Multiplier bootstrap of tail copulas - with applications

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

January 9, 2013

Abstract

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

For these investigations we demonstrate that the common assumption of the existence of continuous partial derivatives in the the literature on tail copula estimation is so restrictive, such that the tail copula corresponding to 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 which are satisfied for many commonly used models. These results are applied in several statistical problems including minimum distance estimation and goodness-of-fit testing.

Keywords and Phrases: tail copula, stable tail dependence function, multiplier bootstrap, minimum distance estimation, comparison of tail copulas, goodness-of-fit

AMS Subject Classification: Primary 62G32 ; secondary 62G20

1 Introduction

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

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

1

arXiv:1102.0110v2 [math.ST] 8 Sep 2011
provided that the limits exist. Here $x = (x_1, x_2) \in \mathbb{R}_+^2 := [0, \infty)^2 \setminus \{(\infty, \infty)\}$, $C$ denotes the (unique) copula of the two-dimensional continuous distribution function $F$, which relates $F$ and its marginals $F_1, F_2$ by

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

[see Sklar (1959)], and $C(u) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2) = \mathbb{P}(F_1(X_1) > 1 - u_1, F_2(X_2) > 1 - u_2)$ denotes the survival copula of $X = (X_1, X_2) \sim F$. The stable tail dependence function $l$ and the upper tail copula $\Lambda_U$ are associated by the relationship

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

Since its introduction various parametric and nonparametric estimates of the tail copulas and of the stable tail dependence function have been proposed in the literature. Several authors assume that the dependence function belongs to some parametric family. Coles and Tawn (1994), Tiago de Oliveira (1980) or Einmahl et al. (1993) imposed restrictions on the marginal distributions to estimate multivariate extreme value distributions. Nonparametric estimates of the stable tail dependence function have been investigated in the pioneering thesis of Huang (1992) and by Qi (1997) and Drees and Huang (1998). Schmidt and Stadtmüller (2006) proposed analogous estimates as in Huang (1992) 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 gave a new proof for the asymptotic behavior of the estimates. More recent work on inference on the stable tail dependence function can be found in Einmahl et al. (2008) and Einmahl et al. (2006), who investigated moment estimators of tail dependence and weighted approximations of tail copula processes, respectively.

The present paper has two main purposes. First we clarify some curiosities in the literature on tail copula estimation, which stem from the fact that most authors assume the existence of continuous partial derivatives of the tail copula [see e.g. Huang (1992), Drees and Huang (1998), Schmidt and Stadtmüller (2006), Einmahl et al. (2006), de Haan and Ferreira (2006), Peng and Qi (2008) or de Haan et al. (2008) among others]. However, the (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

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

where $\Lambda$ denotes either $\Lambda_L$ or $\Lambda_U$, see appendix B for details. As a consequence we provide a result regarding the weak convergence of the empirical tail copula process (and thus also of the empirical stable tail dependence function) under weak smoothness assumptions (see Theorem 2.2 in the following section). The smoothness conditions are nonrestrictive in the sense, that in the case where they are not satisfied, the candidate limiting process does not have continuous trajectories.

Note that similar investigations have recently been carried out by Segers (2011) 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 paper at hand is devoted to the approximation of the distribution of estimators for the tail copulas by new bootstrap methods. In contrast to the problem of estimation of
the stable dependence function and tail copulas, this problem has found much less attention in the literature. Recently, Peng and Qi (2008) considered the tail empirical distribution function and showed the consistency of the bootstrap based on resampling (again under the assumption of continuous partial derivatives). These results were used to construct confidence bands for the tail dependence function. While these authors considered the naive bootstrap, the present paper is devoted to multiplier bootstrap procedures for tail copula estimation. On the one hand, our research is motivated by the observation that the parametric bootstrap, which is commonly applied in goodness-of-fit testing problems [see de Haan et al. (2008)], has very high computational costs, because it heavily relies on random number generation and estimation [see also Kojadinovic and Yan (2010) and Kojadinovic et al. (2010) for a more detailed discussion of the computational efficiency of the multiplier bootstrap]. On the other hand, it was pointed out by Bücher (2011) and Bücher and Dette (2010) in the context of nonparametric copula estimation that some multiplier bootstrap procedures lead to more reliable approximations than the bootstrap based on resampling.

