^{\angle}_\theta(\varphi)\, \Theta^{MD}\right)^2 d\varphi,$$

while under the alternative

$$GOF_n = \int \{H_n^{\angle}(\varphi)\}^2\, d\varphi \xrightarrow{\mathbb{P}^*} \infty.$$

The critical values of the test, which rejects the null hypothesis for large values of GOF_n can be calculated on the basis of the following theorem. The proof is similar to the proof of Theorem 4.2 in the appendix and is therefore omitted.

Theorem 4.4. *If the assumptions of the Theorems 4.1, 3.3 and 3.4 hold and Γ_n denotes either the process α_n^{pdm} (Theorem 3.3) or α_n^{dm} (Theorem 3.4) obtained by the pdm- or dm-bootstrap, respectively, then it holds independently of the hypotheses that*

$$H_n^m := \Gamma_n - \delta_{\hat{\theta}_n^{MD}} \int \gamma_{\hat{\theta}_n^{MD}}(\varphi)\, \Gamma_n^{\angle}(\varphi)\, d\varphi \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\hat{\Lambda}_L} - \delta_{\theta_B}\, \Theta^{MD}.$$

Therefore

$$GOF_n^m = \int \{H_n^{m\angle}(\varphi)\}^2\, d\varphi \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \int \left(\mathbb{G}^{\angle}_{\hat{\Lambda}_L}(\varphi) - \delta^{\angle}_\theta(\varphi)\, \Theta^{MD}\right)^2 d\varphi.$$

In order to investigate the finite sample properties of a goodness-of-fit test on the basis of the multiplier bootstrap we show in Table 7 the simulated level of the *pdm*-bootstrap test

$$GOF_n > q_{1-\alpha}^{(pdm)} \tag{4.8}$$

where $q_{1-\alpha}^{(pdm)}$ denotes the $(1-\alpha)$ quantile of the bootstrap distribution. For the null hypothesis we considered as the parametric class the family of Clayton tail copulas. In particular we investigated three scenarios corresponding to a coefficient of tail dependence varying in $\{0.25, 0.5, 0.75\}$. The results are based on 1000 simulation runs, while the sample size is either $n = 1000$ and $k = 50$ or $n = 3000$ and $k = 200$. For each test $B = 500$ bootstrap replications with $Laplacian(0, 2)$ multipliers have been performed. We observe a reasonable power and approximation of the nominal level. Note that for the sample size $n = 1000$ the *pdm*-bootstrap test is conservative and this effect is increasing with the level of tail dependence.

n	λ_L	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$	n	λ_L	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$
1000	0.25	0.124	0.087	0.037	3000	0.25	0.129	0.090	0.048
	0.5	0.097	0.068	0.032		0.5	0.105	0.069	0.031
	0.75	0.091	0.048	0.018		0.75	0.084	0.056	0.026

Table 7: *Simulated rejection probabilities of the pdm-bootstrap test* (4.8) *for the hypothesis* (4.5) *under the null hypothesis;* $n = 1000$ $(k = 50)$, $n = 3000$ $(k = 200)$.

n	λ_L	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$	n	λ_L	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$
1000	1/12	0.095	0.052	0.017	3000	1/12	0.217	0.155	0.092
	2/12	0.124	0.066	0.029		2/12	0.374	0.292	0.176
	3/12	0.298	0.200	0.088		3/12	0.868	0.819	0.696

Table 8: *Simulated rejection probabilities of the pdm-bootstrap test* (4.8) *for the hypothesis* (4.5) *under the alternative* $C = 1/3C_{clay} + 2/3\Pi$, *where* Π *denotes the independence copula;* $n = 1000$ $(k = 50)$, $n = 3000$ $(k = 200)$.

Acknowledgements The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This work has been supported 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 John Einmahl for pointing out important references on the subject. We are also grateful to Johan Segers for discussing this subject with us in much detail.

Appendix A: Proofs

A.1 Proof of Lemma 2.1

Due to Theorem 1.6.1 in Van der Vaart and Wellner (1996) the proof of weak convergence of $\tilde{\alpha}_n$ in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ can be given for each $l^\infty(T_i)$ separately. To this end we note that every T_i can be written in the form $T = [0, M_1] \times \{\infty\} \cup \{\infty\} \times [0, M_2] \cup [0, M_3]^2$, where $M_1, M_2, M_3 \in \mathbb{N}$, and show weak convergence in $l^\infty(T)$. Recalling the notation of $f_{n,\mathbf{x}}(\mathbf{U}_i)$ in (3.3) we can express $\tilde{\alpha}_n$ as

$$\tilde{\alpha}_n(\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f_{n,\mathbf{x}}(\mathbf{U}_i) - \mathbb{E} f_{n,\mathbf{x}}(\mathbf{U}_i)),$$

and the assertion now follows by an application of Theorem 11.20 in Kosorok (2008). For this purpose we show that the assumptions for this result are satisfied. Let $\mathcal{F}_n = \{f_{n,\mathbf{x}} : \mathbf{x} \in T\}$ be a class of functions changing with n and denote by

$$F_n(\mathbf{u}) = \sqrt{\frac{n}{k}}\, \mathbb{I}\{u_1 \le kM/n \;\text{ or }\; u_2 \le kM/n\},$$

$M = M_1 \vee M_2 \vee M_3$ a corresponding sequence of envelopes of $\mathcal{F}_n$. We have to prove that

(*i*) $(\mathcal{F}_n, F_n)$ satisfies the bounded uniform entropy integral condition

$$\limsup_{n\to\infty} \sup_Q \int_0^1 \sqrt{\log N(\varepsilon ||F_n||_{Q,2}, \mathcal{F}_n, L_2(Q))}\, d\varepsilon < \infty, \tag{A.1}$$

where for each n the supremum ranges over all probability measures Q with finite support and $||F_n||_{Q,2} = \left(\int F_n(x)^2\, dQ(x)\right)^{1/2} > 0$.

(*ii*) The limit $H(\mathbf{x}, \mathbf{y}) = \lim_{n\to\infty} \mathbb{E}[\tilde{\alpha}_n(\mathbf{x})\tilde{\alpha}_n(\mathbf{y})]$ exists for every $\mathbf{x}$ and $\mathbf{y}$ in T .

(*iii*) $\limsup_{n\to\infty} \mathbb{E}F_n^2(\mathbf{U}_1) < \infty$

(*iv*) $\lim_{n\to\infty} \mathbb{E}F_n^2(\mathbf{U}_1)\mathbb{I}\{F_n(\mathbf{U}_1) > \varepsilon\sqrt{n}\} = 0$ for all $\varepsilon > 0$.

(*v*) $\lim_{n\to\infty} \rho_n(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \bar{R}_+^2$, where

$$\rho_n(\mathbf{x}, \mathbf{y}) = \left(\mathbb{E}(f_{n,\mathbf{x}}(U_1)) - f_{n,\mathbf{y}}(U_1))^2\right)^{1/2}. \tag{A.2}$$

Furthermore, for all sequences $(\mathbf{x}_n)_n, (\mathbf{y}_n)_n$ in T the convergence $\rho_n(\mathbf{x}_n, \mathbf{y}_n) \to 0$ holds, provided $\rho(\mathbf{x}_n, \mathbf{y}_n) \to 0$.

