

# Multiplier bootstrap of tail copulas with applications

AXEL BÜCHER\* and HOLGER DETTE\*\*

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany.  
E-mail: \*axel.buecher@ruhr-uni-bochum.de; \*\*holger.dette@ruhr-uni-bochum.de

For the problem of estimating lower tail and upper tail copulas, 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 literature on tail copula estimation is so restrictive, such that the tail copula corresponding to tail 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 that 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:* comparison of tail copulas; goodness-of-fit; minimum distance estimation; multiplier bootstrap; stable tail dependence function; tail copula

## 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.* [11]). The function is closely related to tail copulas (see, e.g., Schmidt and Stadtmüller [24]) and represents the current standard to describe extremal dependence (see Embrechts *et al.* [13] and Malevergne and Sornette [20]). The lower and upper tail copulas are defined by

$$\Lambda_L(\mathbf{x}) = \lim_{t \rightarrow \infty} tC(x_1/t, x_2/t), \quad \Lambda_U(\mathbf{x}) = \lim_{t \rightarrow \infty} t\bar{C}(x_1/t, x_2/t), \quad (1.1)$$

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 (unique) copula of the two-dimensional continuous distribution function  $F$ , which relates  $F$  and its marginals  $F_1, F_2$  by

$$F(\mathbf{x}) = C(F_1(x_1), F_2(x_2)) \quad (1.2)$$

(see Sklar [26]), and  $\bar{C}(\mathbf{u}) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2)$  denotes the survival copula of  $\mathbf{X} = (X_1, X_2) \sim F$ . The stable tail dependence function  $l$  and the upper tail copula  $\Lambda_U$  are

associated through 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 [4], Tiago de Oliveira [28], and Einmahl *et al.* [9] imposed restrictions on the marginal distributions to estimate multivariate extreme value distributions. Nonparametric estimates of the stable tail dependence function were investigated in the pioneering thesis of Huang [15] and by Qi [22] and Drees and Huang [7]. Schmidt and Stadtmüller [24] proposed analogous estimates as in Huang [15] for tail copulas [except for rounding deviations due to the fact that  $F_n(F_n^-(x))$ , with the generalized inverse function  $F_n^-$ , is not exactly equal to  $x$ ] and provided new proof of the asymptotic behavior of the estimates. More recent work on inference on the stable tail dependence function was done by Einmahl *et al.* [11] and Einmahl *et al.* [10], 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 [15], Drees and Huang [7], Schmidt and Stadtmüller [24], Einmahl *et al.* [10], de Haan and Ferreira [5], Peng and Qi [21], de Haan *et al.* [6]). However, the (lower or upper) tail copula corresponding to (lower or upper) tail independence is the only tail copula with this property, because the partial derivatives of a tail copula satisfy

$$\partial_1 \Lambda(0, x) = \begin{cases} \lim_{t \rightarrow \infty} \Lambda(1, t) & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases} \quad (1.3)$$

where  $\Lambda$  denotes either  $\Lambda_L$  or  $\Lambda_U$  (see Appendix B for details). Consequently, we provide a result on 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 next 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.

Note that similar investigations were recently carried out by Segers [25] in the context of nonparametric copula estimation. In that paper it is demonstrated that many (even most) of the most popular copula models do not have continuous partial derivatives on the whole unit square, which has been the usual assumption for the asymptotic behavior of the empirical copula process hitherto. Moreover, it is shown how the assumptions can be suitably relaxed such that the asymptotics are not influenced.

The second objective of the present paper is to approximate 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 received much less attention in the literature. Recently, Peng and Qi [21] considered the tail empirical distribution function and showed the consistency of the bootstrap based on resampling (again under the assumption of continuous partial derivatives). They used their results to construct confidence bands for the tail dependence function. Although the 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.* [6]), has very high computational costs because it relies heavily on random number generation and estimation. (See Kojadinovic and Yan [17] and Kojadinovic *et al.* [18] for a more detailed discussion of the computational efficiency of the multiplier bootstrap.) On the other hand, as pointed out by Bücher [1] and Bücher and Dette [2] in the context of nonparametric copula estimation, some multiplier bootstrap procedures lead to more reliable approximations compared with those from 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, called the *partial derivatives multiplier bootstrap*, uses the structure of the limit distribution of the empirical tail copula process. As a consequence, this approach requires estimation of the partial derivatives of the tail copula. The second method, which we call the *direct multiplier bootstrap*, avoids this problem, using 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 discuss the bootstrap approximations in the context of testing parametric assumptions for the tail copula. We defer all proofs and some of the technical details to the [Appendix](#).

## 2. Empirical tail copulas

Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  denote independent identically distributed random variables with distribution function  $F$  and denote the empirical distribution functions of  $F$  and its marginals  $F_1$  and  $F_2$  by  $F_n(\mathbf{x}) = n^{-1} \sum_{i=1}^n \mathbb{I}\{\mathbf{X}_i \leq \mathbf{x}\}$ ,  $F_{n1}(x_1) = F_n(x_1, \infty)$  and  $F_{n2}(\mathbf{x}) = F_n(\infty, x_2)$ , respectively. Analogously, we define the joint empirical survival function by  $\bar{F}_n(\mathbf{x}) = n^{-1} \sum_{i=1}^n \mathbb{I}\{\mathbf{X}_i > \mathbf{x}\}$  and the marginal empirical survival functions as  $\bar{F}_{n1} = 1 - F_{n1}$  and  $\bar{F}_{n2} = 1 - F_{n2}$ . Following Schmidt and Stadtmüller [24], we consider the estimators

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

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

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

Here  $G^-$  and  $\bar{G}^-$  denote the (left-continuous) generalized inverse functions of some real distribution function  $G$  and its corresponding survival function  $\bar{G} = 1 - G$ , defined by

$$G^-(p) := \begin{cases} \inf\{x \in \mathbb{R} | G(x) \geq p\}, & 0 < p \leq 1, \\ \sup\{x \in \mathbb{R} | G(x) = 0\}, & p = 0, \end{cases}$$

$$\bar{G}^-(p) := \begin{cases} \sup\{x \in \mathbb{R} | \bar{G}(x) \geq p\}, & 0 < p \leq 1, \\ \inf\{x \in \mathbb{R} | \bar{G}(x) = 0\}, & p = 0. \end{cases}$$

It is easy to see that the estimators  $\hat{\Lambda}_L$  and  $\hat{\Lambda}_U$  are asymptotically equivalent to the estimates

$$\frac{1}{k} \sum_{i=1}^n \mathbb{I}\{R(X_{i1}) \leq kx_1, R(X_{i2}) \leq kx_2\} = \hat{\Lambda}_L(\mathbf{x}) + O(1/k), \tag{2.2}$$

$$\frac{1}{k} \sum_{i=1}^n \mathbb{I}\{R(X_{i1}) > n - kx_1, R(X_{i2}) > n - kx_2\} = \hat{\Lambda}_U(\mathbf{x}) + O(1/k), \tag{2.3}$$

where  $R(X_{ij}) = nF_{n1}(X_{ij})$  denotes the rank of  $X_{ij}$  among  $X_{1j}, \dots, X_{nj}$  ( $j = 1, 2$ ) (see Huang [15] for an alternative asymptotic equivalent estimator). Therefore, we introduce analogs of (2.2) and (2.3) where the marginals  $F_1$  and  $F_2$  are assumed known, that is,

$$\tilde{\Lambda}_L(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^n \mathbb{I}\left\{F_1(X_{i1}) \leq \frac{kx_1}{n}, F_2(X_{i2}) \leq \frac{kx_2}{n}\right\}, \tag{2.4}$$

$$\tilde{\Lambda}_U(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^n \mathbb{I}\left\{F_1(X_{i1}) > 1 - \frac{kx_1}{n}, F_2(X_{i2}) > 1 - \frac{kx_2}{n}\right\}. \tag{2.5}$$

For the sake of brevity, we restrict our investigations to the case of lower tail copulas. We assume that this function is non-zero in a single point  $\mathbf{x} \in [0, \infty)^2 \setminus \{(0, 0)\}$ , and as a consequence non-zero everywhere on  $[0, \infty)^2$  (see Theorem 1 in Schmidt and Stadtmüller [24]).

Let  $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$  denote the space of all functions  $f: \bar{\mathbb{R}}_+^2 \rightarrow \mathbb{R}$ , which are locally uniformly bounded on every compact subset of  $\bar{\mathbb{R}}_+^2 = [0, \infty]^2 \setminus \{(0, 0)\}$  (i.e., on closed subsets of  $[0, \infty]^2$  that are bounded away from 0), equipped with the 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 by  $T_i = [0, i]^2 \cup [0, i] \times \{\infty\} \cup \{\infty\} \times [0, i]$  and where  $\|f\|_{T_i} = \sup_{\mathbf{x} \in T_i} |f(\mathbf{x})|$  denotes the sup-norm on  $T_i$ . 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 [30]). Throughout this paper,  $\ell^\infty(T)$  denotes the set of uniformly bounded functions on a set  $T$ ,  $\xrightarrow{\mathbb{P}}$  denotes convergence in (outer) probability, and  $\rightsquigarrow$  denotes weak convergence in the sense of Hoffmann-Jørgensen (see, e.g., Van der Vaart and Wellner [30]).

Schmidt and Stadtmüller [24] assumed that the lower tail copula  $\Lambda_L$  satisfies the second-order condition

$$\lim_{t \rightarrow \infty} \frac{\Lambda_L(\mathbf{x}) - tC(x_1/t, x_2/t)}{A(t)} = g(\mathbf{x}) \tag{2.6}$$

locally uniformly for  $\mathbf{x} = (x_1, x_2) \in \bar{\mathbb{R}}_+^2$ , where  $g$  is a non-constant function and the function  $A: [0, \infty) \rightarrow [0, \infty)$  satisfies  $\lim_{t \rightarrow \infty} A(t) = 0$ . A detailed look at the proofs reveals that (2.6) may be weakened to the condition

$$|\Lambda_L(\mathbf{x}) - tC(x_1/t, x_2/t)| = O(A(t)) \quad (2.7)$$

for  $t \rightarrow \infty$ , locally uniformly for  $\mathbf{x} \in \bar{\mathbb{R}}_+^2$ . Under (2.6) and the additional assumptions  $\Lambda_L \not\equiv 0$ ,  $\sqrt{k}A(n/k) \rightarrow 0$ ,  $k = k(n) \rightarrow \infty$ ,  $k = o(n)$ , Schmidt and Stadtmüller [24] showed that the lower tail copula process with known marginals defined by (2.4) converges weakly in  $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$ , that is,