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

2 Empirical tail copulas

Let \( X_1, \ldots, X_n \) 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 \( F_n(x) = n^{-1} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x) \), \( F_{n1}(x_1) = F_n(x_1, \infty) \) and \( F_{n2}(x) = F_n(\infty, x_2) \), respectively. Analogously, we define the joint empirical survival function by \( \bar{F}_n(x) = n^{-1} \sum_{i=1}^{n} \mathbb{I}(X_i > x) \) and the marginal empirical survival functions as \( \bar{F}_{n1} = 1 - F_{n1} \) and \( \bar{F}_{n2} = 1 - F_{n2} \). Following Schmidt and Stadtmüller (2006) we consider the estimators

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

(2.1)

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

\[
C_n(u) = F_n(F_{n1}^{-1}(u_1), F_{n2}^{-1}(u_2)), \quad \bar{C}_n(u) = \bar{F}_n(F_{n1}^{-1}(u_1), F_{n2}^{-1}(u_2)).
\]
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

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

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

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

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

where $R(X_{ij}) = nF_{ij}(X_{ij})$ denotes the rank of $X_{ij}$ among $X_{1j}, \ldots, X_{nj}$ $(j = 1, 2)$, see also Huang (1992) 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 to be known, that is

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

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

For the sake of brevity we restrict our investigations to the case of lower tail copulas and we assume that this function is non-zero in a single point $x \in [0, \infty)^2$ and as a consequence non-zero everywhere on $[0, \infty)^2$, see Theorem 1 in Schmidt and Stadtmüller (2006).

Let $\mathcal{B}_\infty(\mathbb{R}_2)$ denote the space of all functions $f : \mathbb{R}_2 \to \mathbb{R}$, which are locally uniformly bounded on every compact subset of $\mathbb{R}_2$, 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 recursively by $T_{3i} = T_{3i-1} \cup [0, i]^2, T_{3i-1} = T_{3i-2} \cup ([0, i] \times \{ \infty \}), T_{3i-2} = T_{3(i-1)} \cup (\{ \infty \} \times [0, i]), T_0 = 0$ and where $||f||_{T_i} = \sup_{x \in T_i} |f(x)|$ denotes the sup-norm on $T_i$. Note that with this metric the set $\mathcal{B}_\infty(\mathbb{R}_2^+)$ is a complete metric space and that a sequence $f_n$ in $\mathcal{B}_\infty(\mathbb{R}_2^+)$ converges with respect to $d$ if and only if it converges uniformly on every $T_i$, see Van der Vaart and Wellner (1996).

Throughout this paper $l^\infty(T)$ denotes the set of uniformly bounded functions on a set $T$, $\xrightarrow{T}$ denotes convergence in (outer) probability and $\xrightarrow{w}$ denotes weak convergence in the sense of Hoffmann-Jørgensen, see e.g. Van der Vaart and Wellner (1996).

Schmidt and Stadtmüller (2006) assumed that the lower tail copula $\Lambda_L$ satisfies the second-order condition

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

locally uniformly for $x = (x_1, x_2) \in \mathbb{R}_2^+$, where $g$ is a non-constant function and the function $A : [0, \infty) \to [0, \infty)$ satisfies $\lim_{t \to \infty} A(t) = 0$. Under this and the additional assumptions $\Lambda_L \neq 0$, $\sqrt{k}A(n/k) \to 0$, $\forall x \in [0, \infty)^2$.
where $G_B$ in (2.4) converges weakly in $\mathcal{B}_\infty(\mathbb{R}_+^2)$, that is
\begin{equation}
\sqrt{k} \left( \hat{\Lambda}_L(x) - \Lambda_L(x) \right) \rightsquigarrow G_{\hat{\Lambda}_L}(x),
\end{equation}
where $G_{\hat{\Lambda}_L}$ is a centered Gaussian field with covariance structure given by
\begin{equation}
E G_{\hat{\Lambda}_L}(x)G_{\hat{\Lambda}_L}(y) = \Lambda_L(x_1 \wedge y_1, x_2 \wedge y_2).
\end{equation}
For the empirical tail copula $\hat{\Lambda}_L(x)$ they established the weak convergence
\begin{equation}
\alpha_n(x) = \sqrt{k} \left( \hat{\Lambda}_L(x) - \Lambda_L(x) \right) \rightsquigarrow G_{\hat{\Lambda}_L}(x)
\end{equation}
in $\mathcal{B}_\infty(\mathbb{R}_+^2)$, provided that the tail copula has continuous partial derivatives. Here the limiting process $G_{\hat{\Lambda}_L}$ has the representation
\begin{equation}
G_{\hat{\Lambda}_L}(x) = G_{\hat{\Lambda}_L}(x) - \partial_1 \Lambda_L(x) G_{\hat{\Lambda}_L}(x_1, \infty) - \partial_2 \Lambda_L(x) G_{\hat{\Lambda}_L}(\infty, x_2).
\end{equation}
The assumption of continuous partial derivatives is made in the whole literature on estimation of stable tail dependence functions and tail copulas. However, as demonstrated in (1.3) there does not exist any tail copula $\Lambda_L \neq 0$ with continuous partial derivatives at the origin $(0,0)$. With our first result we will fill this gap and prove weak convergence of the empirical tail copula process under suitable weakened smoothness assumptions. For this purpose we will use a similar approach as in Schmidt and Stadtmüller (2006) since this turns out to be also useful for a proof of consistency of the multiplier bootstrap. First we consider the case of known marginals. Due to the second order condition (2.6) the proof of (2.7) can be given by showing weak convergence of the centered statistic
\begin{equation}
\tilde{\alpha}_n(x) := \sqrt{k} \left( \hat{\Lambda}_L(x) - \frac{n}{k} C(x_1 k/n, x_2 k/n) \right).
\end{equation}