(*vi*) The sequence $\mathcal{F}_n$ of classes is almost measurable Suslin (AMS), i.e. for all $n \geq 1$ there exists a Suslin topological space $T_n \subset T$ with Borel sets $\mathcal{B}_n$ such that

(a) $\mathbb{P}^*(\sup_{\mathbf{x}\in T} \inf_{\mathbf{y}\in T_n} |f_{n,\mathbf{x}}(\mathbf{U}_1) - f_{n,\mathbf{y}}(\mathbf{U}_1)| > 0) = 0$,

(b) $f_{n,\cdot} : [0,1]^2 \times T_n \to \mathbb{R}$ is $\mathcal{B}|_{[0,1]^2} \times \mathcal{B}_n$-measurable for $i = 1, \ldots, n$.

In order to prove the bounded uniform entropy integral condition (i) we decompose $\mathcal{F}_n = \bigcup_{i=1}^3 \mathcal{F}_n^{(i)}$ with $\mathcal{F}_n^{(i)} = \{f_{n,\mathbf{x}}^{(i)}, \mathbf{x} \in T\}$ and

$$f_{n,\mathbf{x}}^{(1)}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}\{U_{i1} \leq kx_1/n\}\, \mathbb{I}\{x_2 = \infty\}, \quad f_{n,\mathbf{x}}^{(2)}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}\{U_{i2} \leq kx_2/n\}\, \mathbb{I}\{x_1 = \infty\},$$
$$f_{n,\mathbf{x}}^{(3)}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}\{U_{i1} \leq kx_1/n, U_{i2} \leq kx_2/n\}\, \mathbb{I}\{x_1 < \infty, x_2 < \infty\}.$$

The corresponding envelopes of the classes $\mathcal{F}_n^{(i)}$ are given by

$$F_n^{(1)}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}(U_{i1} \leq kM/n), \quad F_n^{(2)}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}(U_{i2} \leq kM/n),$$
$$F_n^{(3)}(\mathbf{U}_i) = \sqrt{\frac{n}{k}}\, \mathbb{I}(U_{i1} \leq kM/n, U_{i2} \leq kM/n),$$

so that $F_n(\mathbf{U}_i) = \max_{i=1}^3 \{F_n^{(i)}(\mathbf{U}_i)\}$. If we prove that the sequences $(\mathcal{F}_n^{(i)}, F_n^{(i)})$ satisfy the bounded uniform integral entropy condition given in (A.1), then the condition holds also for $(\mathcal{F}_n, F_n)$ by Lemma

B.1 in the appendix and thus the assertion in (i) is proved. We only consider the (hardest) case of $\mathcal{F}_n^{(3)}$. Note that $\mathcal{F}_n^{(3)} = \{f_{n,\mathbf{x}}, \mathbf{x} \in [0, M_3]^2\} = \mathcal{G}_n^{(1)} \cdot \mathcal{G}_n^{(2)}$, where

$$f_{n,\mathbf{x}} = (n/k)^{1/2}\, \mathbb{I}\{U_{i1} \le kx_1/n, U_{i2} \le kx_2/n\},$$
$$\mathcal{G}_n^{(j)} = \{g_{n,t} = (n/k)^{1/4}\mathbb{I}\{U_{ij} \le kt/n\} \mid t \in [0, M_3]\}$$

for $j = 1, 2$. Since the functions $g_{n,t}$ are increasing in t the $G_n^{(j)}$ are VC-classes with VC-index 2. Thus by Lemma 11.21 in Kosorok (2008) both classes satisfy the bounded uniform integral entropy condition (A.1). Proposition 11.22 in Kosorok (2008) shows that $\mathcal{F}_n^{(3)}$ has the same property and by the discussion at the beginning of this paragraph (i) is satisfied.

For the proof of (ii) note that $\mathbb{E}[\tilde{\alpha}_n(\mathbf{x})\tilde{\alpha}_n(\mathbf{y})] = n/k \left(C(\frac{(\mathbf{x}\wedge\mathbf{y})k}{n}) - C(\frac{\mathbf{x}k}{n})C(\frac{\mathbf{y}k}{n}) \right)$, which converges to $\Lambda_L(\mathbf{x} \wedge \mathbf{y}) =: H(\mathbf{x}, \mathbf{y})$, since $\frac{n}{k}C(\frac{\mathbf{x}k}{n})C(\frac{\mathbf{y}k}{n}) \to 0$.

Regarding (iii) and (iv) we note that $\mathbb{E}F_n(\mathbf{U}_1)^2 = 2M - \frac{n}{k}C(Mk/n, Mk/n)$, which converges to $2M - \Lambda_L(M, M)$. Further,

$$\mathbb{E}F_n^2(\mathbf{U}_1)\mathbb{I}\{F_n(\mathbf{U}_1) > \varepsilon\sqrt{n}\} = \int_{\{F_n(\mathbf{U}_1)>\varepsilon\sqrt{n}\}} F_n^2(\mathbf{U}_1)\, d\mathbb{P}$$
$$\le \frac{n}{k}\mathbb{P}(\frac{1}{k}\mathbb{I}\{U_{11} \le kM/n \text{ or } U_{12} \le kM/n\} > \varepsilon) = 0$$

for sufficiently large n, such that $k > 1/\varepsilon$. For (v) we note that

$$\rho_n(\mathbf{x}, \mathbf{y}) = \left(\mathbb{E}(f_{n,\mathbf{x}}(\mathbf{U}_1) - f_{n,\mathbf{y}}(\mathbf{U}_1))^2\right)^{1/2} = \sqrt{\frac{n}{k}}\, (C(\mathbf{x}k/n) - 2C((\mathbf{x} \wedge \mathbf{y})k/n) + C(\mathbf{y}k/n))^{1/2}$$
$$\to (\Lambda_L(\mathbf{x}) - 2\Lambda_L(\mathbf{x} \wedge \mathbf{y}) + \Lambda_L(\mathbf{y}))^{1/2} =: \rho(\mathbf{x}, \mathbf{y}).$$

Due to Theorem 1 in Schmidt and Stadtmüller (2006) we have locally uniform convergence in the latter expression, which yields the second condition stated in (v).

For the proof of condition (vi) we use Lemma 11.15 and the discussion on page 224 in Kosorok (2008) and show separability of $\mathcal{F}_n$, i.e. for every $n \ge 1$ there exists a countable subset $T_n \subset T$ such that

$$\mathbb{P}^* \left(\sup_{\mathbf{x}\in T} \inf_{\mathbf{y}\in T_n} |f_{n,\mathbf{y}}(\mathbf{U}_1) - f_{n,\mathbf{x}}(\mathbf{U}_1)| > 0 \right) = 0.$$

Choose $T_n = (\mathbb{Q} \cap [0, M_1] \times \{\infty\}) \cup (\{\infty\} \times \mathbb{Q} \cap [0, M_2]) \cup (\mathbb{Q}^2 \cap [0, M_3]^2)$, then we have (note that the functions $f_{n,\mathbf{x}}$ are built by indicators) that for every ω and every $\mathbf{x} \in T$ there is an $\mathbf{y} \in T_n$ with $|f_{n,\mathbf{x}}(\mathbf{U}_1(\omega)) - f_{n,\mathbf{y}}(\mathbf{U}_1(\omega))| = 0$. This yields the assertion and thus the proof of Lemma 2.1 is finished. □

A.2 Proof of Theorem 2.2

Let $\mathcal{B}_\infty(\mathbb{R}_+)$ denote the set of functions $f : \mathbb{R}_+ \to \mathbb{R}$ that are uniformly bounded on compact sets (equipped with the topology of uniform convergence on compact sets), define $\mathcal{B}_\infty^{I,0}(\mathbb{R}_+)$ as the subset of all non-decreasing functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ with $f(0) = 0$ and set

$$\mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) := \left\{\gamma \in \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2) \mid \gamma(\cdot, \infty) \in \mathcal{B}_\infty^{I,0}(\mathbb{R}_+), \gamma(\infty, \cdot) \in \mathcal{B}_\infty^{I,0}(\mathbb{R}_+)\right\}.$$