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

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

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

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

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

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

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

The assumption of continuous partial derivatives is made in the literature on estimation of stable tail dependence functions and tail copulas. However, as demonstrated in (1.3), there does not exist any tail copula  $\Lambda_L \not\equiv 0$  with continuous partial derivatives at the origin  $(0, 0)$ . With our first result, we fill this gap and prove weak convergence of the empirical tail copula process under suitable weakened smoothness assumptions. To do so, we use a similar approach as in Schmidt and Stadtmüller [24], because this turns out to be useful for a proof of consistency of the multiplier bootstrap as well. We first consider the case of known marginals. Because of the second-order condition (2.7), the proof of (2.8) can be given by showing weak convergence of the centered statistic

$$\tilde{\alpha}_n(\mathbf{x}) := \sqrt{k} \left( \tilde{\Lambda}_L(\mathbf{x}) - \frac{n}{k} C(x_1 k/n, x_2 k/n) \right). \quad (2.12)$$

**Lemma 2.1.** *If  $\Lambda_L \not\equiv 0$  and the second-order condition (2.7) holds with  $\sqrt{k}A(n/k) \rightarrow 0$ , where  $k = k(n) \rightarrow \infty$  and  $k = o(n)$ , then we have, as  $n$  tends to infinity,*

$$\tilde{\alpha}_n(\mathbf{x}) = \sqrt{k} \left( \tilde{\Lambda}_L(\mathbf{x}) - \frac{n}{k} C(x_1 k/n, x_2 k/n) \right) \rightsquigarrow \mathbb{G}_{\tilde{\Lambda}_L}(\mathbf{x}) \quad (2.13)$$

in  $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$ . Here  $\mathbb{G}_{\hat{\Lambda}_L}$  is a tight centered Gaussian field concentrated on  $\mathcal{C}_\rho(\bar{\mathbb{R}}_+^2)$  with covariance structure given in (2.9) and  $\rho$  is a pseudo-metric on the space  $\bar{\mathbb{R}}_+^2$  defined by

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

$\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2), \mathbf{x} \wedge \mathbf{y} = (x_1 \wedge y_1, x_2 \wedge y_2)$  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  $\rho$ -continuous on every  $T_i$ .

This assertion is proved in Theorem 4 of Schmidt and Stadtmüller [24] by showing convergence of the finite-dimensional distributions and tightness. For an alternative proof based on Donsker classes, see Remark A.2 in the Appendix. For a proof of a corresponding result for the empirical tail copula process with estimated marginals as defined in (2.10), we use the functional delta method in (2.8) with some suitable functional.

**Theorem 2.2.** *Let  $\Lambda_L \not\equiv 0$  be a lower tail copula whose first-order partial derivatives satisfy the condition*

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

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

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

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  $\{\mathbf{x} \in \bar{\mathbb{R}}_+^2 \mid x_p \in \{0, \infty\}\}$ .

Theorem 2.2 has been proven by Schmidt and Stadtmüller [24], Theorem 6, under the additional assumption that the tail copula has continuous partial derivatives. As pointed out earlier, there is no tail copula  $\Lambda_L \not\equiv 0$  with this property. We point out that in the case where  $d = 2$ , it can be shown that all tail copulas with continuous partial derivatives on the interior as required in (2.14) also have continuous partial derivatives on the axes, except for the origin. Therefore, in the two-dimensional case, our condition just makes verification issues easier. Nevertheless, in higher dimensions, the condition (as considered, e.g., Einmahl *et al.* [12]) becomes more meaningful. Note that a careful inspection of our proof in the Appendix reveals that a generalization to higher dimensions is not a trivial extension (especially the proof of Lemma A.1). Nevertheless, we are convinced that such an extension is possible. Recently, Bücher and Volgushev [3] derived a similar extension to the  $d$ -dimensional case for the usual empirical copula process. This program should be transferred to the empirical tail copula process and is deferred to future research.

### 3. Multiplier bootstrap approximation

#### 3.1. Asymptotic theory

In this section we construct multiplier bootstrap approximations of the Gaussian limit distributions  $\mathbb{G}_{\hat{\Lambda}_L}$  and  $\mathbb{G}_{\hat{\Lambda}_L}$  specified in (2.8) and (2.10), respectively. Toward this end, let  $\xi_i$  be inde-

pendent identically distributed positive random variables, independent of the  $\mathbf{X}_i$ , with mean  $\mu$  in  $(0, \infty)$  and finite variance  $\tau^2$ . We first deal with the case of known marginals and define a multiplier bootstrap analog of (2.4) by

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

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

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

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

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

and

$$\mathbf{U}_i = (U_{i1}, U_{i2}); \quad U_{ip} = F_p(X_{ip}) \quad \text{for } p = 1, 2.$$

Throughout this paper, we use the notation

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

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

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

and

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

where

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

denotes the set of all Lipschitz-continuous functions bounded by 1. The subscript  $\xi$  in the expectations indicates conditional expectation over the weights  $\xi = (\xi_1, \dots, \xi_n)$  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  $\xi$ . 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 [29]). The following result shows that the process (3.2) provides a valid bootstrap approximation of the process defined in (2.12).

**Theorem 3.1.** *If  $\Lambda_L \not\equiv 0$  and the second-order condition (2.7) holds with  $\sqrt{k}A(n/k) \rightarrow 0$ ,  $k = k(n) \rightarrow \infty$  and  $k = o(n)$  we have, as  $n$  tends to infinity,*

$$\tilde{\alpha}_n^m = \frac{\mu}{\tau} \sqrt{k} (\tilde{\Lambda}_L^\xi - \tilde{\Lambda}_L) \xrightarrow[\xi]{\mathbb{P}} \mathbb{G}_{\tilde{\Lambda}_L}$$

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

Because 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  $\xi_1, \dots, \xi_n$  and denote them by  $\xi_{1,b}, \dots, \xi_{n,b}$ . Compute the statistics  $\tilde{\alpha}_{n,b}^m = \tilde{\alpha}_n^m(\xi_{1,b}, \dots, \xi_{n,b})$  ( $b = 1, \dots, B$ ) and use the empirical distribution of  $\tilde{\alpha}_{n,1}^m, \dots, \tilde{\alpha}_{n,B}^m$  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, Theorem 3.1 cannot be used in many statistical applications. We have developed two consistent bootstrap approximation for the limiting distribution of the process (2.10) that do not require knowledge of the marginals. Intuitively, it is natural to replace the unknown marginal distributions in (3.1) by their empirical counterparts, that is,

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

which yields the process

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

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

$$\beta_n = \frac{\mu}{\tau} \sqrt{k} (\hat{\Lambda}_L^{\xi, \cdot} - \hat{\Lambda}_L) \xrightarrow[\xi]{\mathbb{P}} \mathbb{G}_{\tilde{\Lambda}_L}$$

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  $\beta_n$  cannot 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  $\beta_n$  can be calculated

from the data without knowing the marginal distributions. Consequently, we obtain an approximation for the unknown distribution of the process  $\mathbb{G}_{\hat{\Lambda}_L}$ . To get an approximation of  $\mathbb{G}_{\hat{\Lambda}_L}$ , we follow Rémillard and Scaillet [23] and estimate the derivatives of the tail copula by

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

where  $\mathbf{e}_p$  denotes the  $p$ th unit vector ( $p = 1, 2$ ) and  $h \sim k^{-1/2}$  tends to 0 with increasing sample size. In the Appendix (see the proof of the following theorem in Appendix A), we show that these estimates are consistent, and thus 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  $\alpha_n^{pdm}$  depends only on the data and the multipliers  $\xi_1, \dots, \xi_n$ . Consequently, a bootstrap sample can be readily generated as described in the previous paragraph; in what follows, we call this method the *partial derivatives multiplier bootstrap (pdm-bootstrap)*. 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 introduces multiplier random variables not only in the two-dimensional distribution function, but also in the inner estimators of the marginals. To be precise, define

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

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

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

and consider the process

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

We call this bootstrap method the *direct multiplier bootstrap (dm bootstrap)*.

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

$$\alpha_n^{dm} = \frac{\mu}{\tau} \sqrt{k} (\hat{\Lambda}_L^{\xi, \xi} - \hat{\Lambda}_L) \underset{\xi}{\overset{\mathbb{P}}{\rightsquigarrow}} \mathbb{G}_{\hat{\Lambda}_L} \quad \text{in } \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2). \tag{3.10}$$

**Remark 3.5.** As pointed out by a referee, an alternative multiplier bootstrap could be obtained by multiplying each summand with  $\xi_i - 1$ , where  $\xi_1, \dots, \xi_n$  are i.i.d. with  $E[\xi_1] = \text{Var}(\xi_1) = 1$  (see Kosorok [19], Section 11.4.2).

### 3.2. Finite-sample results

In this section we present a small comparison of the finite-sample properties of the two bootstrap approximations given in this section. We also study the impact of the choice of the parameter  $k$  on the properties of the estimates and the bootstrap procedure. For the sake of brevity, we only consider data generated from the Clayton copula with a coefficient of lower tail dependence  $\lambda_L = 0.25$ . The Clayton copula, defined by

$$C(\mathbf{u}; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \quad \theta > 0, \tag{3.11}$$

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

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

Tables 1 and 2 show the accuracy of the bootstrap approximation of the covariances of the limiting variable  $\mathbb{G}_{\hat{\Lambda}_L}$ .

We chose three points on the unit circle,  $\{e^{i\varphi}, \varphi = \ell\pi/8 \text{ with } \ell = 1, 2, 3\}$ , and present 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  $\alpha_n$  on the basis of  $5 \cdot 10^5$  simulation runs, with a sample size of  $n = 1000$  and the parameter  $k$  chosen as 50. This choice is motivated by the left panel of Figure 1, which plots the sum of the squared bias, the variance, and the mean squared

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

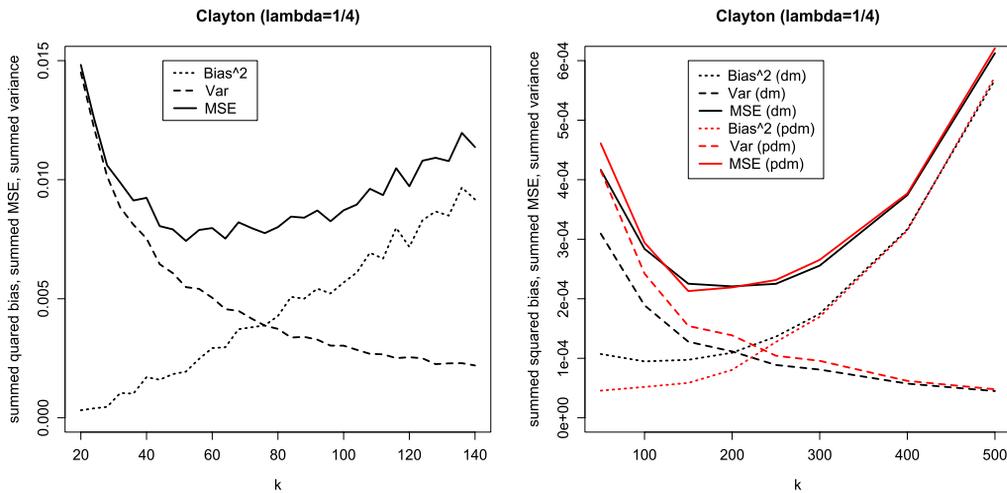
	True			$\alpha_n$		
	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$
$\frac{\pi}{8}$	0.0874	0.0754	0.0516	0.0889	0.0737	0.0476
$2\frac{\pi}{8}$		0.1160	0.0754		0.1218	0.0741
$3\frac{\pi}{8}$			0.0874			0.0892