Lemma 2.1. If $\Lambda_L \neq 0$ and the second order condition [2.6] holds with $\sqrt{k} A(n/k) \to 0$, where $k = k(n) \to \infty$ and $k = o(n)$, then we have, as $n$ tends to infinity
\begin{equation}
\tilde{\alpha}_n(x) = \sqrt{k} \left( \hat{\Lambda}_L(x) - \frac{n}{k} C(x_1 k/n, x_2 k/n) \right) \rightsquigarrow G_{\hat{\Lambda}_L}(x)
\end{equation}
in $\mathcal{B}_\infty(\mathbb{R}_+^2)$. Here $G_{\hat{\Lambda}_L}$ is a tight centered Gaussian field concentrated on $\mathcal{C}_\rho(\mathbb{R}_+^2)$ with covariance structure given in (2.8) and $\rho$ is a pseudometric on the space $\mathbb{R}_+^2$ defined by
\begin{equation}
\rho(x, y) = E \left[ (G_{\hat{\Lambda}_L}(x) - G_{\hat{\Lambda}_L}(y))^2 \right]^{1/2} = (\Lambda_L(x) - 2\Lambda_L(x \wedge y) + \Lambda_L(x))^{1/2},
\end{equation}
$x = (x_1, x_2), y = (y_1, y_2), x \wedge y = (x_1 \wedge y_1, x_2 \wedge y_2)$ and $\mathcal{C}_\rho(\mathbb{R}_+^2) \subset \mathcal{B}_\infty(\mathbb{R}_+^2)$ 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 (2006) 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 (2.9) we will use the functional delta method in (2.7) with some suitable functional.
Theorem 2.2. Let $\Lambda_L \neq 0$ be a lower tail copula whose first order partial derivatives satisfy the condition
\[
\partial_p \Lambda_L \text{ exists and is continuous on } \{x \in \mathbb{R}_+^2 \mid 0 < x_p < \infty\}
\tag{2.13}
\]
for $p = 1, 2$. If additionally the assumptions of Lemma 2.1 are satisfied then we have
\[
\alpha_n(x) = \sqrt{k} \left( \hat{\Lambda}_L(x) - \Lambda_L(x) \right) \Rightarrow G_{\hat{\Lambda}_L}(x)
\]
in $B_{\infty}(\mathbb{R}_+^2)$, where the process $G_{\hat{\Lambda}_L}$ is defined in (2.10) and $\partial_p \Lambda_L, p = 1, 2$ is defined as 0 on the set $\{x \in \mathbb{R}_+^2 \mid x_p \in \{0, \infty\}\}$.

Theorem 2.2 has been proved by Schmidt and Stadtmüller (2006) [see Theorem 6, therein] under the additional assumption that the tail copula has continuous partial derivatives. As pointed out in the previous paragraphs there does not exist any tail copula $\Lambda_L \neq 0$ with this property.