We define a map $\Phi : \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \to \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$ by

$$\gamma \longmapsto \Phi(\gamma) = \begin{cases} \gamma(\gamma^-(x,\infty), \gamma^-(\infty,y)) & , \text{if} \quad x, y \neq \infty \\ \gamma(\gamma^-(x,\infty), \infty) & , \text{if} \quad y = \infty \\ \gamma(\infty, \gamma^-(\infty, y)) & , \text{if} \quad x = \infty. \end{cases},$$

where f^- denotes the adjusted generalized inverse function of $f \in \mathcal{B}_\infty^{I,0}(\mathbb{R}_+)$ defined by

$$f^-(z) = \min\big\{ \inf\{x \in \mathbb{R} \,|\, f(x) \geq z\}, \sup(\operatorname{ran} f)\big\},$$

see also Schmidt and Stadtmüller (2006). Observing that the adjusted generalized inverse of $\tilde{\Lambda}_L(x,\infty)$ is ($\mathbb{P}$-almost surely) given by $\frac{n}{k}F_1(F_{n1}^-(kx/n))$, one can conclude that $\Phi(\Lambda_L) = \Lambda_L$ and $\Phi(\tilde{\Lambda}_L) = \hat{\Lambda}_L$ ($\mathbb{P}$-almost surely) and the proof of Theorem 2.2 follows from the functional delta method [Theorem 3.9.4 in Van der Vaart and Wellner (1996)] and the following Lemma, which is an extension of the result in the proof of Theorem 5 in Schmidt and Stadtmüller (2006).

Lemma A.5. *Let Λ_L be a lower tail copula whose partial derivatives satisfy the following first order properties*

$$\partial_p \Lambda_L \text{ exists on } \{\mathbf{x} \in \bar{\mathbb{R}}_+^2 \,|\, x_p < \infty\} \text{ and is continuous on } \{\mathbf{x} \in \bar{\mathbb{R}}_+^2 \,|\, 0 < x_p < \infty\}$$

for $p = 1, 2$. Then Φ is Hadamard-differentiable at Λ_L tangentially to the set

$$\mathcal{C}^0(\bar{\mathbb{R}}_+^2) = \big\{\gamma \in \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2) \mid \gamma \text{ continuous with } \gamma(\cdot, 0) = \gamma(0, \cdot) = 0\big\}.$$

Its derivative at Λ_L in $\gamma \in \mathcal{C}^0(\bar{\mathbb{R}}_+^2)$ is given by

$$\Phi'_{\Lambda_L}(\gamma)(\mathbf{x}) = \gamma(\mathbf{x}) - \partial_1\Lambda_L(\mathbf{x})\gamma(x_1, \infty) - \partial_2\Lambda_L(\mathbf{x})\gamma(\infty, x_2) \tag{A.3}$$

where $\partial_p \Lambda_L$, $p = 1, 2$ is defined as 0 on the set $\{\mathbf{x} \in \bar{\mathbb{R}}_+^2 \,|\, x_p = \infty\}$.

Proof. Decompose $\Phi = \Phi_3 \circ \Phi_2 \circ \Phi_1$ where

$$\begin{aligned} \Phi_1 :& \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \to \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \\ & \gamma \longmapsto (\gamma, \gamma(\cdot, \infty), \gamma(\infty, \cdot)) \\ \Phi_2 :& \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \to \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \\ & (\gamma, f, g) \longmapsto (\gamma, f^-, g^-)) \\ \Phi_3 :& \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \to \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2) \\ & (\gamma, f, g) \longmapsto \begin{cases} \gamma(f(x), g(y)) & , \text{if} \quad x, y \neq \infty \\ \gamma(f(x), \infty) & , \text{if} \quad y = \infty \\ \gamma(\infty, g(y)) & , \text{if} \quad x = \infty. \end{cases} \end{aligned}$$

Now Φ_1 is Hadamard-differentiable at Λ_L tangentially to $\mathcal{C}^0(\bar{\mathbb{R}}_+^2)$ since it is linear and continuous. The second map Φ_2 is Hadamard-differentiable at $(\Lambda_L, \mathrm{id}_{\mathbb{R}_+}, \mathrm{id}_{\mathbb{R}_+})$ tangentially to $\mathcal{C}^0(\bar{\mathbb{R}}_+^2) \times \mathcal{C}^0(\mathbb{R}_+) \times \mathcal{C}^0(\mathbb{R}_+)$

where $\mathcal{C}^0(\mathbb{R}_+)$ consists of all continuous functions f on $\mathbb{R}_+ = [0,\infty)$ with $f(0) = 0$ and the derivative of Φ_2 at $(\Lambda_L, \mathrm{id}_{\mathbb{R}_+}, \mathrm{id}_{\mathbb{R}_+})$ in (γ, f, g) is given by $\Phi'_{2,(\Lambda_L,\mathrm{id}_{\mathbb{R}_+},\mathrm{id}_{\mathbb{R}_+})}(\gamma, f, g) = (\gamma, -f, -g)$ [see Schmidt and Stadtmüller (2006), p. 321]. Some more efforts are necessary to show that Φ_3 is Hadamard-differentiable at $(\Lambda_L, \mathrm{id}_{\mathbb{R}_+}, \mathrm{id}_{\mathbb{R}_+})$ tangentially to $\mathcal{C}^0(\bar{\mathbb{R}}^2_+) \times \mathcal{C}^0(\mathbb{R}_+) \times \mathcal{C}^0(\mathbb{R}_+)$ with derivative

$$\Phi'_{3,(\Lambda_L,\mathrm{id}_{\mathbb{R}_+},\mathrm{id}_{\mathbb{R}_+})}(\gamma, f, g)(\mathbf{x}) = \gamma(\mathbf{x}) + \partial_1\Lambda_L(\mathbf{x})f(x_1) + \partial_2\Lambda_L(\mathbf{x})g(x_2).$$

To see this let $t_n \to 0$, $(\gamma_n, f_n, g_n) \in \mathcal{B}_\infty(\bar{\mathbb{R}}^2_+) \times \mathcal{B}_\infty(\mathbb{R}_+) \times \mathcal{B}_\infty(\mathbb{R}_+)$ with $(\gamma_n, f_n, g_n) \to (\gamma, f, g) \in \mathcal{C}^0(\bar{\mathbb{R}}^2_+) \times \mathcal{C}^0(\mathbb{R}_+) \times \mathcal{C}^0(\mathbb{R}_+)$ such that $(\Lambda_L + t_n\gamma_n, \mathrm{id}_{\mathbb{R}_+} + t_n f_n, \mathrm{id}_{\mathbb{R}_+} + t_n g_n) \in \mathcal{B}^{I,0}_\infty(\bar{\mathbb{R}}^2_+) \times \mathcal{B}^{I,0}_\infty(\mathbb{R}_+) \times \mathcal{B}^{I,0}_\infty(\mathbb{R}_+)$. Now Φ_3 is linear in its first argument and we introduce the decomposition

$$t_n^{-1}\left\{\Phi_3(\Lambda_L + t_n\gamma_n, \mathrm{id}_{\mathbb{R}_+} + t_n f_n, \mathrm{id}_{\mathbb{R}_+} + t_n g_n) - \Phi_3(\Lambda_L, \mathrm{id}_{\mathbb{R}_+}, \mathrm{id}_{\mathbb{R}_+})\right\} = L_{n1} + L_{n2},$$

where

$$\begin{aligned} L_{n1} &= t_n^{-1}\left\{\Phi_3(\Lambda_L, \mathrm{id}_{\mathbb{R}_+} + t_n f_n, \mathrm{id}_{\mathbb{R}_+} + t_n g_n) - \Phi_3(\Lambda_L, \mathrm{id}_{\mathbb{R}_+}, \mathrm{id}_{\mathbb{R}_+})\right\} \\ L_{n2} &= \Phi_3(\gamma_n, \mathrm{id}_{\mathbb{R}_+} + t_n f_n, \mathrm{id}_{\mathbb{R}_+} + t_n g_n). \end{aligned}$$