**Table 2.** Averaged sample covariances (rows 3–5) and  $MSE \times 10^4$  (rows 6–8) of the bootstrap approximations  $\alpha_n^{pdm}$ ,  $\alpha_n^{dm}$  and  $\alpha_n^{res}$  of  $\mathbb{G}_{\hat{\Lambda}_L}$  under the conditions of Table 1

	$\alpha_n^{pdm}$			$\alpha_n^{dm}$			$\alpha_n^{res}$		
	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$	$\frac{\pi}{8}$	$2\frac{\pi}{8}$	$3\frac{\pi}{8}$
$\frac{\pi}{8}$	0.094	0.072	0.046	0.100	0.071	0.045	0.100	0.070	0.043
$2\frac{\pi}{8}$		0.130	0.072		0.136	0.707		0.136	0.070
$3\frac{\pi}{8}$			0.094			0.099			0.099
$\frac{\pi}{8}$	3.67	4.68	3.65	3.86	3.49	2.72	4.21	3.85	3.21
$2\frac{\pi}{8}$		8.11	4.87		8.89	3.25		8.73	3.64
$3\frac{\pi}{8}$			3.70			3.77			3.90

error (MSE) of the estimators  $\hat{\Lambda}_L(e^{i\ell\pi/4})$  for  $\Lambda_L(e^{i\ell\pi/4})$  ( $\ell = 1, 2, 3$ ). The MSE is minimized for values of  $k$  in a neighborhood of the point 50. Note also that the literature provides several data-adaptive proposals for the choice of the parameter  $k$  (see, e.g., Drees and Kaufmann [8] oder Gomes and Oliveira [14]) in the univariate context.

The data in Table 1 serve as a benchmark for the multiplier bootstrap approximations of the covariances stated in Table 2, where we investigate the quality of the approximation by various bootstrap methods. The distribution of the multipliers in the  $dm$  and  $pdm$  bootstrap procedures was chosen according to Bücher and Dette [2] as  $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 0.5$ , such that  $\mu =$



**Figure 1.** Left: Averaged MSE, variance, and squared bias for the estimation of  $\Lambda_L(e^{i\ell\pi/4})$  ( $\ell = 1, 2, 3$ ) by its empirical counterpart  $\hat{\Lambda}_L(e^{i\ell\pi/4})$  against the parameter  $k$ . Right: Averaged MSE, variance, and squared bias for the bootstrap estimation of the covariances of  $\mathbb{G}_{\hat{\Lambda}_L}$ .

$\tau = 1$ . For the sake of completeness, we also investigate the resampling bootstrap considered in Peng and Qi [21], hereinafter denoted by  $\alpha_n^{res}$ . The estimated covariances given in the first part (rows 3–5) of Table 2 were calculated by 1000 simulation runs, where in each run the covariance is estimated on the basis of  $B = 500$  bootstrap replications. The second part (rows 6–8) of Table 2 shows the corresponding MSE.

As the figures show, all bootstrap procedures yield approximations of comparable quality. Considering only the bias in Table 2, the *pdm* bootstrap demonstrates slight advantages in all cases, whereas there are basically no differences between the *dm* and the resampling bootstraps. A comparison of the MSE in Table 2 shows that the *pdm* bootstrap has the best performance on the diagonal. On the other hand, it yields a less accurate approximation for the off-diagonal covariances, for which the *dm* bootstrap yields the best results.

The right panel of Figure 1 illustrates the sensitivity of the accuracy of the estimators for the covariances with respect to the choice of  $k$ . For this purpose, we calculated the sum of the MSE values given in Table 2 (as well as the variance and squared bias) for various choices of  $k$ . As can be seen, the best choices for  $k$  lie in an interval of approximate length 100 around the center  $k = 200$ . Compared with the “best” value  $k = 50$ , for estimating  $\Lambda_L$ , the optimal values for estimating the covariances of  $\mathbb{G}_{\hat{\Lambda}_L}$  are approximately four times larger for both the *pdm* bootstrap and the *dm* bootstrap. This increase may be explained by the fact that the large bias of  $\hat{\Lambda}_L(\mathbf{x})$ ,  $\hat{\Lambda}_L^{\xi, \cdot}(\mathbf{x})$  and  $\hat{\Lambda}_L^{\xi, \xi}(\mathbf{x})$  for estimating  $\Lambda_L(\mathbf{x})$  cancels out if the difference  $\hat{\Lambda}_L^{\xi, \cdot}(\mathbf{x}) - \hat{\Lambda}_L(\mathbf{x})$  or  $\hat{\Lambda}_L^{\xi, \xi}(\mathbf{x}) - \hat{\Lambda}_L(\mathbf{x})$  is calculated. As a result, we may choose larger values of  $k$ , resulting in a notable decay of the variance.

These findings have ambiguous consequences. If we are interested only in the covariances of the limiting variable  $\mathbb{G}_{\hat{\Lambda}_L}$ , then a larger value of  $k$  for the bootstrap is advisable. However, because the bootstrap is not able to capture the true bias of the empirical tail copula process, some care is needed if we are interested in approximations of the whole distribution of  $\sqrt{k}(\hat{\Lambda}_L - \Lambda_L)$  (as is required in Section 4). In this case a careful choice of  $k$  for estimating the tail copula becomes even more important; we are confronted with the strong requirement of a small bias for this estimator. Finally, a comparison of the variance and the bias of the two bootstrap procedures investigated in Figure 1 reveals that the *pdm* bootstrap has a smaller bias, but a slightly larger variance, than the *dm* bootstrap. On the other hand, the differences with respect to the MSE are nearly undetectable.

## 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  $\mathbf{X}_1, \dots, \mathbf{X}_{n_1}$  and  $\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}$  denote two independent samples of i.i.d. random variables (we relax the assumption of independence between the samples later) with continuous cumulative

distribution function  $F = C(F_1, F_2)$  and  $H = D(H_1, H_2)$ , respectively. We assume that for both distributions, the corresponding lower tail copulas, say  $\Lambda_{L,X}$  and  $\Lambda_{L,Y}$ , exist and do not vanish and the tails of the corresponding copulas converge to  $\Lambda_L$  at the rate specified in (2.7). We are interested in a test for the hypothesis

$$\mathcal{H}_0: \Lambda_{L,X} \equiv \Lambda_{L,Y} \quad \text{vs.} \quad \mathcal{H}_1: \Lambda_{L,X} \not\equiv \Lambda_{L,Y}. \tag{4.1}$$

Given the homogeneity of tail copulas, we have  $\Lambda_L(t\mathbf{x}) = t\Lambda_L(\mathbf{x})$  for all  $t > 0, \mathbf{x} \in [0, \infty)^2$ , 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  $\varrho$  is defined by

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

and we use the notation  $\Lambda_{L,X}^\angle(\varphi) = \Lambda_{L,X}(\cos \varphi, \sin \varphi), \Lambda_{L,Y}^\angle(\varphi) = \Lambda_{L,Y}(\cos \varphi, \sin \varphi)$ . Note that the integration in (4.2) over the unit circle with respect to the Euclidian norm is rather a matter of taste. An integration along the unit circle with respect to the sup-norm on  $\mathbb{R}^2$  (i.e., along  $\{\mathbf{x} \in [0, \infty)^2: \max(x_1, x_2) = 1\}$ ) would be possible as well.

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

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

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

$$k_p \rightarrow \infty, \quad k_p = o(n_p) \quad (p = 1, 2) \quad \text{and} \quad k_1/(k_1 + k_2) \rightarrow \lambda \in (0, 1).$$

We assume that the tail copulas  $\Lambda_{L,X}$  and  $\Lambda_{L,Y}$  satisfy a second-order condition as in (2.7) (with  $A$  replaced by  $A_p$ ), and that  $k_p$  is chosen appropriately, that is,  $\sqrt{k_p} A_p(k_p/n_p) = o(1)$ . 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

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

Because the two samples  $X$  and  $Y$  are independent, we obtain, independent of the hypotheses, that

$$\mathcal{E}_n \rightsquigarrow \sqrt{1-\lambda} \mathbb{G}_{\hat{\Lambda}_{L,X}} - \sqrt{\lambda} \mathbb{G}_{\hat{\Lambda}_{L,Y}} =: \mathcal{E}. \tag{4.3}$$

in the metric space  $\mathcal{B}_\infty(\mathbb{R}_+^2)$ , where the stochastically independent two-dimensional centered Gaussian fields  $\mathbb{G}_{\hat{\Lambda}_{L,X}}$  and  $\mathbb{G}_{\hat{\Lambda}_{L,Y}}$  are as 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

$$\mathcal{S}_n \rightsquigarrow_{\mathcal{H}_0} \mathcal{T}, \quad \mathcal{S}_n \xrightarrow{\mathbb{P}}_{\mathcal{H}_1} \infty, \tag{4.4}$$

which shows that a test that rejects the null hypothesis (4.1) for large values of  $\mathcal{T}_n$  is consistent. Note that the latter convergence depends crucially on the assumption that  $\sqrt{k}A(n/k) \rightarrow 0$ . If this assumption does not hold, then a large value of  $\mathcal{S}_n$  could occur even under the null hypothesis, because of the biasedness of  $\hat{\Lambda}_L$ . Thus, as discussed at the end of Section 3.2, the choice of a small  $k$  corresponding to a small bias is of prime importance.