3 Multiplier bootstrap approximation

3.1 Asymptotic theory

In this section we will construct multiplier bootstrap approximations of the Gaussian limit distributions $G_{\tilde{\Lambda}_L}$ and $G_{\hat{\Lambda}_L}$ specified in (2.7) and (2.9), respectively. To this end let $\xi_i$ be independent identically distributed positive random variables, independent of the $X_i$, with mean $\mu$ in $(0, \infty)$ and finite variance $\tau^2$, which additionally satisfy $||\xi||_{2,1} := \int_0^\infty \sqrt{P(|\xi| > x)} \, dx < \infty$. We will first deal with the case of known marginals and define a multiplier bootstrap analogue of (2.4) by
\[
\tilde{\Lambda}_L^\xi(x) = \frac{1}{k} \sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n} \mathbb{I}\{F_1(X_{i1}) \leq \frac{k x_1}{n}, F_2(X_{i2}) \leq \frac{k x_2}{n}\}
\tag{3.1}
\]
where $\bar{\xi}_n = n^{-1} \sum_{i=1}^n \xi_i$ denotes the mean of $\xi_1, \ldots, \xi_n$. We have
\[
\tilde{\alpha}_n^m(x) = \frac{\mu}{\tau} \sqrt{n} \sum_{i=1}^n \left( \frac{\xi_i}{\bar{\xi}_n} - 1 \right) f_{n,x}(U_i) = \frac{\mu}{\tau} \sqrt{k} (\tilde{\Lambda}_L^\xi - \Lambda_L),
\tag{3.2}
\]
where the function $f_{n,x}(U_i)$ is defined by
\[
f_{n,x}(U_i) = \sqrt{\frac{n}{k}} \mathbb{I}\left\{U_{i1} \leq k x_1/n, U_{i2} \leq k x_2/n\right\}
\tag{3.3}
\]
and
\[
U_i = (U_{i1}, U_{i2}); \quad U_{ip} = F_p(X_{ip}) \text{ for } p = 1, 2.
\]

Throughout this paper we use the notation
\[
G_n \overset{\mathbb{P}}{\rightarrow} G \text{ in } \mathbb{D}
\tag{3.4}
\]
for conditional weak convergence in a metric space $(\mathbb{D}, d)$ in the sense of Kosorok (2008), page 19. To be precise, (3.4) holds for some random variables $G_n = G_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n), G \in \mathbb{D}$ if and only if
\[
\sup_{h \in BL_1(\mathbb{D})} \mathbb{P}\left|\mathbb{E}_\xi h(G_n) - \mathbb{E}h(G)\right| = 0 \tag{3.5}
\]
\[ \mathbb{E}_{\xi} h(G_n)^* - \mathbb{E}_{\xi} h(G_n) \xrightarrow{p} 0 \quad \text{for every } h \in BL_1(\mathbb{D}), \]

where

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

denotes the set of all Lipschitz-continuous functions bounded by 1. The subscript \( \xi \) in the expectations indicates conditional expectation over the weights \( \xi = (\xi_1, \ldots, \xi_n) \) given the data and \( h(G_n)^* \) and \( h(G_n) \) denote measurable majorants and minorants with respect to the joint data, including the weights \( \xi \). 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 (1998). The following result shows that the process (3.2) provides a valid bootstrap approximation of the process defined in (2.11).

**Theorem 3.1.** If \( \Lambda_L \neq 0 \) and the second order condition (2.6) holds with \( \sqrt{k}A(n/k) \to 0, k = k(n) \to \infty \) and \( k = o(n) \) we have, as \( n \) tends to infinity,

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

in the metric space \( B_{\infty}(\mathbb{R}_+^2) \).

Since Theorem 3.1 states that we have weak convergence of \( \hat{\alpha}_n^m \) to \( \mathbb{G}_{\hat{\Lambda}_L} \) conditional on the data \( U_i \), it provides a bootstrap approximation of the empirical tail copula in the case where the marginal distributions are known. To be precise, consider \( B \in \mathbb{N} \) independent replications of the random variables \( \xi_1, \ldots, \xi_n \) and denote them by \( \xi_{1,b}, \ldots, \xi_{n,b} \). Compute the statistics \( \hat{\alpha}_n^m(b) = \hat{\alpha}_n^m(\xi_{1,b}, \ldots, \xi_{n,b}) \) \((b = 1, \ldots, B)\) and use the empirical distribution of \( \hat{\alpha}_n^{m,1}, \ldots, \hat{\alpha}_n^{m,B} \) as an approximation for the limiting distribution of \( \mathbb{G}_{\hat{\Lambda}_L} \).

Because in most cases of practical interest there will be no information about the marginals one cannot use Theorem 3.1 in many statistical applications. We will now develop two consistent bootstrap approximation for the limiting distribution of the process (2.9) which 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

\[ \hat{\Lambda}_L^\xi(x) = \frac{1}{k} \sum_{i=1}^n \xi_i I\{X_i \leq F_{n_1}(kx_1/n), X_{i2} \leq F_{n_2}(kx_2/n)\} \]

which yields the process

\[ \beta_n(x) = \frac{\mu}{\tau} \sqrt{k}(\hat{\Lambda}_L^\xi - \hat{\Lambda}_L) = \frac{\mu}{\tau} \sqrt{k} \sum_{i=1}^n \left( \frac{\xi_i}{\xi_n} - 1 \right) I\{X_i \leq F_{n_1}(kx_1/n), X_{i2} \leq F_{n_2}(kx_2/n)\}. \]