By the definition of d it suffices to show uniform convergence on sets T of the form $T = [0, M_1] \times \{\infty\} \cup \{\infty\} \times [0, M_2] \cup [0, M_3]^2$, where $M_1, M_2, M_3 \in \mathbb{N}$. Since $T \subset \bar{\mathbb{R}}^2_+$ is compact (f_n, g_n) converges uniformly and γ is uniformly continuous; hence L_{n2} uniformly converges to γ.
Considering L_{n1} we split the investigation into six different cases. First, let $\mathbf{x} \in (0, M_3]^2$. A series expansion at $\mathbf{x}$ yields

$$L_{n1} = \partial_1\Lambda_L(\mathbf{x})f_n(x_1) + \partial_2\Lambda_L(\mathbf{x})g_n(x_2) + r_n(\mathbf{x}),$$

where the error term r_n can be written as

$$r_n(\mathbf{x}) = \big(\partial_1\Lambda_L(\mathbf{y}) - \partial_1\Lambda_L(\mathbf{x})\big)f_n(x_1) + \big(\partial_2\Lambda_L(\mathbf{y}) - \partial_2\Lambda_L(\mathbf{x})\big)g_n(x_2)$$

with some intermediate point $\mathbf{y} = \mathbf{y}(n)$ between $\mathbf{x}$ and $(x_1 + t_n f_n(x_1), x_2 + t_n f_n(x_2))$. The dominating term converges uniformly to $\partial_1\Lambda_L(\mathbf{x})f(x_1) + \partial_2\Lambda_L(\mathbf{x})g(x_2)$, hence it remains to show that $r_n(\mathbf{x})$ converges to 0 uniformly in $\mathbf{x}$. For a given $\varepsilon > 0$ uniform convergence of f_n and uniform continuity of f on $[0, M_3]$ as well as the fact that $f(0) = 0$ allows to choose a $\delta > 0$ such that $|f_n(x_1)| < \varepsilon$ for all $x_1 < \delta$. Since partial derivatives of tail copulas are bounded by 1, the first term of $r_n(\mathbf{x})$ is uniformly small for $x_1 < \delta$. On the quadrangle $[\delta, M_3] \times (0, M_3]$ the partial derivative $\partial_1\Lambda_L$ is uniformly continuous which yields the desired convergence under consideration of $\mathbf{y}(n) \to \mathbf{x}$ and boundedness of f. The same arguments apply for the second derivative and the case $\mathbf{x} \in (0, M_3]^2$ is finished.
Now consider the case $\mathbf{x} \in (0, M_3] \times \{0\}$. By Lipschitz-continuity of Λ_L on $\mathbb{R}^2_+$ we get

$$\begin{aligned} |L_{n1}(x_1, 0)| &= t_n^{-1}|\Lambda_L(x_1 + t_n f_n(x_1), t_n g_n(0))| = t_n^{-1}|\Lambda_L(x_1 + t_n f_n(x_1), t_n g_n(0)) - \Lambda_L(x_1 + t_n f_n(x_1), 0)| \\ &\le |g_n(0)| \to g(0) = 0. \end{aligned}$$

Since $\partial_1 \Lambda_L(x_1,0)f(x_1) + \partial_2\Lambda_L(x_1,0)g(0) = 0$ this yields the assertion. For the cases $\mathbf{x} = (0,0)$ and $\mathbf{x} \in \{0\} \times (0, M_3]$ the arguments are similar and we proceed with $\mathbf{x} \in [0, M_1] \times \{\infty\}$ (and analogously $\mathbf{x} \in \{\infty\} \times [0, M_2]$)

$$L_{n1}(x_1,\infty) = t_n^{-1}(\Lambda_L(x_1 + t_n f_n(x_1), \infty) - \Lambda_L(x_1,\infty)) = f_n(x_1) \to f(x_1).$$

By $\partial_1\Lambda_L(x_1,\infty) = 1$ and $\partial_2\Lambda_L(x_1,\infty) = 0$ this yields the assertion. To conclude, Φ_3 is Hadamard-differentiable as asserted.
An application of the chain rule [see Lemma 3.9.3 in Van der Vaart and Wellner (1996)] completes the proof of the Lemma. □

A.3 Proof of Theorem 3.1

Due to Lemma B.2 in the Appendix B we can proceed as in the proof of Lemma 2.1 and consider just convergence in $l^\infty(T)$ with $T = [0, M_1] \times \{\infty\} \cup \{\infty\} \times [0, M_2] \cup [0, M_3]^2$, where M_1, M_2 and M_3 are arbitrary constants in $\mathbb{N}$. The assertion now follows from Theorem 11.23 in Kosorok (2008), because the corresponding sufficient conditions have already been established in the proof of Lemma 2.1. □

A.4 Proof of Theorem 3.4

For technical reasons we give a proof of Theorem 3.4 in advance of Theorem 3.2 and 3.3. The proof is essentially a consequence of a bootstrap version of the functional delta method, see Theorem 12.1 in Kosorok (2008). Since this result only holds for Banach space valued stochastic processes some adjustments have to be made. Note that the space $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ is a complete topological vector space with a metric d and some care is necessary whenever technical results depending on the norm are used.
Due to Lemma 2.1 and Theorem 3.1 we have

$$\sqrt{k}\big(\tilde{\Lambda}_L - \Lambda_L\big) \rightsquigarrow \mathbb{G}_{\tilde{\Lambda}_L}, \ \sqrt{k}\,\frac{\mu}{\tau}\big(\tilde{\Lambda}_L^\xi - \tilde{\Lambda}_L\big) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\tilde{\Lambda}_L}$$

in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$. Observing that the generalized inverses of $\tilde{\Lambda}_L(x,\infty)$ and $\tilde{\Lambda}_L^\xi(x,\infty)$ are ($\mathbb{P}$-almost surely) given by $\frac{n}{k}F_1(F_{n1}^-(kx/n))$ and $\frac{n}{k}F_1(F_{n1}^{\xi-}(kx/n))$, respectively, one can conclude that $\Phi(\Lambda_L) = \Lambda_L, \Phi(\tilde{\Lambda}_L) = \hat{\Lambda}_L$ and $\Phi(\tilde{\Lambda}_L^\xi) = \hat{\Lambda}_L^{\xi,\xi}$ ($\mathbb{P}$-almost surely). By Lemma A.5 Φ is Hadamard-differentiable on $\mathcal{B}^I_\infty(\bar{\mathbb{R}}^2_+)$ at $\gamma_0 = \Lambda_L$ tangentially to $\mathcal{C}^0(\bar{\mathbb{R}}^2_+) \subset \mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$. Therefore it remains to argue why Theorem 12.1 in Kosorok (2008) can be applied in the present context.
A careful inspection of the proof of Theorem 12.1 in Kosorok (2008) shows that properties going beyond our specific assumptions (i.e. the complete topological vector space $(\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+), d)$) are used only three times. First of all the mapping Φ'_{Λ_L} needs to be extended to the whole space $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$, which is possible using equation (A.3) as the defining identity. Secondly, the proof of Theorem 12.1 in Kosorok (2008) uses the usual functional delta method as stated in Theorem 2.8 in the same reference, but this result can be replaced by Theorem 3.9.4 in Van der Vaart and Wellner (1996), which provides a functional delta method holding in general metrizable topological vector spaces. Finally, the proof of Theorem 12.1 in

Kosorok (2008) makes use of a bootstrap continuous mapping theorem, see Theorem 10.8 in Kosorok (2008), which would yield that

$$\sqrt{k}\frac{\mu}{\tau}(\tilde{\Lambda}_L^{\xi} - \tilde{\Lambda}_L) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\tilde{\Lambda}_L} \quad \Rightarrow \quad \Phi'_{\Lambda_L}(\sqrt{k}\frac{\mu}{\tau}(\tilde{\Lambda}_L^{\xi} - \tilde{\Lambda}_L)) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \Phi'_{\Lambda_L}(\mathbb{G}_{\tilde{\Lambda}_L}).$$

In our specific context this statement follows immediately from the Lipschitz continuity of the derivative Φ'_{Λ_L} and an application of Lemma B.3 in Appendix B. □

A.5 Proof of Theorem 3.2.