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 consider the *pdm* bootstrap (the extension to the *dm* bootstrap is straightforward) using the definition in equation (3.10) and denote, for any  $b \in \{1, \dots, B\}$ ,  $\xi_{1,b}, \dots, \xi_{n_1,b}, \zeta_{1,b}, \dots, \zeta_{n_2,b}$  i.i.d. nonnegative random variables with mean  $\mu_1$  (resp.,  $\mu_2$ ) and variance  $\tau_1^2$  (resp.,  $\tau_2^2$ ). For each  $b$  and both samples, we compute the bootstrap statistics as given in (3.8), that is,

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

where

$$\begin{aligned} \beta_{X,n_1,b}(\mathbf{x}) &= \frac{\mu_1}{\tau_1} \frac{1}{\sqrt{k_1}} \sum_{i=1}^{n_1} \left( \frac{\xi_{i,b}}{\bar{\xi}_{\cdot, b_{n_1}}} - 1 \right) \mathbb{I}\{X_{i1} \leq F_{n_1}^-(k_1 x_1/n_1), X_{i2} \leq F_{n_1}^-(k_1 x_2/n_1)\}, \\ \beta_{Y,n_2,b}(\mathbf{x}) &= \frac{\mu_2}{\tau_2} \frac{1}{\sqrt{k_2}} \sum_{i=1}^{n_2} \left( \frac{\zeta_{i,b}}{\bar{\zeta}_{\cdot, b_{n_2}}} - 1 \right) \mathbb{I}\{Y_{i1} \leq H_{n_2}^-(k_2 x_1/n_2), Y_{i2} \leq H_{n_2}^-(k_2 x_2/n_2)\}, \end{aligned}$$

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

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

By Theorem 3.3 and Theorem 10.8 in Kosorok [19], it follows that for every  $b \in \{1, \dots, B\}$ ,

$$\widehat{\mathcal{T}}_n^{(pdm,b)} \xrightarrow[\xi]{\mathbb{P}} \mathcal{T}^{(b)},$$

where  $\mathcal{T}^{(b)}$  is an independent copy of  $\mathcal{T}$ . (Note that we consider the processes  $\widehat{\mathcal{E}}_n^{(pdm,b)}$  in the Banach space  $\ell^\infty([0, 1]^2)$ .) Thus, from (4.4), we obtain a consistent asymptotic level  $\alpha$  test for the null hypothesis (4.1) by rejecting  $\mathcal{H}_0$  for large values of  $\mathcal{S}_n$ , that is,

$$\mathcal{S}_n > q_{1-\alpha}^{pdm}, \tag{4.5}$$

where  $q_{1-\alpha}^{pdm}$  denotes the  $(1 - \alpha)$ -quantile of the c.d.f.  $K_n^{pdm}(s) = B^{-1} \sum_{b=1}^B \mathbb{I}\{\widehat{\mathcal{T}}_n^{(pdm,b)} \leq s\}$ .

So far, we have focused our discussion on the case of two independent samples. It is easy to check that our methodology also applies in cases of paired observations, that is,  $\mathbf{X}_i$  is not independent of  $\mathbf{Y}_i$ , but  $n_1 = n_2 = n$ . In that case we must 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,

$$\begin{aligned} C(u_1, u_2) &= \mathcal{C}(u_1, u_2, 1, 1), & D(v_1, v_2) &= \mathcal{C}(1, 1, v_1, v_2), \\ \mathcal{C}_n(u_1, u_2) &= \mathcal{C}_n(u_1, u_2, 1, 1), & D_n(v_1, v_2) &= \mathcal{C}_n(1, 1, v_1, v_2). \end{aligned}$$

If we set  $\Lambda_{L,Z}(\mathbf{x}, \mathbf{y}) = \lim_{t \rightarrow \infty} t \mathcal{C}(\mathbf{x}/t, \mathbf{y}/t)$ ,  $\widehat{\Lambda}_{L,Z}(\mathbf{x}, \mathbf{y}) = \frac{n}{k} \mathcal{C}_n(\frac{n\mathbf{x}}{k}, \frac{n\mathbf{y}}{k})$ , then we obtain

$$\begin{aligned} \Lambda_{L,X}(\mathbf{x}) &= \Lambda_{L,Z}(\mathbf{x}, \infty, \infty), & \Lambda_{L,Y}(\mathbf{y}) &= \Lambda_{L,Z}(\infty, \infty, \mathbf{y}), \\ \widehat{\Lambda}_{L,X}(\mathbf{x}) &= \widehat{\Lambda}_{L,Z}(\mathbf{x}, \infty, \infty), & \widehat{\Lambda}_{L,Y}(\mathbf{y}) &= \widehat{\Lambda}_{L,Z}(\infty, \infty, \mathbf{y}). \end{aligned}$$

Under a second-order condition on the joint tail copula  $\Lambda_{L,Z}$ , the asymptotic properties of the process  $\widehat{\Lambda}_{L,Z}$  can be derived along similar lines as before; we omit the details for the sake of brevity. As the only difference from the preceding discussion, note that the occurring limiting fields  $\mathbb{G}_{\widehat{\Lambda}_{L,X}}$  and  $\mathbb{G}_{\widehat{\Lambda}_{L,Y}}$  are no longer independent. Because the asymptotic behavior of the multiplier bootstrap approximations can be shown to reflect this dependence, we still obtain consistency of the test; we again omit the details.

To investigate the finite-sample property, we consider two independent samples of i.i.d. random variables with Clayton copula (see (3.11)) with a coefficient of lower tail dependence  $\lambda_L$  varying in the set  $\{0.25, 0.5, 0.75\}$ .

**Table 3.** Simulated rejection probabilities of the bootstrap tests defined in (4.5) for the hypothesis (4.1)

$k$	$\lambda_{L,X}$	$\lambda_{L,Y}$	$pdm$			$dm$		
			$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$
50	0.25	0.25	0.143	0.098	0.054	0.125	0.091	0.052
	0.5	0.5	0.140	0.099	0.047	0.108	0.069	0.036
	0.75	0.75	0.117	0.078	0.029	0.068	0.051	0.023
	0.25	0.5	0.764	0.706	0.605	0.713	0.643	0.529
	0.5	0.75	0.896	0.856	0.783	0.869	0.822	0.713
	0.25	0.75	1	1	1	0.999	0.999	0.997
200	0.25	0.25	0.145	0.107	0.052	0.125	0.084	0.044
	0.5	0.5	0.128	0.083	0.037	0.140	0.097	0.051
	0.75	0.75	0.141	0.092	0.041	0.103	0.068	0.035
	0.25	0.5	0.991	0.978	0.948	0.979	0.971	0.950
	0.5	0.75	1	1	1	1	1	1
	0.25	0.75	1	1	1	1	1	1

Table 3 presents the simulated rejection probabilities of the  $pdm$  and  $dm$  bootstrap tests defined in (4.5) for various nominal levels on the basis of 1000 simulation runs. The sample size was  $n_1 = n_2 = n = 1000$  and,  $B = 500$  bootstrap replications with  $\mathcal{U}(\{0, 2\})$  multipliers (i.e.,  $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 0.5$ , such that  $\mu = \tau = 1$ ) were used. The parameter  $k$  was chosen as either  $k = 50$  or  $k = 200$  as suggested by the discussion in the preceding paragraph.

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 tends to be conservative. Of note, the approximation of the nominal level is rather robust with respect to the choice of  $k$ . A comparison of the performance of the two bootstrap procedures shows that the  $dm$  bootstrap test is slightly more conservative, and that this effect increases with the coefficient of tail dependence.

The alternative of different lower tail copulas is detected with reasonable power. Both tests yield rather similar results, with a slight advantage for the  $pdm$  bootstrap. Evaluation of the impact of the choice of the parameter  $k$  under the alternative shows some advantages for  $k = 200$ . This again may be explained by the fact that bias terms cancel out if the difference  $\hat{\Lambda}_{L,X} - \hat{\Lambda}_{L,Y}$  is calculated.

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

In this section we estimate the tail copula of  $\mathbf{X}$  under the additional assumption that it is an element of some parametric class, say  $\mathcal{L} = \{\Lambda_L(\cdot; \theta) | \theta \in \Theta\}$ . Estimation of parametric classes of tail copulas and stable tail dependence functions was recently investigated by de Haan *et al.* [6] and Einmahl *et al.* [11], who proposed a censored likelihood estimator and a moment-based

estimator, respectively. Here we investigate a different, based on the minimum distance method. To be precise, let  $\Lambda_L$  denote an arbitrary lower tail copula and  $\Lambda_L(\cdot; \theta)$  denote an element in the parametric class  $\mathcal{L}$ , and consider the parameter corresponding to the best approximation by the distance  $\varrho$  defined in (4.2)

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

where  $\varrho$  is as defined in (4.2). We call  $\hat{\theta}_n^{\text{MD}} = T(\hat{\Lambda}_L)$  a minimum distance estimator for  $\theta$ , where  $\hat{\Lambda}_L$  is the empirical lower tail copula defined in (2.1). Note that  $\theta_B$  is the “true” parameter if and only if the null hypothesis is satisfied.

Throughout this section, we let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  denote i.i.d. bivariate random variables with c.d.f.  $F = C(F_1, F_2)$  and existing lower tail copula  $\Lambda_L$ . (For a proof of the following result, see Section 4.5 in Bücher [1].)

**Theorem 4.1.** *Suppose that the standard conditions of minimum distance estimation are satisfied. (For a precise formulation of these conditions, see Bücher [1], pages 89ff.) If the true tail copula  $\Lambda_L$  satisfies the first-order condition (2.14) of Theorem 2.2, and if the second-order condition (2.7) holds with  $\sqrt{k}A(n/k) \rightarrow 0$ , where  $k = k(n) \rightarrow \infty$  and  $k = o(n)$ , then the minimum distance estimator  $\hat{\theta}_n^{\text{MD}}$  is consistent for the parameter  $\theta_B$  corresponding to the best approximation with respect to the distance  $\varrho$ . Moreover,*

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

where  $\Lambda_L^{\angle}(\varphi) = \Lambda_L(\cos \varphi, \sin \varphi)$ ,  $\hat{\Lambda}_L^{\angle} = \hat{\Lambda}_L(\cos \varphi, \sin \varphi)$ ,  $\gamma_{\theta_B}(\varphi) = A_{\theta_B}^{-1} \delta_{\theta_B}^{\angle}(\varphi)$ ,  $\delta_{\theta}^{\angle}(\varphi) = \partial_{\theta} \Lambda_L(\cos \varphi, \sin \varphi, \theta)$ ,  $\mathbb{G}_{\hat{\Lambda}_L}^{\angle}(\varphi) = \mathbb{G}_{\hat{\Lambda}_L}(\cos \varphi, \sin \varphi)$ , and

$$A_{\theta_B} := \int \delta_{\theta_B}^{\angle}(\varphi) \delta_{\theta_B}^{\angle}(\varphi)^T + \partial_{\theta} \delta_{\theta_B}^{\angle}(\varphi) (\Lambda_L^{\angle}(\varphi; \theta_B) - \Lambda_L^{\angle}(\varphi)) \, d\varphi,$$

with  $\Lambda_L^{\angle}(\varphi; \theta) = \Lambda_L(\cos \varphi, \sin \varphi; \theta)$ . The limiting variable  $\Theta^{\text{MD}}$  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).

To make use of the latter result in statistical applications, we need 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* bootstraps yield a valid approximation of the distribution of the random variable  $\Theta^{\text{MD}}$ .