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

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

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

in the metric space \( B_{\infty}(\mathbb{R}_+^2) \).
Although Theorem [3.2] provides a negative result and shows, that the distribution of $\beta_n$ can not be used for approximating the limiting law $G_{\hat{A}_L}$, 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. As a consequence, we obtain an approximation for the unknown distribution of the process $G_{\hat{A}_L}$. In order to get an approximation of $G_{\hat{A}_L}$ we follow Rémillard and Scaillet (2009) and estimate the derivatives of the tail copula by

$$\hat{\partial}_p \Lambda_L(x) := \begin{cases} \frac{\Lambda_L(x+he_p) - \Lambda_L(x-he_p)}{2h}, & x_p < h \\ \frac{\Lambda_L(x + (h - x_p)e_p)}{x_p}, & x_p = h \\ 0, & \text{else} \end{cases}$$

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

$$\alpha_{n}^{pdm}(x) = \beta_n(x) - \hat{\partial}_1 \Lambda_L(x) \beta_n(x_1, \infty) - \hat{\partial}_2 \Lambda_L(x) \beta_n(\infty, x_2).$$

(3.8)

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

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

$$\alpha_{n}^{pdm} \overset{\mathbb{P}}{\to} G_{\hat{A}_L}$$

in the metric space $B_\infty(\mathbb{R}^n_+)$.

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

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

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

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

and consider the process

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

(3.9)

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

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

$$\alpha_{n}^{dm}(x) = \frac{\mu}{\tau} \sqrt{k} \left( \hat{\Lambda}_L^\xi(x) - \Lambda_L(x) \right) \overset{\mathbb{P}}{\to} G_{\hat{A}_L} \quad \text{in} \quad B_\infty(\mathbb{R}^n_+).$$

(3.10)
3.2 Finite sample results

In this section we will present a small comparison of the finite sample properties of the two bootstrap approximations given in this section. 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 form the Clayton copula with a coefficient of lower tail dependence $\lambda_L = 0.25$. The Clayton copula is defined by

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

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

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

In Tables 1 and 2 we investigate the accuracy of the bootstrap approximation of the covariances of the limiting variable $\hat{G}_{\Lambda_L}$. We chose three points on the unit circle $\{e^{i\varphi}, \varphi = \ell \pi/8$ with $\ell = 1, 2, 3\}$ and present in the first four columns of Table 1 the true covariances of the limiting process $\hat{G}_{\Lambda_L}$. The remaining columns show the simulated covariances of the process $\alpha_n$ on the basis of $5 \cdot 10^5$ simulation runs, where the sample size is $n = 1000$ and the parameter $k$ is chosen as 50. This choice is motivated by the left panel of Figure 1 where we plot the sum of the squared bias, the variance and the mean squared error 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 neighbourhood of the point 50.

Note also that the literature provides several data-adaptive proposals for the choice of the parameter $k$, see for example Drees and Kaufmann (1998) oder Gomes and Oliveira (2001) in the univariate context. Table 1 now serves 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 procedure has been chosen according to Bücher and Dette (2010) as $P(\xi = 0) = P(\xi = 2) = 0.5$, such that $\mu = \tau = 1$. For the sake of completeness we also investigate the resampling bootstrap considered in Peng and Qi (2008) [which is hereafter denoted by $\alpha_{n^{res}}$]. The estimated covariances given in the first part (lines 3–5) of Table 2 have been calculated by 1000 simulation runs, where in each run the covariance is estimated on the basis of $B = 500$ bootstrap replications. The second part (lines 6–8) of Table 2 shows the corresponding mean squared error.

As one can see all bootstrap procedures yield approximations of comparable quality. Considering only the bias in Table 2 the $pdm$-bootstrap has slight advantages in all cases, while there are basically no differences between the $dm$- and the resampling bootstrap. A comparison of the mean squared error in Table 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. In this case, the $dm$-bootstrap yields the best results.