Consider the mapping $\Psi : \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \longrightarrow \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$ defined by $\Psi = \Phi_3 \circ \Phi_2 \circ \Psi_1$, where Φ_3 and Φ_2 are defined in the proof of Lemma A.5 and Ψ_1 is given by

$$\begin{aligned}\Psi_1 : &\mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \to \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{I,0}(\mathbb{R}_+) \\ &(\beta, \gamma) \longmapsto (\beta, \gamma(\cdot, \infty), \gamma(\infty, \cdot)).\end{aligned}$$

Note that we obtain the representations $\Psi(\Lambda_L, \Lambda_L) = \Lambda_L$, $\Psi(\tilde{\Lambda}_L, \tilde{\Lambda}_L) = \hat{\Lambda}_L$ and $\Psi(\tilde{\Lambda}_L^{\xi}, \tilde{\Lambda}_L) = \hat{\Lambda}_L^{\xi,\cdot}$ ($\mathbb{P}$-almost surely). Clearly, Ψ_1 is Hadamard-differentiable at (Λ_L, Λ_L) since it is linear and continuous. Φ_2 and Φ_3 are Hadamard-differentiable tangentially to suitable subspaces as well, see the proof of Lemma A.5. By an application of the chain rule, see Lemma 3.9.3 in Van der Vaart and Wellner (1996), we can conclude that Ψ is Hadamard-differentiable (Λ_L, Λ_L) tangentially to $\mathcal{C}^0(\bar{\mathbb{R}}_+^2) \times \mathcal{C}^0(\bar{\mathbb{R}}_+^2)$ with derivative

$$\Psi'_{(\Lambda_L,\Lambda_L)}(\beta, \gamma)(\mathbf{x}) = \beta(\mathbf{x}) - \partial_1\Lambda_L(\mathbf{x})\gamma(x_1, \infty) - \partial_2\Lambda_L(\mathbf{x})\gamma(\infty, x_2). \tag{A.4}$$

Note that, unlike in the previous proof, we do not have weak convergence (resp. weak conditional convergence) of $\sqrt{k}\left((\tilde{\Lambda}_L, \tilde{\Lambda}_L) - (\Lambda_L, \Lambda_L)\right)$ and $\frac{\mu}{\tau}\sqrt{k}\left((\tilde{\Lambda}_L^{\xi}, \tilde{\Lambda}_L) - (\tilde{\Lambda}_L, \tilde{\Lambda}_L)\right)$ towards the same limiting field, which would be necessary for an application of the functional delta method for the bootstrap [see for example Theorem 12.1 in Kosorok (2008)]. Nevertheless, we can mimic certain steps in the proof of this theorem to conclude the result. To be precise, note that we obtain by analogous arguments as on page 236 of Kosorok (2008) that

$$\sqrt{k}\begin{pmatrix} \tilde{\Lambda}_L^{\xi} - \Lambda_L \\ \tilde{\Lambda}_L - \Lambda_L \end{pmatrix} \rightsquigarrow \begin{pmatrix} c^{-1}\mathbb{G}_1 + \mathbb{G}_2 \\ \mathbb{G}_2 \end{pmatrix},$$

unconditionally, where $\mathbb{G}_1$ and $\mathbb{G}_2$ denote independent copies of $\mathbb{G}_{\tilde{\Lambda}_L}$ and $c = \mu\tau^{-1}$. Hadamard-differentiability of the mapping $(\beta, \gamma) \mapsto (\Psi(\beta, \gamma), \Psi(\gamma, \gamma), (\beta, \gamma), (\gamma, \gamma))$ and the usual functional delta method [Theorem 3.9.4 in Van der Vaart and Wellner (1996)] yields

$$\sqrt{k}\begin{pmatrix} \Psi(\tilde{\Lambda}_L^{\xi}, \tilde{\Lambda}_L) - \Psi(\Lambda_L, \Lambda_L) \\ \Psi(\tilde{\Lambda}_L, \tilde{\Lambda}_L) - \psi(\Lambda_L, \Lambda_L) \\ (\tilde{\Lambda}_L^{\xi}, \tilde{\Lambda}_L) - (\Lambda_L, \Lambda_L) \\ (\tilde{\Lambda}_L, \tilde{\Lambda}_L) - (\Lambda_L, \Lambda_L) \end{pmatrix} \rightsquigarrow \begin{pmatrix} \Psi'_{(\Lambda_L,\Lambda_L)}(c^{-1}\mathbb{G}_1 + \mathbb{G}_2, \mathbb{G}_2) \\ \Psi'_{(\Lambda_L,\Lambda_L)}(\mathbb{G}_2, \mathbb{G}_2) \\ (c^{-1}\mathbb{G}_1 + \mathbb{G}_2, \mathbb{G}_2) \\ (\mathbb{G}_2, \mathbb{G}_2) \end{pmatrix}.$$

Observing that $\Psi'_{(\Lambda_L,\Lambda_L)}$ is linear we can conclude that

$$c\sqrt{k}\begin{pmatrix}\Psi(\tilde\Lambda_L^\xi,\tilde\Lambda_L)-\Psi(\tilde\Lambda_L,\tilde\Lambda_L)\\ (\tilde\Lambda_L^\xi,\tilde\Lambda_L)-(\tilde\Lambda_L,\tilde\Lambda_L)\end{pmatrix}\rightsquigarrow\begin{pmatrix}\Psi'_{(\Lambda_L,\Lambda_L)}(\mathbb{G}_1,0)\\ (\mathbb{G}_1,0)\end{pmatrix}=\begin{pmatrix}\mathbb{G}_1\\ (\mathbb{G}_1,0)\end{pmatrix}.$$

Continuity of the map $(\alpha,\beta,\gamma)\mapsto d(\alpha,\beta)$ yields

$$d\left(c\sqrt{k}\left(\Psi(\tilde\Lambda_L^\xi,\tilde\Lambda_L)-\Psi(\tilde\Lambda_L,\tilde\Lambda_L)\right),c\sqrt{k}\left(\tilde\Lambda_L^\xi-\tilde\Lambda_L\right)\right)\longrightarrow 0$$

in outer probability and thus by boundedness of the metric d also in outer expectation. Since $c\sqrt{k}(\tilde\Lambda_L^\xi-\tilde\Lambda_L)\underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}}\mathbb{G}_1$ we obtain the assertion by Lemma B.4. □

A.6 Proof of Theorem 3.3.