**Table 4.** Simulated coverage probability of the confidence intervals based on the pdm bootstrap for  $n = 1000$ . In the last two columns, the parameter  $k$  is chosen as 50 for the estimation of  $\theta$ , whereas  $k$  is chosen as 200 for the bootstrap approximation of  $\mathbb{G}_{\hat{\Lambda}_L}$

$\lambda_L$	$k$					
	50		200		50/200	
	90%	95%	90%	95%	90%	95%
0.25	0.895	0.955	0.014	0.044	0.830	0.915
0.5	0.893	0.936	0.779	0.882	0.888	0.934
0.75	0.838	0.887	0.900	0.949	0.863	0.894

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

$$\Theta_n^{MD,m} := \int \gamma_{\hat{\theta}_n^{MD}}(\varphi) \Gamma_n^{\angle}(\varphi) \, d\varphi \overset{\mathbb{P}}{\rightsquigarrow} \Theta^{MD},$$

where  $\Gamma_n^{\angle}(\varphi) = \Gamma_n(\cos \varphi, \sin \varphi)$ ,  $\gamma_{\hat{\theta}_n^{MD}} = \hat{A}_{\hat{\theta}_n^{MD}}^{-1} \delta_{\hat{\theta}_n^{MD}}^{\angle}(\varphi)$ , and

$$\hat{A}_{\hat{\theta}_n^{MD}} := \int \delta_{\hat{\theta}_n^{MD}}^{\angle}(\varphi) \delta_{\hat{\theta}_n^{MD}}^{\angle}(\varphi)^T + \partial_{\theta} \delta_{\hat{\theta}_n^{MD}}^{\angle}(\varphi) (\Lambda_L^{\angle}(\varphi; \hat{\theta}_n^{MD}) - \hat{\Lambda}_L^{\angle}(\varphi)) \, d\varphi.$$

Based on this result, it is possible to construct asymptotic confidence regions for the parameter  $\theta$ , as well as to test point hypotheses regarding the parameter. Table 4 presents a small simulation study regarding the finite-sample coverage probabilities of some confidence intervals for the parameter of a Clayton tail copula. This interval is defined as  $K I_{1-\alpha} = [\hat{\theta}_n^{MD} - k^{-1/2} \hat{q}_{1-\alpha/2}, \hat{\theta}_n^{MD} + k^{-1/2} \hat{q}_{\alpha/2}]$ , where  $\hat{q}_{\beta}$  denotes the estimated  $\beta$ -quantile of the distribution of  $\Theta_n^{MD}$  based on the bootstrap approximation provided by Theorem 4.2. The sample size is  $n = 1000$ , and  $B = 500$  bootstrap replications are used for calculating the quantiles. All coverage probabilities are calculated by 1000 simulation runs. The parameter of the Clayton tail copula is chosen such that the tail dependence coefficient varies in the set  $\{1/4, 2/4, 3/4\}$ .

To investigate the impact of the choice of  $k$ , we chose three different scenarios:  $k = 50$ ,  $k = 200$ , and two different values of  $k$ , namely  $k = 50$  for the estimator  $\hat{\theta}_n^{MD}$  and  $k = 200$  for the bootstrap estimator of the quantiles  $\hat{q}_{\beta}$ . This choice was motivated by the findings in Section 3.2, which indicate that a smaller value of  $k$  should be used in the estimator  $\hat{\Lambda}_L$ .

The tables reveal that there is no unique ‘‘optimal’’ choice for  $k$ . For  $\lambda_L = 0.25$ , the best results are obtained for the scenario with  $k = 50$ , followed by the scenario with two different values of  $k$ . Compare these findings with the results of Section 3.2. For  $k = 200$ , the large bias of  $\hat{\theta}_n^{MD}$  (compare the left side of Figure 1) demonstrates that the true parameter does not lie in the estimated confidence interval for more than 95% of the repetitions. For stronger tail dependence,  $\lambda_L = 0.5, 0.75$ , the choice  $k = 200$  yields better results, with almost perfect coverage probabilities for  $\lambda_L = 0.75$ . Also note that the case with  $k = 50$  in the estimator  $\hat{\Lambda}_L$  and  $k = 200$  in the

**Table 5.** Relative efficiency of the method-of-moments type estimate to the minimum distance estimate

$\lambda$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E_\lambda$	0.86	0.87	0.89	0.92	0.97	1.03	1.14	1.38	2.11

corresponding bootstrap statistic does not yield improvements with respect to the approximation of the coverage probability compared with the case where  $k = 50$ .

**Remark 4.3.** As pointed out at the beginning of this section, there are two alternative estimators for parametric classes of tail copulas. De Haan *et al.* [6] proposed a censored maximum likelihood estimator and proved weak convergence to a normal distribution, which involves the partial derivatives up to the sixth order of the stable tail dependence function. Einmahl *et al.* [11] proposed a method-of-moments type estimator and proved a similar statement as given in Theorem 4.1 for the minimum distance estimate. Table 5 compares the asymptotic variances of the method-of-moments and minimum distance estimators for the parameter  $\theta$  in the Clayton family chosen such that the coefficient of tail dependence  $\lambda$  varies in the set  $\{0.1, \dots, 0.9\}$ . The calculated values  $E_\lambda$  are defined as

$$E_\lambda = \frac{\text{Asymptotic variance of the minimum distance estimate}}{\text{Asymptotic variance of the moment type estimate}}.$$

(Note that we were not able to obtain the asymptotic variances for the censored maximum likelihood estimator, because of the complicated structure of the limiting distribution.) The method-of-moments estimator requires the specification of a function  $g$ , which was chosen as in Einmahl *et al.* [11] as the indicator of the set  $\{\mathbf{x} \in [0, 1]^2: x_1 + x_2 \leq 1\}$ . We observe that none of the two estimates is globally preferable over the other. For small amounts of tail dependence, the minimum distance estimate performs slightly better, whereas for increasing tail dependence, the method-of-moments type estimator is more qualified from an asymptotic standpoint.

Also of note, the *dm* and *pdm* bootstraps 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.* [6] and Einmahl *et al.* [11]. The main argument for proving consistency is that the limiting distribution of the method-of-moments and minimum distance estimators can be represented in the form  $\Phi(\mathbb{G}_{\hat{\Lambda}_L}, \Lambda_L, \partial\Lambda_L)$  for some appropriate functional  $\Phi$  depending on the method of estimation. Here  $\partial\Lambda_L$  denotes any vector of partial derivatives of  $\Lambda_L$  with respect to its coordinates or the parameter. Given that the functional  $\Phi$  is suitable smooth, the bootstrap approximation is then obtained by  $\Phi(\alpha_n, \hat{\Lambda}_L, \widehat{\partial\Lambda}_L)$ , where  $\alpha_n$  is  $\alpha_n^{pdm}$  of  $\alpha_n^{dm}$  and  $\widehat{\partial\Lambda}_L$  is a consistent estimate of  $\partial\Lambda_L$ .

In what follows, we 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, that is,

$$\mathcal{H}_0: \Lambda_L \in \mathcal{L} = \{\Lambda_L(\cdot, \theta) | \theta \in \Theta\}, \quad \mathcal{H}_1: \Lambda_L \notin \mathcal{L}. \tag{4.7}$$

This problem was also been discussed by de Haan *et al.* [6] and Einmahl *et al.* [11], who proposed a comparison between a nonparametric estimate 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  $\theta_B$ . Although Einmahl *et al.* [11] did not propose any bootstrap approximation, de Haan *et al.* [6] proposed using the parametric bootstrap. However, Kojadinovic and Yan [17] and Kojadinovic *et al.* [18] pointed out that for copula models, approximations based on multiplier bootstraps are computationally more efficient, especially for large sample sizes. We now illustrate how the multiplier bootstrap can be successfully applied to the problem of testing the hypothesis (4.7).

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

$$GOF_n := k_{\mathcal{Q}}(\hat{\Lambda}_L, \Lambda_L(\cdot; \hat{\theta}_n^{\text{MD}})) = k \int (\hat{\Lambda}_L^{\leftarrow}(\varphi) - \Lambda_L^{\leftarrow}(\varphi; \hat{\theta}_n^{\text{MD}}))^2 d\varphi,$$

where  $\hat{\theta}_n^{\text{MD}}$  denotes the minimum distance estimate. If the standard assumptions of minimum distance estimation are satisfied (see page 89 in Bücher [1] for details), then we obtain for the process  $H_n = \sqrt{k}(\hat{\Lambda}_L - \Lambda_L(\cdot; \hat{\theta}_n^{\text{MD}}))$  under the null hypothesis  $\mathcal{H}_0: \Lambda_L = \Lambda_L(\cdot; \theta_B)$

$$\begin{aligned} H_n &= \sqrt{k}(\hat{\Lambda}_L - \Lambda_L - \delta_{\theta}(\hat{\theta}_n^{\text{MD}} - \theta)) + o_{\mathbb{P}}(1) \\ &= \sqrt{k}\left(\hat{\Lambda}_L - \Lambda_L - \delta_{\theta} \int \gamma_{\theta}(\varphi)(\hat{\Lambda}_L^{\leftarrow}(\varphi) - \Lambda_L^{\leftarrow}(\varphi)) d\varphi\right) + o_{\mathbb{P}}(1) \\ &\rightsquigarrow \mathbb{G}_{\hat{\Lambda}_L} - \delta_{\theta} \int \gamma_{\theta}(\varphi)\mathbb{G}_{\hat{\Lambda}_L}^{\leftarrow}(\varphi) d\varphi = \mathbb{G}_{\hat{\Lambda}_L} - \delta_{\theta}\Theta^{\text{MD}}. \end{aligned}$$

Under the alternative hypothesis, we get an additional summand,

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

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

$$GOF_n = \int \{H_n^{\leftarrow}(\varphi)\}^2 d\varphi \rightsquigarrow Z := \int (\mathbb{G}_{\hat{\Lambda}_L}^{\leftarrow}(\varphi) - \delta_{\theta}^{\leftarrow}(\varphi)\Theta^{\text{MD}})^2 d\varphi, \tag{4.8}$$

whereas under the alternative,  $GOF_n = \int \{H_n^{\leftarrow}(\varphi)\}^2 d\varphi \xrightarrow{\mathbb{P}} \infty$ .

The critical values of the test, which rejects the null hypothesis for large values of  $GOF_n$ , can be calculated based on the following theorem. For a proof, see Theorem 4.10 in Bücher [1].

**Table 6.** Simulated rejection probabilities of the *pdm* bootstrap test (4.9) for the hypothesis (4.7). The first three rows represent models from the null hypothesis, and the next eight rows represent alternative models. The sample size is  $n = 1000$ , and  $B = 500$  bootstrap replications were performed.  $\lambda_L$  denotes the lower tail dependence coefficient

Model	$k$					
	50			200		
	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.15$	$\alpha = 0.1$	$\alpha = 0.05$
Clayton ( $\lambda_L = 0.25$ )	0.124	0.087	0.037	0.174	0.108	0.049
Clayton ( $\lambda_L = 0.5$ )	0.097	0.068	0.032	0.117	0.073	0.039
Clayton ( $\lambda_L = 0.75$ )	0.091	0.048	0.018	0.091	0.058	0.024
Convex ( $\lambda_L = 1/12$ )	0.095	0.052	0.017	0.386	0.291	0.179
Convex ( $\lambda_L = 2/12$ )	0.124	0.066	0.029	0.502	0.401	0.253
Convex ( $\lambda_L = 3/12$ )	0.298	0.200	0.088	0.880	0.828	0.700
Aneglog ( $\lambda_L = 0.2$ )	0.119	0.071	0.028	0.257	0.185	0.109
Aneglog ( $\lambda_L = 0.4$ )	0.241	0.174	0.105	0.625	0.534	0.416
Aneglog ( $\lambda_L = 0.6$ )	0.874	0.833	0.732	1.000	1.000	1.000
Mixed ( $\lambda_L = 0.1$ )	0.118	0.069	0.022	0.523	0.424	0.268
Mixed ( $\lambda_L = 0.3$ )	0.148	0.068	0.032	0.405	0.315	0.187

**Theorem 4.5.** If the assumptions of Theorem 4.1 hold and  $\Gamma_n$  denotes either the process  $\alpha_n^{pdm}$  (Theorem 3.3) or  $\alpha_n^{dm}$  (Theorem 3.4) obtained by the *pdm* bootstrap and the *dm* bootstrap, respectively, then, independent of the hypotheses, it holds that

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

Therefore,  $GOF_n^m = \int \{H_n^m \angle(\varphi)\}^2 d\varphi \overset{\mathbb{P}}{\rightsquigarrow}_{\xi} Z$ , where  $Z$  is as defined in (4.8).