In the right panel of Figure 1 we investigate 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 one can see the best choices for $k$ lie in an interval of approximate length of 100 around the center $k = 200$. Compared to the “best” value $k = 50$ for estimating $\Lambda_L$ the optimal values for estimating the covariances of $\hat{G}_{\Lambda_L}$ are approximately four times larger for both the $pdm$- and the $dm$-bootstrap. This increase may be explained by the fact that the large bias of $\hat{\Lambda}_L(x)$, $\hat{\Lambda}_L^\xi(x)$ and $\hat{\Lambda}_L^{\xi\xi}(x)$ for estimating $\Lambda_L(x)$ cancels out if
Table 1: Left part: True covariances of $G_{\Lambda_L}$ for the Clayton Copula with $\lambda_L = 0.25$. Right part: sample covariances of the empirical tail copula process $\alpha_n$ with sample size $n = 1000$ and parameter $k = 50$.

|       | $\pi N$ | $2\pi N$ | $3\pi N$ | $\pi N$ | $2\pi N$ | $3\pi N$ | $\pi N$ | $2\pi N$ | $3\pi N$ |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| True  | 0.0874  | 0.0754  | 0.0516  | 0.0889  | 0.0737  | 0.0476  |
| $\alpha_n$ | 0.1160  | 0.0754  | 0.1218  | 0.0741  |
| $\alpha_n$ | 0.0874  | 0.0892  | 0.0892  |

Table 2: Averaged sample covariances (lines 3–5) and mean squared error $\times 10^4$ (lines 6–8) of the Bootstrap approximations $\alpha_n^{pdm}$, $\alpha_n^{dm}$ and $\alpha_n^{res}$ of $G_{\Lambda_L}$ under the conditions of Table 1.

|       | $\pi N$ | $2\pi N$ | $3\pi N$ | $\pi N$ | $2\pi N$ | $3\pi N$ | $\pi N$ | $2\pi N$ | $3\pi N$ |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| $\alpha_{n}^{pdm}$ | 0.094  | 0.072  | 0.046  | 0.100  | 0.071  | 0.045  | 0.100  | 0.070  | 0.043  |
| $\alpha_{n}^{dm}$ | 0.130  | 0.072  | 0.136  | 0.707  | 0.136  | 0.707  | 0.136  | 0.707  |
| $\alpha_{n}^{res}$ | 3.67   | 4.68   | 3.65   | 3.86   | 3.49   | 2.72   | 4.21   | 3.85   | 3.21   |
|       | 8.11   | 4.87   | 3.70   | 8.89   | 3.25   | 3.77   | 3.77   | 3.25   |

Figure 1: Left panel: 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 $\hat{\Lambda}_L(e^{i\pi/4})$ against the parameter $k$. Right panel: Averaged MSE, variance and squared bias for the bootstrap estimation of the covariances of $G_{\Lambda_L}$.
one calculates the difference $\hat{\Lambda}^{\xi}\xi(x) - \hat{\Lambda}_L(x)$ or $\hat{\Lambda}^{\xi}\xi L(x) - \hat{\Lambda}_L(x)$. As a result we may choose larger values of $k$ which yields a notable decay of the variance. These results indicate that a larger value of $k$ should be used in the bootstrap procedure.

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 than the dm-bootstrap, whereas the variance is slightly larger. On the other hand the differences with respect to the MSE are nearly not visible.

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 $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ denote two independent samples of independent identically distributed random variables [we will relax the assumption of independence between the samples later on] 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. We are interested in a test for the hypothesis

$$H_0 : \Lambda_{L,X} \equiv \Lambda_{L,Y} \text{ vs. } H_1 : \Lambda_{L,X} \neq \Lambda_{L,Y}$$

(4.1)

Due to the homogeneity of tail copulas we have $\Lambda_{L}(tx) = t\Lambda_{L}(x)$ for all $t > 0, x \in [0, \infty)^2$, and the hypotheses are equivalent to

$$H_0 : \varrho(\hat{\Lambda}_{L,X}, \hat{\Lambda}_{L,Y}) = 0 \text{ vs. } H_1 : \varrho(\hat{\Lambda}_{L,X}, \hat{\Lambda}_{L,Y}) > 0,$$

where the distance $\varrho$ is defined by

$$\varrho(\hat{\Lambda}_{L,X}, \hat{\Lambda}_{L,Y}) := \int_{0}^{\pi/2} (\hat{\Lambda}_{L,X}(\cos \varphi, \sin \varphi) - \hat{\Lambda}_{L,Y}(\cos \varphi, \sin \varphi))^2 d\varphi$$

(4.2)

and we have used the notation $\Lambda_{L,X}(\varphi) = \Lambda_{L,X}(\cos \varphi, \sin \varphi)$, $\Lambda_{L,Y}(\varphi) = \Lambda_{L,Y}(\cos \varphi, \sin \varphi)$. We propose to base the test for the hypothesis $[4.1]$ on the distance between the empirical tail copulas and define

$$S_n = \frac{k_1k_2}{k_1 + k_2} \varrho(\hat{\Lambda}_{L,X}, \hat{\Lambda}_{L,Y}) = \frac{k_1k_2}{k_1 + k_2} \int_{0}^{\pi/2} (\hat{\Lambda}_{L,X}(\varphi) - \hat{\Lambda}_{L,Y}(\varphi))^2 d\varphi,$$