Let T be a set of the form $T=[0,M_1]\times\{\infty\}\cup\{\infty\}\times[0,M_2]\cup[0,M_3]^2$, see also the beginning of the proof of Lemma 2.1. We start the proof with an assertion regarding consistency of $\widehat{\partial_p\Lambda_L}$ and claim that for any $\delta\in(0,1)$

$$\sup_{\mathbf{x}\in T:x_p\ge\delta}\left|\widehat{\partial_p\Lambda_L}(\mathbf{x})-\partial_p\Lambda_L(\mathbf{x})\right|\longrightarrow 0 \tag{A.5}$$

in outer probability. For a proof of (A.5) split T into three subsets as indicated by its definition and then proceed similar as in the proof of Lemma 4.1 in Segers (2010). The details are omitted. Regarding the assertion of the Theorem we set

$$\bar\alpha_n^{pdm}(\mathbf{x})=\beta_n(\mathbf{x})-\partial_1\Lambda_L(\mathbf{x})\beta_n(x_1,\infty)-\partial_2\Lambda_L(\mathbf{x})\beta_n(\infty,x_2).$$

Under consideration of Lemma B.4 it suffices to prove that $d(\alpha_n^{pdm},\bar\alpha_n^{pdm})$ converges to 0 in outer probability. By the definition of d we have to show uniform convergence on the set T. Since $|\alpha_n^{pdm}-\bar\alpha_n^{pdm}|\le D_{n1}+D_{n2}$, where

$$D_{n1}=\left|\widehat{\partial_1\Lambda_L}-\partial_1\Lambda_L\right||\beta_n(\cdot,\infty)|,\qquad D_{n2}=\left|\widehat{\partial_2\Lambda_L}-\partial_2\Lambda_L\right||\beta_n(\infty,\cdot)|$$

we can consider both summands D_{np} separately and deal with D_{n1} exemplarily. First consider the case $\mathbf{x}\in[0,M_3]^2$, then for arbitrary $\varepsilon>0$ and $\delta\in(0,1)$

$$\mathbb{P}^*\left(\sup_{\mathbf{x}\in[0,M_3]^2}D_{n1}(\mathbf{x})>\varepsilon\right)\le\mathbb{P}^*\left(\sup_{\mathbf{x}\in[0,M_3]^2,x_1\ge\delta}D_{n1}(\mathbf{x})>\varepsilon/2\right)+\mathbb{P}^*\left(\sup_{\mathbf{x}\in[0,M_3]^2,x_1<\delta}D_{n1}(\mathbf{x})>\varepsilon/2\right). \tag{A.6}$$

Since $\widehat{\partial_1\Lambda_L}$ is uniformly consistent on $\{\mathbf{x}\in[0,M_3]^2\,|\,x_1\ge\delta\}$ and since β_n is asymptotically tight in $l^\infty(T)$ [β_n converges unconditionally by the results in Chapter 10 of Kosorok (2008)] the first probability on the right-hand side converges to zero.

Regarding the second summand note that $F_{n1}^{-}(kx/n)=X_{\lceil kx\rceil:n,1}$ (where $\lceil x\rceil=\min\{k\in\mathbb{Z}\,|\,k\ge x\}$) so that

$$\sup_{\mathbf{x}\in[0,M_3]^2}\left|\widehat{\partial_1\Lambda_L}(\mathbf{x})\right|\le\sup_{\mathbf{x}\in[0,M_3]^2,x_1\ge h}\frac{\lceil k(x_1+h)\rceil-\lceil k(x_1-h)\rceil}{2h}\le 1+\frac{M_3}{2kh}\le 2$$

for sufficiently large n. Hence the right-hand side of equation (A.6) is bounded by

$$\mathbb{P}^*\left(\sup_{\mathbf{x}\in[0,M_3]^2, x_1<\delta} |\beta_n(\mathbf{x})| > \varepsilon/4\right),$$

eventually. As $\beta_n \rightsquigarrow \mathbb{G}_{\tilde{\Lambda}_L}$ (unconditionally) the lim sup of this outer probability is bounded by

$$\mathbb{P}\left(\sup_{\mathbf{x}\in[0,M_3]^2, x_1<\delta} |\mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x})| > \varepsilon/4\right).$$

Since $\mathbb{G}_{\tilde{\Lambda}_L}$ has continuous trajectories and $\mathbb{G}_{\tilde{\Lambda}_L}(0,x_2)=0$ (almost surely) this probability can be made arbitrary small by choosing δ sufficiently small. The case $\mathbf{x}\in[0,M_3]^2$ is finished. For $\mathbf{x}\in[0,M_1]\times\{\infty\}$ the arguments are similar, while for $\mathbf{x}\in\{\infty\}\times[0,M_2]$ we have $D_{n1}=0$ and nothing has to be shown. To conclude, $\sup_{\mathbf{x}\in T} D_{n1}(\mathbf{x})$ converges to zero in outer probability and because the term $\sup_{\mathbf{x}\in T} D_{n_2}$ can be treated similarly the proof is finished. □

A.7 Proof of Theorem 4.1

The consistency follows by Theorem 5.9 in Van der Vaart (1998) observing the inequality

$$\sup_{\theta\in\Theta}||\psi_n(\theta)-\psi(\theta)|| \le 2\int \sup_{\theta\in\Theta}||\delta_\theta^\angle(\varphi)||\,|\hat{\Lambda}_L^\angle(\varphi)-\Lambda_L^\angle(\varphi)|\,d\varphi = o_{\mathbb{P}^*}(1)$$

and the consistency of the empirical tail copula.
Regarding the asymptotic normality note that by a Taylor expansion we have

$$0=\psi_n(\hat{\theta}_n^{MD})=\psi_n(\theta_b)+\partial_\theta\,\psi_n(\bar{\theta})(\hat{\theta}_n^{MD}-\theta_B),$$

where $\parallel \bar{\theta}-\theta_B \parallel \le \parallel \theta_B-\hat{\theta}_n^{MD} \parallel$. Due to consistency of both the empirical tail copula $\hat{\Lambda}_L$ and the MD-estimator $\hat{\theta}_n^{MD}$ we can conclude (note that the functions $\Lambda_L(\cdot\,;\theta)$, δ_θ and $\partial_\theta\,\delta_\theta$ are continuous in θ_B) that

$$\partial_\theta\,\psi_n(\bar{\theta}) = 2\int \delta_{\bar{\theta}}^\angle(\varphi)\delta_{\bar{\theta}}^\angle(\varphi)^T + \partial_\theta\,\delta_{\bar{\theta}}^\angle(\varphi)(\Lambda_L^\angle(\varphi;\bar{\theta})-\hat{\Lambda}_L^\angle(\varphi))\,d\varphi \stackrel{\mathbb{P}^*}{\longrightarrow} 2\,A_{\theta_B}.$$

Since $0=\psi(\theta_B)=2\int\delta_{\theta_B}^\angle(\varphi)(\Lambda_L^\angle(\varphi,\theta_B)-\Lambda_L^\angle(\varphi))\,d\varphi$ we obtain

$$2\int \delta_{\theta_B}^\angle(\varphi)(\hat{\Lambda}_L^\angle(\varphi)-\Lambda_L^\angle(\varphi))\,d\varphi = -\psi_n(\theta_B) = 2\,(A_{\theta_B}+o_{\mathbb{P}^*}(1))(\hat{\theta}_n^{MD}-\theta_B).$$

The probability that $(2\,A_{\theta_B}+o_{\mathbb{P}^*}(1))$ is invertible, converges to one, which yields the assertion by multiplying the last equality with $\sqrt{k}\,1/2(A_{\theta_B}+o_{\mathbb{P}^*}(1))^{-1}$. □