To investigate the finite-sample properties of a goodness-of-fit test on the basis of the multiplier bootstrap, Table 6 presents the simulated rejection probabilities of the *pdm* bootstrap test,

$$GOF_n > q_{1-\alpha}^{(pdm)}, \tag{4.9}$$

where  $q_{1-\alpha}^{(pdm)}$  denotes the  $(1 - \alpha)$  quantile of the bootstrap distribution. For the null hypothesis, we considered the family of Clayton tail copulas as the parametric class. In particular, we investigated three scenarios corresponding to a coefficient of tail dependence  $\Lambda_L(1, 1)$  varying in  $\{0.25, 0.5, 0.75\}$ . Under the alternative, we consider three models:

- (1) A convex combination of the independence copula  $\Pi(\mathbf{u}) = u_1 u_2$  and a Clayton copula (with convex parameter  $1/3$ ), such that the tail copula is given by  $\Lambda_L(\mathbf{x}) = 1/3(x_1^{-\theta} + x_2^{-\theta})^{-1/\theta}$ . The parameter  $\theta$  is chosen such that  $\lambda_L = \Lambda_L(1, 1) = 1/3 \times 2^{-1/\theta}$  varies in the set  $\{1/12, 2/12, 3/12\}$ .

(2) The asymmetric negative logistic model (see Joe [16]), defined by

$$\Lambda_L(1 - t, t) = \{(\psi_1(1 - t))^{-\theta} + (\psi_2 t)^{-\theta}\}^{-1/\theta}, \quad t \in [0, 1],$$

with parameters  $\psi_1 = 2/3, \psi_2 = 1$  and  $\theta \in (0, \infty)$  chosen such that  $\lambda_L = \Lambda_L(1, 1)$  varies in the set  $\{0.2, 0.4, 0.6\}$ .

(3) The mixed model (see Tawn [27]), given by

$$\Lambda_L(1 - t, t) = \theta t(1 - t), \quad t \in [0, 1],$$

where the parameter  $\theta \in [0, 1]$  is chosen such that  $\lambda_L = \Lambda_L(1, 1) = \theta/2$  equals 0.1 or 0.3.

The results are based on 1000 simulation runs, a sample size of  $n = 1000$ , and two choices,  $k = 50, 200$ , of the parameter  $k$ . For each scenario, the critical values were calculated by  $B = 500$  bootstrap replications with  $\mathcal{U}(\{0, 2\})$ -multipliers. We observe a reasonable power and approximation of the nominal level in most cases. Under the null hypothesis, the test is conservative, and this effect is increasing with the level of tail dependence. For the mixed model with  $k = 50$ , the power of the test is close to the nominal level. This observation can be explained by the fact that for  $\lambda_L = 0.5$  (which corresponds to the case where  $\theta = 1$ ), the model is exactly the same as the Clayton model with parameter 1; that is, we get close to the null hypothesis with increasing tail dependence. Finally, we note that a choice of larger  $k$  results in substantially better power properties, whereas we noted no notable differences in the quality of the approximation of the nominal level. Again, this may be explained by the fact that bias terms in  $GOF_n$  cancel out when calculating the difference  $H_n = \sqrt{k}(\hat{\Lambda}_L - \Lambda_L(\cdot; \hat{\theta}_n^{MD}))$ . Therefore, we propose using rather large values for  $k$  in applications of the goodness-of-fit test.

## Appendix A: Proofs

### A.1. Proof of Theorem 2.2

Let  $\mathcal{B}_\infty(\mathbb{R}_+)$  denote the set of functions  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  (where  $\mathbb{R}_+ = [0, \infty)$ ) that are uniformly bounded on compact sets (equipped with the topology of uniform convergence on compact sets), and define  $\mathcal{B}_\infty^{I,0}(\mathbb{R}_+)$  as the subset of all nondecreasing functions  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  that satisfy  $f(0+) = 0$  and for which  $\sup \text{ran } f < \infty$  (with  $\text{ran } f$  denoting the range of  $f$ ) implies that there exists a  $x_0$  with  $f(x_0) = \sup \text{ran } f$ . The latter condition implies that the adjusted generalized inverse function, defined by

$$f^-(z) = \begin{cases} \sup\{x \in \mathbb{R}_+ \mid f(x) = 0\}, & z = 0, \\ \inf\{x \in \mathbb{R}_+ \mid f(x) \geq z\}, & 0 < z < \sup \text{ran } f, \\ \inf\{x \in \mathbb{R}_+ \mid f(x) = \sup \text{ran } f\}, & z \geq \sup \text{ran } f, \end{cases}$$

remains in  $\mathcal{B}_\infty(\mathbb{R}_+)$  for every  $f \in \mathcal{B}_\infty^{I,0}(\mathbb{R}_+)$ . Further, set

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

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

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

(see also Schmidt and Stadtmüller [24]). Observing that  $\tilde{\Lambda}_L \in \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2)$ , and that the adjusted generalized inverse of  $\tilde{\Lambda}_L(x, \infty)$  is given by  $\frac{n}{k}F_1(F_{n1}^-(kx/n))$ , we 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 of Van der Vaart and Wellner [30]) and the following Lemma, which is an extension of the result in the proof of Theorem 5 of Schmidt and Stadtmüller [24] to our weaker conditions.

**Lemma A.1.** *Let  $\Lambda_L$  be a lower tail copula whose partial derivatives satisfy the first-order property (2.14) for  $p = 1, 2$ . Then  $\Phi$  is Hadamard-differentiable at  $\Lambda_L$  tangentially to the set*

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

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

$$\Phi'_{\Lambda_L}(\gamma)(\mathbf{x}) = \gamma(\mathbf{x}) - \partial_1 \Lambda_L(\mathbf{x})\gamma(x_1, \infty) - \partial_2 \Lambda_L(\mathbf{x})\gamma(\infty, x_2), \tag{A.1}$$

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

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

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

where  $\mathcal{B}_\infty^{I,0,-}(\mathbb{R}_+)$  denotes the set of all adjusted generalized inverse functions  $f^-$  with  $f \in \mathcal{B}_\infty^{I,0}(\mathbb{R}_+)$ . Now  $\Phi_1$  is Hadamard-differentiable at  $\Lambda_L$  tangentially to  $\mathcal{C}^0(\bar{\mathbb{R}}_+^2)$ , because it is linear and continuous. The second map,  $\Phi_2$ , is Hadamard-differentiable at  $(\Lambda_L, \text{id}_{\mathbb{R}_+}, \text{id}_{\mathbb{R}_+})$  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  $\mathbb{R}_+$  with  $f(0) = 0$  and its derivative at  $(\Lambda_L, \text{id}_{\mathbb{R}_+}, \text{id}_{\mathbb{R}_+})$  in  $(\gamma, f, g)$  is given by  $\Phi'_{2,(\Lambda_L, \text{id}_{\mathbb{R}_+}, \text{id}_{\mathbb{R}_+})}(\gamma, f, g) = (\gamma, -f, -g)$ . The proof follows along similar lines as the proof of Theorem 5 in Schmidt and Stadtmüller [24], page 321, and thus is omitted; we simply note that

$(\text{id}_{\mathbb{R}_+} + t_n f_n)^-(x) > 0$  for all  $x > 0$  is implied by the additional assumption of continuity in 0 for functions in the set  $\mathcal{B}^{I,0}(\mathbb{R}_+)$ . More effort is needed to show that  $\Phi_3$  is Hadamard-differentiable at  $(\Lambda_L, \text{id}_{\mathbb{R}_+}, \text{id}_{\mathbb{R}_+})$  tangentially to  $\mathcal{C}^0(\bar{\mathbb{R}}_+^2) \times \mathcal{C}^0(\mathbb{R}_+) \times \mathcal{C}^0(\mathbb{R}_+)$  with derivative

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

To see this, let  $t_n \rightarrow 0$ ,  $(\gamma_n, f_n, g_n) \in \mathcal{B}_\infty(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty(\mathbb{R}_+) \times \mathcal{B}_\infty(\mathbb{R}_+)$  with  $(\gamma_n, f_n, g_n) \rightarrow (\gamma, f, g) \in \mathcal{C}^0(\bar{\mathbb{R}}_+^2) \times \mathcal{C}^0(\mathbb{R}_+) \times \mathcal{C}^0(\mathbb{R}_+)$  such that  $(\Lambda_L + t_n \gamma_n, \text{id}_{\mathbb{R}_+} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n) \in \mathcal{B}_\infty^{I,0}(\bar{\mathbb{R}}_+^2) \times \mathcal{B}_\infty^{I,0,-}(\mathbb{R}_+) \times \mathcal{B}_\infty^{I,0,-}(\mathbb{R}_+)$ . Now  $\Phi_3$  is linear in its first argument, and we introduce the decomposition

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

where

$$L_{n1} = t_n^{-1}\{\Phi_3(\Lambda_L, \text{id}_{\mathbb{R}_+} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n) - \Phi_3(\Lambda_L, \text{id}_{\mathbb{R}_+}, \text{id}_{\mathbb{R}_+})\},$$

$$L_{n2} = \Phi_3(\gamma_n, \text{id}_{\mathbb{R}_+} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n).$$

By the definition of  $d$ , it suffices to show uniform convergence on the sets  $T_i$ ,  $i \in \mathbb{N}$ . Because  $T_i \subset \bar{\mathbb{R}}_+^2$  is compact,  $(f_n, g_n)$  converges uniformly and  $\gamma$  is uniformly continuous; thus  $L_{n2}$  uniformly converges to  $\gamma$ .