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

$$k_p \to \infty, k_p = o(n_p) \text{ (p = 1, 2) } \text{ and } k_1/(k_1 + k_2) \to \lambda \in (0, 1).$$
We assume that the tail copulas $\Lambda_{L,X}$ and $\Lambda_{L,Y}$ satisfy a second order condition as in (2.6) (with $A$ replaced by $A_p$) and that $k_p$ is chosen appropriately, i.e. $\sqrt{k_p}A_p(k_p/n_p) = o(1)$. Under the null hypothesis (4.1) of equality between the tail copulas we have $S_n = T_n$ with

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

where

$$\mathcal{E}_n(x) = \sqrt{\frac{k_2}{k_1 + k_2}} \sqrt{\mathcal{E}_1(\hat{\Lambda}_{L,X}(x) - \Lambda_{L,X}(x))} - \sqrt{\frac{k_1}{k_1 + k_2}} \sqrt{\mathcal{E}_2(\hat{\Lambda}_{L,Y}(x) - \Lambda_{L,Y}(x))}.$$

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

$$\mathcal{E}_n \sim \sqrt{\frac{\beta}{\alpha}} \mathcal{G}_{\hat{\Lambda}_{L,X}} - \sqrt{\frac{\beta}{\alpha}} \mathcal{G}_{\hat{\Lambda}_{L,Y}} := \mathcal{E}.$$  \hspace{1cm} (4.3)

in the metric space $\mathcal{B}_{\infty}(\mathbb{R}^2_+)$, where the stochastically independent two-dimensional centered Gaussian fields $\mathcal{G}_{\hat{\Lambda}_{L,X}}$ and $\mathcal{G}_{\hat{\Lambda}_{L,Y}}$ are defined in (2.10). This yields by the continuous mapping theorem

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

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

$$S_n \sim_{H_0} \mathcal{T}, \quad S_n \xrightarrow{p} \mathcal{T}, \quad \text{for all } b \in \{1, \ldots, B\}$$

which shows that a test, which rejects the null hypothesis $4.1$ for large values of $T_n$, is consistent.

In order to determine critical values for the test we approximate the limiting distribution $\mathcal{T}$ by the multiplier bootstrap proposed in Section 3. For this purpose we exemplarily consider the $\text{pdm}$-bootstrap (the extension to the $\text{dm}$-bootstrap is straightforward) using the definition in equation (3.10) and denote for any $b \in \{1, \ldots, B\}$ by $\xi_{1,b}, \ldots, \xi_{n_1,b}, \xi_{1,b}, \ldots, \xi_{n_2,b}$ independent and identically distributed non-negative random variables with mean $\mu_1$ (resp. $\mu_2$) and variance $\tau_1^2$ (resp. $\tau_2^2$). We compute for each $b$ and both samples the bootstrap statistics as given in (3.8), i.e.

$$\alpha_{X,n_1,b}^{\text{pdm}}(x) = \beta_{X,n_1,b}(x) - \hat{\partial}_1 \Lambda_{L,X}(x) \beta_{X,n_1,b}(x_1, \infty) - \hat{\partial}_2 \Lambda_{L,X}(x) \beta_{X,n_1,b}(\infty, x_2),$$

$$\alpha_{Y,n_2,b}^{\text{pdm}}(x) = \beta_{Y,n_2,b}(x) - \hat{\partial}_1 \Lambda_{L,Y}(x) \beta_{Y,n_2,b}(x_1, \infty) - \hat{\partial}_2 \Lambda_{L,Y}(x) \beta_{Y,n_2,b}(\infty, x_2),$$

where

$$\beta_{X,n_1,b}(x) = \frac{\mu_1}{\tau_1} \frac{1}{\sqrt{k_1}} \sum_{i=1}^{n_1} \left( \frac{\xi_{i,b}}{\xi_{i,b}} - 1 \right) I\{X_{i} \leq F_{n_{1}}(k_{1}x_{1}/n_{1}), X_{i} \leq F_{n_{1}}(k_{2}x_{2}/n_{1})\},$$

$$\beta_{Y,n_2,b}(x) = \frac{\mu_2}{\tau_2} \frac{1}{\sqrt{k_2}} \sum_{i=1}^{n_2} \left( \frac{\xi_{i,b}}{\xi_{i,b}} - 1 \right) I\{Y_{i} \leq H_{n_{2}}(k_{1}x_{1}/n_{2}), Y_{i} \leq H_{n_{2}}(k_{2}x_{2}/n_{2})\},$$