A.8 Proof of Theorem 4.2

Since $\Gamma_n \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\hat{\Lambda}_L}$ and $\gamma_{\hat{\theta}_n^{MD}} \stackrel{\mathbb{P}^*}{\to} \gamma_{\theta_B}$ (consistency of $\hat{\theta}_n^{MD}$ for θ_B) it is easy to see that

$$\sup_{\mathbf{x}\in[0,1]^2} |\gamma_{\hat{\theta}_n^{MD}}\Gamma_n(\mathbf{x})-\gamma_{\theta_B}\Gamma_n(\mathbf{x})|$$

converges to zero in outer probability. Hence, by Lemma B.4, $\gamma_{\hat{\theta}_n^{MD}}\Gamma_n \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\hat{\Lambda}_L}$ in $l^\infty[0,1]^2$ and the assertion follows invoking the Lipschitz-continuous mapping Theorem for the bootstrap, Lemma B.3. □

B Appendix B: Auxiliary results

Lemma B.1. *Suppose $\mathcal{G}_n$ and $\mathcal{H}_n$ are sequences of measurable functions with envelopes G_n and H_n, so that $(\mathcal{G}_n, G_n)$ and $(\mathcal{H}_n, H_n)$ satisfy the bounded uniform integral entopry condition as stated in (A.1). Then the bounded uniform entropy integral condition (A.1) holds also for $\mathcal{F}_n = \mathcal{G}_n \cup \mathcal{H}_n$, with envelopes $F_n = G_n \vee H_n$.*

Proof. Note that

$$\begin{aligned} N &:= N(\varepsilon||F_n||_{Q,2}, \mathcal{F}_n, L_2(Q)) \\ &\leq N(\varepsilon||F_n||_{Q,2}, \mathcal{G}_n, L_2(Q)) + N(\varepsilon||F_n||_{Q,2}, \mathcal{H}_n, L_2(Q)) \leq N_1 + N_2, \end{aligned}$$

where $N_1 = N(\varepsilon||G_n||_{Q,2}, \mathcal{G}_n, L_2(Q))$ and $N_2 = N(\varepsilon||H_n||_{Q,2}, \mathcal{H}_n, L_2(Q))$. By monotonicity and subadditivity of $\log(n)$ and $\sqrt{n}$ for $n \geq 2$ we obtain the inequality

$$\sqrt{\log N} \leq \sqrt{\log N_1} + \sqrt{\log N_2},$$

which yields the assertion. □

Lemma B.2. *Suppose $G_n = G_n(\mathbf{X}_1, \dots \mathbf{X}_n, \xi_1, \dots \xi_n)$ is some statistic taking values in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$. Then a conditional version of Theorem 1.6.1 in Van der Vaart and Wellner (1996) holds , namely $G_n \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} G$ in $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ is equivalent to $G_n \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} G$ in $l^\infty(T_i)$ for every $i \in \mathbb{N}$.*

Proof. We first show that

$$a_n = \sup_{h \in BL_1(\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+))} |\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)| \xrightarrow{\mathbb{P}^*} 0 \tag{B.1}$$

is equivalent to

$$a_n(T_i) = \sup_{h \in BL_1(l^\infty(T_i))} |\mathbb{E}_\xi h(G_n|_{T_i}) - \mathbb{E}h(G|_{T_i})| \xrightarrow{\mathbb{P}^*} 0 \tag{B.2}$$

for all $i \in \mathbb{N}$.
Suppose first that (B.1) holds. For arbitrary $h \in BL_1(l^\infty(T_i))$ we consider the mapping

$$h' : \mathcal{B}_\infty(\bar{\mathbb{R}}^2_+) \to \mathbb{R}, \alpha \mapsto h(\alpha|_{T_i}).$$

Then

$$\begin{aligned} |h'(\alpha) - h'(\beta)| &\leq \sup_{\mathbf{x} \in T_i} |\alpha(\mathbf{x}) - \beta(\mathbf{x})| \wedge 2 \leq 2 \sum_{j=1}^{i} \left(\sup_{\mathbf{x} \in T_j} |\alpha(\mathbf{x}) - \beta(\mathbf{x})| \wedge 1 \right) \\ &\leq 2^{i+1} \sum_{j=1}^{i} 2^{-j} \left(\sup_{\mathbf{x} \in T_j} |\alpha(\mathbf{x}) - \beta(\mathbf{x})| \wedge 1 \right) \leq 2^{i+1} d(\alpha, \beta), \end{aligned}$$

and therefore the mapping $h'' := 2^{-i-1}h'$ is an element of $BL_1(\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2))$. Observing

$$\begin{aligned}|\mathbb{E}_\xi h(G_n|_{T_i}) - \mathbb{E}h(G|_{T_i})| &= 2^{i+1}|\mathbb{E}_\xi h''(G_n) - \mathbb{E}h''(G)| \\ &\leq 2^{i+1} \sup_{h\in BL_1(\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2))} |\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)|,\end{aligned}$$

this yields the assertion.
Now suppose that (B.2) holds for all $i \in \mathbb{N}$. For $h \in BL_1(\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2))$ define $h_i(\alpha) = h(\alpha\,\mathbb{I}_{T_i})$, then $h_i \in BL_1(l^\infty(T_i))$ by the following reasoning: Obviously $||h_i|| \leq 1$ and

$$\begin{aligned}|h_i(\alpha) - h_i(\beta)| = |h(\alpha\,\mathbb{I}_{T_i}) - h(\beta\,\mathbb{I}_{T_i})| &\leq \sum_{j=1}^\infty 2^{-j}\left(\sup_{\mathbf{x}\in T_j}\big|(\alpha\,\mathbb{I}_{T_i})(\mathbf{x}) - (\beta\,\mathbb{I}_{T_i})(\mathbf{x})\big| \wedge 1\right) \\ &= \sum_{j=1}^i 2^{-j}\left(\sup_{\mathbf{x}\in T_j}\big|\alpha(\mathbf{x}) - \beta(\mathbf{x})\big| \wedge 1\right) + \left(\sup_{\mathbf{x}\in T_i}\big|\alpha(\mathbf{x}) - \beta(\mathbf{x})\big| \wedge 1\right)\sum_{j=i+1}^\infty 2^{-j} \\ &\leq \left(\sup_{\mathbf{x}\in T_i}\big|\alpha(\mathbf{x}) - \beta(\mathbf{x})\big| \wedge 1\right)\sum_{j=1}^\infty 2^{-j} \\ &\leq \sup_{\mathbf{x}\in T_i}\big|\alpha(\mathbf{x}) - \beta(\mathbf{x})\big|.\end{aligned}$$

Now we choose for $\varepsilon > 0$ an i_0 with $\sum_{j=i_0+1}^\infty 2^{-j} < \varepsilon$, then for any $i \geq i_0$

$$\begin{aligned}|h(\alpha) - h_i(\alpha)| = |h(\alpha) - h(\alpha\,\mathbb{I}_{T_i})| \leq d(\alpha, \alpha\,\mathbb{I}_{T_i}) &= \sum_{j=1}^\infty 2^{-j}\sup_{\mathbf{x}\in T_j}\big|\alpha(\mathbf{x}) - (\alpha\,\mathbb{I}_{T_i})(\mathbf{x})\big| \wedge 1 \\ &= \sum_{j=i+1}^\infty 2^{-j}\left(\sup_{\mathbf{x}\in T_j}\big|\alpha(\mathbf{x})\big| \wedge 1\right) = \sum_{j=i+1}^\infty 2^{-j} < \varepsilon.\end{aligned}$$