Considering  $L_{n1}$ , we split the investigation into six different cases, depending on the position of  $\mathbf{x} \in T_i$ . First, let  $\mathbf{x} \in (0, i]^2$ . 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

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

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  $\partial_1 \Lambda_L(\mathbf{x})f(x_1) + \partial_2 \Lambda_L(\mathbf{x})g(x_2)$ ; thus 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  $[0, i]$ , as well as the fact that  $f(0) = 0$ , allows us to choose a  $\delta > 0$  such that  $|f_n(x_1)| < \varepsilon$  for all  $x_1 < \delta$ . Because 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  $[\delta, i] \times (0, i]$ , the partial derivative  $\partial_1 \Lambda_L$  is uniformly continuous, which yields the desired convergence under consideration of  $\mathbf{y}(n) \rightarrow \mathbf{x}$  and boundedness of  $f$ . The same arguments apply for the second derivative, and the case  $\mathbf{x} \in (0, i]^2$  is finished.

Now consider the case  $\mathbf{x} \in (0, i] \times \{0\}$ . By Lipschitz continuity of  $\Lambda_L$  on  $\mathbb{R}_+^2$  (see Theorem 1 in Schmidt and Stadtmüller [24]), we get

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

Because  $\partial_1 \Lambda_L(x_1, 0)f(x_1) + \partial_2 \Lambda_L(x_1, 0)g(0) = 0$ , this yields the assertion. The arguments are similar for the cases  $\mathbf{x} = (0, 0)^T$  and  $\mathbf{x} \in \{0\} \times (0, i]$ , and we proceed with  $\mathbf{x} \in [0, i] \times \{\infty\}$  (and analogously  $\mathbf{x} \in \{\infty\} \times [0, i]$ )

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

This yields the assertion by  $\partial_1 \Lambda_L(x_1, \infty) = 1$  and  $\partial_2 \Lambda_L(x_1, \infty) = 0$ . To conclude,  $\Phi_3$  is Hadamard-differentiable as asserted.

An application of the chain rule (see Lemma 3.9.3 in Van der Vaart and Wellner [30]) completes the proof of the lemma. □

### A.2. Proof of Theorem 3.1

In light of Lemma C.2 in Appendix B (an analog of Theorem 1.6.1 in [30] for the case of conditional weak convergence), the proof of conditional weak convergence of  $\tilde{\alpha}_n^m$  in  $\mathcal{B}_\infty(\mathbb{R}_+^2)$  can be given for each  $\ell^\infty(T_i)$  separately. For brevity, we suppress the index  $i$  and write  $T = T_i$ . Recalling the notation of  $f_{n,\mathbf{x}}(\mathbf{U}_i)$  in (3.3) we can express  $\tilde{\alpha}_n^m$  as

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

and the assertion now follows by an application of Theorem 11.23 in Kosorok [19]. 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 let

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

denote a corresponding sequence of envelopes of  $\mathcal{F}_n$ . We must prove the following:

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

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

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

- (ii) The limit  $H(\mathbf{x}, \mathbf{y}) = \lim_{n \rightarrow \infty} \mathbb{E}[\tilde{\alpha}_n(\mathbf{x})\tilde{\alpha}_n(\mathbf{y})]$  exists for every  $\mathbf{x}$  and  $\mathbf{y}$  in  $T$ .
- (iii)  $\limsup_{n \rightarrow \infty} \mathbb{E}F_n^2(\mathbf{U}_1) < \infty$
- (iv)  $\lim_{n \rightarrow \infty} \mathbb{E}F_n^2(\mathbf{U}_1) \mathbb{I}\{F_n(\mathbf{U}_1) > \varepsilon\sqrt{n}\} = 0$  for all  $\varepsilon > 0$ .
- (v)  $\lim_{n \rightarrow \infty} \rho_n(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x}, \mathbf{y})$  for all  $\mathbf{x}, \mathbf{y} \in T$ , where

$$\rho_n(\mathbf{x}, \mathbf{y}) = (\mathbb{E}(f_{n,\mathbf{x}}(U_1) - f_{n,\mathbf{y}}(U_1))^2)^{1/2}. \tag{A.3}$$

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

- (vi) The sequence  $\mathcal{F}_n$  of classes is almost measurable Suslin (AMS), that is, 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)  $f_{n,\cdot} : [0, 1]^2 \times T_n \rightarrow \mathbb{R}$  is  $\mathcal{B}|_{[0,1]^2} \times \mathcal{B}_n$ -measurable for  $i = 1, \dots, n$ .

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

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

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

$$\begin{aligned}
 F_n^{(1)}(\mathbf{U}_i) &= \sqrt{\frac{n}{k}} \mathbb{I}(U_{i1} \leq ki/n), \\
 F_n^{(2)}(\mathbf{U}_i) &= \sqrt{\frac{n}{k}} \mathbb{I}(U_{i2} \leq ki/n), \\
 F_n^{(3)}(\mathbf{U}_i) &= \sqrt{\frac{n}{k}} \mathbb{I}(U_{i1} \leq ki/n, U_{i2} \leq ki/n),
 \end{aligned}$$

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.2), then the condition also holds for  $(\mathcal{F}_n, F_n)$  by Lemma C.1 in the Appendix, and thus the assertion in (i) is proved. We consider only the (hardest) case of  $\mathcal{F}_n^{(3)}$ . Note that  $\mathcal{F}_n^{(3)} = \{f_{n,\mathbf{x}}, \mathbf{x} \in [0, i]^2\} = \mathcal{G}_n^{(1)} \cdot \mathcal{G}_n^{(2)}$ , where

$$\begin{aligned}
 f_{n,\mathbf{x}} &= (n/k)^{1/2} \mathbb{I}\{U_{i1} \leq kx_1/n, U_{i2} \leq kx_2/n\}, \\
 \mathcal{G}_n^{(j)} &= \{g_{n,t} = (n/k)^{1/4} \mathbb{I}\{U_{ij} \leq kt/n\} | t \in [0, i]\}
 \end{aligned}$$

for  $j = 1, 2$ . Because the functions  $g_{n,t}$  are increasing in  $t$ , the  $\mathcal{G}_n^{(j)}$  are VC classes with VC index 2. Thus, by Lemma 11.21 in Kosorok [19], both classes satisfy the bounded uniform integral entropy condition (A.2). Proposition 11.22 in Kosorok [19] 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  $\mathbb{E}[\tilde{\alpha}_n(\mathbf{x})\tilde{\alpha}_n(\mathbf{y})] = n/k(C(\frac{\mathbf{x} \wedge \mathbf{y}}{n}) - C(\frac{\mathbf{x}}{n})C(\frac{\mathbf{y}}{n}))$ , which converges to  $\Lambda_L(\mathbf{x} \wedge \mathbf{y}) =: H(\mathbf{x}, \mathbf{y})$ , because  $\frac{n}{k} C(\frac{\mathbf{x}}{n})C(\frac{\mathbf{y}}{n}) \rightarrow 0$ .

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

$$\begin{aligned} \mathbb{E}F_n^2(\mathbf{U}_1)\mathbb{I}\{F_n(\mathbf{U}_1) > \varepsilon\sqrt{n}\} &= \int_{\{F_n(\mathbf{U}_1) > \varepsilon\sqrt{n}\}} F_n^2(\mathbf{U}_1) d\mathbb{P} \\ &\leq \frac{n}{k}\mathbb{P}\left(\frac{1}{k}\mathbb{I}\{U_{11} \leq ki/n \text{ or } U_{12} \leq ki/n\} > \varepsilon\right) = 0 \end{aligned}$$

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

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

In light of Theorem 1 in Schmidt and Stadtmüller [24], 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 [19] and show separability of  $\mathcal{F}_n$ ; that is, for every  $n \geq 1$ , 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  $T_n = (\mathbb{Q} \cap [0, i] \times \{\infty\}) \cup (\{\infty\} \times \mathbb{Q} \cap [0, i]) \cup (\mathbb{Q}^2 \cap [0, i]^2)$ . We then have (note that the functions  $f_{n,\mathbf{x}}$  are built by indicators) that for every  $\omega$  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 Theorem 3.1 is complete.

**Remark A.2.** Given that

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

and that in Section A.2 we provided the sufficient conditions for an application of Theorem 11.20 in Kosorok [19], we obtain an alternative proof of Lemma 2.1.

### A.3. Proof of Theorem 3.4

For technical reasons, we give a proof of Theorem 3.4 in advance of the proofs of Theorems 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 [19]). Because this result holds only for Banach space-valued stochastic processes, some adjustments must 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.

Given Lemma 2.1 and Theorem 3.1, we have

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

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, we 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.1,  $\Phi$  is Hadamard-differentiable on  $\mathcal{B}_\infty^I(\bar{\mathbb{R}}_+^2)$  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 [19] can be applied in the present context.

A careful inspection of the proof of Theorem 12.1 in Kosorok [19] 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, the mapping  $\Phi'_{\Lambda_L}$  needs to be extended to the whole space  $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$ , which is possible by using equation (A.1) as the defining identity. Second, the proof of Theorem 12.1 in Kosorok [19] 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 [30], which provides a functional delta method that holds in general metrizable topological vector spaces. Finally, the proof of Theorem 12.1 in Kosorok [19] makes use of a bootstrap continuous mapping theorem (see Theorem 10.8 in Kosorok [19]), which yields

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

In our specific context, this statement follows immediately from the Lipschitz continuity of the derivative  $\Phi'_{\Lambda_L}$  and an application of Lemma C.3 in Appendix C.

### A.4. 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  $\Phi_3$  and  $\Phi_2$  are defined in the proof of Lemma A.1 and  $\Psi_1$  is given by

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

Note that we obtain  $\Psi(\Lambda_L, \Lambda_L) = \Lambda_L$ ,  $\Psi(\tilde{\Lambda}_L, \tilde{\Lambda}_L) = \hat{\Lambda}_L$  and  $\Psi(\tilde{\Lambda}_L^\xi, \tilde{\Lambda}_L) = \hat{\Lambda}_L^{\xi, \cdot}$  ( $\mathbb{P}$ -almost surely). Clearly,  $\Psi_1$  is Hadamard-differentiable at  $(\Lambda_L, \Lambda_L)$ , because it is linear and continuous.  $\Phi_2$  and  $\Phi_3$  are Hadamard-differentiable tangentially to suitable subspaces as well (see the proof of Lemma A.1). By an application of the chain rule (see Lemma 3.9.3 in [30]), we can conclude that  $\Psi$  is Hadamard-differentiable  $(\Lambda_L, \Lambda_L)$  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  $\sqrt{k}((\tilde{\Lambda}_L, \tilde{\Lambda}_L) - (\Lambda_L, \Lambda_L))$  and  $\frac{\mu}{\tau}\sqrt{k}((\tilde{\Lambda}_L^\xi, \tilde{\Lambda}_L) - (\tilde{\Lambda}_L, \tilde{\Lambda}_L))$  toward the same limiting field, which is necessary for an application of the functional delta method to the bootstrap (see, e.g., Theorem 12.1 in Kosorok [19]). Nevertheless, we can mimic certain steps in the proof of this theorem to obtain the result. To be precise, note that, by arguments analogous to those on page 236 of Kosorok [19], we obtain 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 [30]) yield

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

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

$$d(c\sqrt{k}(\Psi(\tilde{\Lambda}_L^\xi, \tilde{\Lambda}_L) - \Psi(\tilde{\Lambda}_L, \tilde{\Lambda}_L)), c\sqrt{k}(\tilde{\Lambda}_L^\xi - \tilde{\Lambda}_L)) \longrightarrow 0$$

in outer probability and thus by boundedness of the metric  $d$  also in outer expectation. Since  $c\sqrt{k}(\tilde{\Lambda}_L^\xi - \tilde{\Lambda}_L) \xrightarrow{\mathbb{P}}_{\xi} \mathbb{G}_1$  we obtain the assertion by Lemma C.4.

### A.5. Proof of Theorem 3.3

Again, write  $T = T_i$ . We start the proof with an assertion regarding consistency of  $\widehat{\partial_p \Lambda_L}$  and claim that for any  $\delta \in (0, 1)$ ,

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

in outer probability. For the proof of (A.5), split  $T$  into three subsets as indicated by its definition, and then proceed as for the proof of Lemma 4.1 in [25]. The details are omitted. Regarding the assertion of the Theorem, we set

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

In light of Lemma C.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 need to show uniform convergence on the set  $T$ . Because  $|\alpha_n^{pdm} - \bar{\alpha}_n^{pdm}| \leq D_{n1} + D_{n2}$ , where

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

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

$$\mathbb{P}^* \left( \sup_{\mathbf{x} \in [0, i]^2} D_{n1}(\mathbf{x}) > \varepsilon \right) \leq \mathbb{P}^* \left( \sup_{\mathbf{x} \in [0, i]^2, x_1 \geq \delta} D_{n1}(\mathbf{x}) > \varepsilon \right) + \mathbb{P}^* \left( \sup_{\mathbf{x} \in [0, i]^2, x_1 < \delta} D_{n1}(\mathbf{x}) > \varepsilon \right). \quad (\text{A.6})$$

Because  $\widehat{\partial_1 \Lambda_L}$  is uniformly consistent on  $\{\mathbf{x} \in [0, i]^2 | x_1 \geq \delta\}$ , and because  $\beta_n$  is asymptotically tight in  $\ell^\infty(T)$  ( $\beta_n$  converges unconditionally by the results in Chapter 10 of [19]), the first probability on the right-hand side converges to 0.

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

$$\sup_{\mathbf{x} \in [0, i]^2} |\widehat{\partial_1 \Lambda_L}(\mathbf{x})| \leq \sup_{\mathbf{x} \in [0, i]^2, x_1 \geq h} \frac{\lceil k(x_1 + h) \rceil - \lceil k(x_1 - h) \rceil}{2kh} \leq 1 + \frac{1}{2kh} \leq 2$$

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

$$\mathbb{P}^* \left( \sup_{\mathbf{x} \in [0, i]^2, x_1 < \delta} |\beta_n(\mathbf{x})| > \varepsilon/3 \right)$$

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

$$\mathbb{P} \left( \sup_{\mathbf{x} \in [0, i]^2, x_1 < \delta} |\mathbb{G}_{\bar{\Lambda}_L}(\mathbf{x})| > \varepsilon/3 \right).$$

Because  $\mathbb{G}_{\bar{\Lambda}_L}$  has continuous trajectories and  $\mathbb{G}_{\bar{\Lambda}_L}(0, x_2) = 0$  (almost surely), this probability can be made arbitrary small by choosing  $\delta$  sufficiently small. The case  $\mathbf{x} \in [0, i]^2$  is completed. For  $\mathbf{x} \in [0, i] \times \{\infty\}$ , the arguments are similar, whereas for  $\mathbf{x} \in \{\infty\} \times [0, i]$ , we have  $D_{n1} = 0$ , and nothing needs to be shown. To conclude,  $\sup_{\mathbf{x} \in T} D_{n1}(\mathbf{x})$  converges to 0 in outer probability, and because the term  $\sup_{\mathbf{x} \in T} D_{n2}$  can be treated similarly, the proof is complete.

## Appendix B: Partial derivatives of tail copulas

**Proposition B.1.** *The first partial derivative of a (lower or upper) tail copula  $\Lambda$  satisfies*

$$\partial_1 \Lambda(0, x) = \begin{cases} \lim_{t \rightarrow \infty} \Lambda(1, t) & \text{if } x \in (0, \infty), \\ 0 & \text{if } x = 0. \end{cases}$$

Consequently, the only tail copula that admits for continuous partial derivatives in the origin is the tail copula corresponding to tail independence, that is,  $\Lambda \equiv 0$  for either the lower or the upper tail.

**Proof.** By the groundedness and homogeneity of  $\Lambda$  (see Theorem 1 in Schmidt and Stadtmüller [24]), we have

$$\partial_1 \Lambda(0, x) = \lim_{h \rightarrow 0} \frac{\Lambda(h, x) - \Lambda(0, x)}{h} = \lim_{h \rightarrow 0} \Lambda(1, x/h) = \lim_{t \rightarrow \infty} \Lambda(1, t)$$

for all  $x \in (0, \infty)$ . Similarly,  $\partial_1 \Lambda(0, 0) = 0$ . The addendum follows by Theorem 1(iv) in Schmidt and Stadtmüller [24].  $\square$

As an example, note that for the Clayton copula given in (3.11), we obtain  $\partial_1 \Lambda_L(0, x) = 1$  for all  $\theta > 0$ .

### Appendix C: Auxiliary results

Here we present several technical details. We omit the proofs of the assertions and refer the reader to Bücher [1], pages 103ff.

**Lemma C.1.** *Suppose that  $\mathcal{G}_n$  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 entropy condition as stated in (A.2). Then the bounded uniform entropy integral condition (A.2) also holds for  $\mathcal{F}_n = \mathcal{G}_n \cup \mathcal{H}_n$ , with envelopes  $F_n = G_n \vee H_n$ .*

**Lemma C.2.** *Suppose that  $G_n = G_n(\mathbf{X}_1, \dots, \mathbf{X}_n, \xi_1, \dots, \xi_n)$  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 [30] holds – namely,  $G_n \overset{\mathbb{P}}{\rightsquigarrow}_\xi G$  in  $\mathcal{B}_\infty(\bar{\mathbb{R}}_+^2)$  is equivalent to  $G_n \overset{\mathbb{P}}{\rightsquigarrow}_\xi G$  in  $\ell^\infty(T_i)$  for every  $i \in \mathbb{N}$ .*

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

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

### Acknowledgements

The authors thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. They also thank two unknown referees and the associate editor for their constructive

comments on an earlier version of this manuscript, John Einmahl for pointing out important references on the subject, and Johan Segers for detailed discussions. This work was supported by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823) of the German Research Foundation (DFG).

## References

- [1] Bücher, A. (2011). Statistical inference for copulas and extremes. Ph.D. thesis, Ruhr-Univ. Bochum, Germany.
- [2] Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statist. Probab. Lett.* **80** 1925–1932. [MR2734261](#)
- [3] Bücher, A. and Volgushev, S. (2011). Empirical and sequential empirical copula processes under serial dependence. Preprint. Available at [arXiv:1111.2778](#).
- [4] Coles, S.G. and Tawn, J.A. (1994). Statistical methods for multivariate extremes: An application to structural design. *J. Appl. Stat.* **43** 1–48.
- [5] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering*. New York: Springer. [MR2234156](#)
- [6] de Haan, L., Neves, C. and Peng, L. (2008). Parametric tail copula estimation and model testing. *J. Multivariate Anal.* **99** 1260–1275. [MR2419346](#)
- [7] Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal.* **64** 25–47. [MR1619974](#)
- [8] Drees, H. and Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Process. Appl.* **75** 149–172. [MR1632189](#)
- [9] Einmahl, J.H.J., de Haan, L. and Huang, X. (1993). Estimating a multidimensional extreme-value distribution. *J. Multivariate Anal.* **47** 35–47. [MR1239104](#)
- [10] Einmahl, J.H.J., de Haan, L. and Li, D. (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *Ann. Statist.* **34** 1987–2014. [MR2283724](#)
- [11] Einmahl, J.H.J., Krajina, A. and Segers, J. (2008). A method of moments estimator of tail dependence. *Bernoulli* **14** 1003–1026. [MR2543584](#)
- [12] Einmahl, J.H.J., Krajina, A. and Segers, J. (2011). An M-estimator for tail dependence in arbitrary dimensions. Discussion Paper 2011-013, CentER. Available at <http://ssrn.com/abstract=1760022>.
- [13] Embrechts, P., Lindskog, F. and McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance* (S. Rachev, ed.) 329–384. Amsterdam: Elsevier.
- [14] Gomes, M.I. and Oliveira, O. (2001). The bootstrap methodology in statistics of extremes – choice of the optimal sample fraction. *Extremes* **4** 331–358 (2002). [MR1924234](#)
- [15] Huang, X. (1992). Statistics of bivariate extreme values. Ph.D. thesis, Tinbergen Institute Research Series, Netherlands.
- [16] Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statist. Probab. Lett.* **9** 75–81. [MR1035994](#)
- [17] Kojadinovic, I. and Yan, J. (2011). A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Stat. Comput.* **21** 17–30. [MR2746600](#)
- [18] Kojadinovic, I., Yan, J. and Holmes, M. (2011). Fast large-sample goodness-of-fit tests for copulas. *Statist. Sinica* **21** 841–871. [MR2829858](#)
- [19] Kosorok, M.R. (2008). *Introduction to Empirical Processes and Semiparametric Inference. Springer Series in Statistics*. New York: Springer. [MR2724368](#)

- [20] Malevergne, Y. and Sornette, D. (2004). How to account for extreme co-movements between individual stocks and the market. *J. Risk* **6** 71–116.
- [21] Peng, L. and Qi, Y. (2008). Bootstrap approximation of tail dependence function. *J. Multivariate Anal.* **99** 1807–1824. [MR2444820](#)
- [22] Qi, Y. (1997). Almost sure convergence of the stable tail empirical dependence function in multivariate extreme statistics. *Acta Math. Appl. Sin. Engl. Ser.* **13** 167–175. [MR1443948](#)
- [23] Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *J. Multivariate Anal.* **100** 377–386. [MR2483426](#)
- [24] Schmidt, R. and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence. *Scand. J. Stat.* **33** 307–335. [MR2279645](#)
- [25] Segers, J. (2011). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. Preprint. Available at arXiv:[1012.2133](#).
- [26] Sklar, M. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8** 229–231. [MR0125600](#)
- [27] Tawn, J.A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika* **75** 397–415. [MR0967580](#)
- [28] Tiago de Oliveira, J. (1980). Bivariate extremes: Foundations and statistics. In *Multivariate Analysis, V (Proc. Fifth Internat. Sympos., Univ. Pittsburgh, Pittsburgh, Pa., 1978)* (P.R. Krishnaiah, ed.) 349–366. Amsterdam: North-Holland. [MR0566351](#)
- [29] van der Vaart, A.W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. [MR1652247](#)
- [30] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics. Springer Series in Statistics*. New York: Springer. [MR1385671](#)

Received January 2011 and revised January 2012