This yields, for any $i \geq i_0$

$$\begin{aligned}|\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)| &\leq |\mathbb{E}_\xi h(G_n) - \mathbb{E}_\xi h_i(G_n)| + |\mathbb{E}_\xi h_i(G_n) - \mathbb{E}h_i(G)| + |\mathbb{E}h_i(G) - \mathbb{E}h(G)| \\ &\leq 2\varepsilon + |\mathbb{E}_\xi h_i(G_n) - \mathbb{E}h_i(G)| \\ &\leq 2\varepsilon + \sup_{h\in BL_1(l^\infty(T_i))} |\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)|\end{aligned}$$

and the latter summand converges to 0 in outer probability by assumption. Since $\varepsilon > 0$ was arbitrary the assertion follows.
It remains to show that $\mathbb{E}_\xi h(G_n)^* - \mathbb{E}_\xi h(G)_* \xrightarrow{\mathbb{P}} 0$ for all $h \in BL_1(\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2))$ if and only if $\mathbb{E}_\xi h(G_n|_{T_i})^* - \mathbb{E}_\xi h(G_n|_{T_i})_* \xrightarrow{\mathbb{P}} 0$ for every $h \in BL_1(l^\infty(T_i))$ and every $i \in \mathbb{N}$. This assertion can be proved along similar lines as in the proof of the first equivalence given above. The necessity follows from

$$\mathbb{E}_\xi h(G_n|_{T_i})^* - \mathbb{E}_\xi h(G_n|_{T_i})_* = 2^{i+1}\mathbb{E}_\xi h''(G_n)^* - \mathbb{E}_\xi h''(G_n)_* \xrightarrow{\mathbb{P}} 0,$$

while sufficiency can be concluded from

$$\mathbb{E}_\xi h(G_n)^* - \mathbb{E}_\xi h(G_n)_*$$

$$\le \mathbb{E}_\xi |h(G_n)^* - h_i(G_n)^*| + \mathbb{E}_\xi h_i(G_n)^* - \mathbb{E}_\xi h_i(G_n)_* + \mathbb{E}_\xi |h_i(G_n)_* - h(G_n)_*|$$
$$\le 2\varepsilon + \mathbb{E}_\xi h_i(G_n)^* - \mathbb{E}_\xi h_i(G_n)_* \xrightarrow{\mathbb{P}} 2\varepsilon,$$

where we estimated the first and the last summand in the second line under consideration of Lemma 1.2.2 in Van der Vaart and Wellner (1996). □

Lemma B.3. *Suppose that $g : \mathbb{D}_1 \longrightarrow \mathbb{D}_2$ is a Lipschitz-continuous map between metrized topological vector spaces. If $G_n = G_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} G$ in $\mathbb{D}_1$, where G is tight, then $g(G_n) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} g(G)$ in $\mathbb{D}_2$.*

Proof. Let $c_0 \ge 1$ denote a Lipschitz constant for g. For arbitrary $h \in BL_1(\mathbb{D}_2)$ we have $|hg(\alpha) - hg(\beta)| \le d(g\alpha, g\beta) \le c_0 d(\alpha, \beta)$, which yields that the mapping $h' := c_0^{-1} hg$ lies in $BL_1(\mathbb{D}_1)$ and therefore

$$\sup_{h \in BL_1(\mathbb{D}_2)} |\mathbb{E}_\xi h(g(G_n)) - \mathbb{E} h(g(G))| \le c_0 \sup_{h \in BL_1(\mathbb{D}_1)} |\mathbb{E}_\xi h(G_n) - \mathbb{E} h(G)| \xrightarrow{\mathbb{P}^*} 0.$$

The asymptotic measurability follows along similar lines. □

Lemma B.4. *Let $Y_n = Y_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n)$ and $Z_n = Z_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n)$ be two (bootstrap) statistics in a metric space $(\mathbb{D}, d)$, depending on the data $X_1, \ldots, X_n$ and on some multipliers $\xi_1, \ldots, \xi_n$. If $Y_n \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} Y$ in $\mathbb{D}$, where Y is tight, and $d(Y_n, Z_n) \xrightarrow{\mathbb{P}^*} 0$, then also $Z_n \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} Y$ in $\mathbb{D}$.*

Proof. We only prove (3.5), the assertion about the asymptotic measurability in (3.6) follows along similar lines. Observing the estimate

$$\sup_{h \in BL_1(\mathbb{D})} |\mathbb{E}_\xi h(Z_n) - \mathbb{E} h(Y)| \le \mathbb{E}_\xi \left[d(Y_n, Z_n)^* \wedge 2 \right] + \sup_{h \in BL_1(\mathbb{D})} |\mathbb{E}_\xi h(Y_n) - \mathbb{E} h(Y)|$$

it suffices to show that $\mathbb{E}_\xi \left[d(Y_n, Z_n)^* \wedge 2 \right]$ converges to 0 in outer probability. Now $d(Y_n, Z_n)^* \wedge 2$ is uniformly integrable and converges in probability by assumption, hence it also converges in L^1. We finally use Markov's inequality to obtain $\mathbb{E}_\xi [d(Y_n, Z_n)^* \wedge 2] \xrightarrow{\mathbb{P}^*} 0$, which proves the assertion. □

References

Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statist. Probab. Lett.* in press.

Coles, S. G. and Tawn, J. A. (1994). Statistical methods for multivariate extremes: An application to structural design. *Journal of Applied Statistics*, 43:1–48.

de Haan, L. and Ferreira, A. (2006). *Extreme value theory.* Springer Series in Operations Research and Financial Engineering. Springer, New York.

de Haan, L., Neves, C., and Peng, L. (2008). Parametric tail copula estimation and model testing. *J. Multivar. Anal.*, 99(6):1260–1275.

Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimates of the stable tail dependence functions. *Journal of Multivariate Analysis*, 64:25–47.

Einmahl, H. J., de Haan, L., and Li, D. (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *Annals of Statistics*, 34:1987–2014.

Einmahl, H. J., de Haan, L., and Xin, H. (1993). Estimating a multidimensional extreme-value distribution. *Journal of Multivariate Analysis*, 47:35–47.

Einmahl, H. J., Krajina, A., and Segers, J. (2008). A method of moments estimator of tail dependence. *Bernoulli*, 14:1003–1026.

Embrechts, P., Lindskog, F., and McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In Rachev, S., editor, *Handbook of heavy tailed distributions in finance*, pages 329–384. Elsevier, Amsterdam.

Huang, X. (1992). *Statistics of bivariate extreme values.* PhD thesis, Tinbergen Institute Research Series, Netherlands.

Kojadinovic, I. and Yan, J. (2010). A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing.* in press.

Kojadinovic, I., Yan, J., and Holmes, M. (2010). Fast large sample goodness-of-fit tests for copulas. *Statistica Sinica.* to appear.

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

Malevergne, Y. and Sornette, D. (2004). How to account for extreme co-movements between individual stocks and the market. *J. Risk*, 6:71–116.

Peng, L. and Qi, Y. (2008). Bootstrap approximation of tail dependence function. *Journal of Multivariate Analysis*, 99:1807–1824.

Qi, Y. (1997). Almost sure convergence of the stable tail empirical dependence function in multivariate extreme statistics. *Acta Math. Appl. Sinica (English Ser.)*, 13:167–175.

Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100:377–386.

Schmidt, R. and Stadtmüller, U. (2006). Nonparametric estimation of tail dependence. *Scandinavian Journal of Statistics*, 33:307–335.

Segers, J. (2010). Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions. *arXiv:1012.2133v.*

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

Tiago de Oliveira, J. (1980). Bivariate extremes: Foundations and statistics. In Krishnaiah, P. R., editor, *Multivariate Analysis 5*, pages 349–366. North-Holland, Amsterdam.

Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.

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