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

$$\hat{\mathcal{E}}_n^{(\text{pdm}, b)}(x) := \sqrt{\frac{k_2}{k_1 + k_2}} \alpha_{X,n_1,b}^{\text{pdm}}(x) - \sqrt{\frac{k_1}{k_1 + k_2}} \alpha_{Y,n_2,b}^{\text{pdm}}(x),$$

12
\[ \hat{T}_n^{(pdm,b)} := \int_0^{\pi/2} \left\{ \hat{\mathcal{E}}_n^{(pdm,b)}(\cos \varphi, \sin \varphi) \right\}^2 d\varphi. \]

By Theorem 3.3 and Theorem 10.8 in Kosorok (2008), it follows for every \( b \in \{1, \ldots, B\} \)
\[ \hat{T}_n^{(pdm,b)} \xrightarrow{P} \xi^{(b)}, \]
where \( \xi^{(b)} \) is an independent copy of \( \xi \) [note that we consider the processes \( \hat{\mathcal{E}}_n^{(pdm,b)} \) in the Banach space \( l^\infty([0,1]^2] \)]. From (4.4) we therefore obtain a consistent asymptotic level \( \alpha \) test for the null hypothesis (4.1) by rejecting \( H_0 \) for large values of \( S_n \), that is
\[ S_n > q_{1-\alpha}^{pdm}. \tag{4.5} \]
where \( q_{1-\alpha}^{pdm} \) denotes the \((1-\alpha)\) quantile of the distribution function \( K_n^{pdm}(s) = B^{-1} \sum_{b=1}^{B} \mathbb{I}\{ \hat{T}_n^{(pdm,b)} \leq s \} \). The discussion present so far refers to two independent samples. Nevertheless it is easy to check that the methodology of the previous sections also applies if we are faced with paired observations, i.e. \( X \), is not independent of \( Y \), but \( n_1 = n_2 = n \). In that case we have to set \( \xi_{i,b} = \xi_i \) for all \( i = 1, \ldots, n \) and \( b = 1, \ldots, B \). To see this, set \( Z_i = (X_{i1}, X_{i2}, Y_{i1}, Y_{i2}) \) and denote the (empirical) copula of \( Z_i \) by \((C_n)\). Clearly,
\[ C(u_1, u_2) = C(u_1, u_2, 1, 1), \quad D(v_1, v_2) = D(1, 1, v_1, v_2), \]
\[ C_n(u_1, u_2) = C_n(u_1, u_2, 1, 1), \quad D_n(v_1, v_2) = D_n(1, 1, v_1, v_2). \]
If we set \( \Lambda_{L,Z}(x,y) = \lim_{t \to \infty} t C(x/t, y/t), \hat{\Lambda}_{L,Z}(x,y) = \frac{n}{n} C_n(\frac{nx}{t}, \frac{ny}{t}) \), we obtain
\[ \hat{\Lambda}_{L,X}(x) = \Lambda_{L,Z}(x, \infty, \infty), \quad \Lambda_{L,Y}(x) = \Lambda_{L,Z}(\infty, \infty, x), \]
\[ \hat{\Lambda}_{L,X}(x) = \hat{\Lambda}_{L,Z}(x, \infty, \infty), \quad \hat{\Lambda}_{L,Y}(x) = \hat{\Lambda}_{L,Z}(\infty, \infty, x). \]

Under a second order condition on the joint tail copula \( \Lambda_{L,Z} \) the asymptotic properties of the process \( \hat{\Lambda}_{L,Z} \) can be derived along similar lines as given before and the details are omitted for the sake of brevity. As the only difference to the preceding discussion note that the occurring limiting fields \( \mathbb{G}_{\hat{\Lambda}_{L,X}} \) and \( \mathbb{G}_{\hat{\Lambda}_{L,Y}} \) are not independent anymore. Since the asymptotic behavior of the multiplier bootstrap approximations can be shown to reflect this dependence we still obtain consistency of the test, the details are again omitted. For an investigation of the finite sample property we consider two independent samples of independent identically distributed random variables with Clayton copula, see (3.11), with a coefficient of lower tail dependence \( \lambda_L \) varying in the set \( \{0.25, 0.5, 0.75\} \).

In Table 3 we show the simulated rejection probabilities of the pdm- and dm-bootstrap test defined in (4.5) for various nominal levels on the basis of 1000 simulation runs. The sample size is \( n_1 = n_2 = n = 1000 \) 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 \( \mu = \tau = 1 \)] have been used. The parameter \( k \) is 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. It is worthwhile to mention that 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 this effect is increasing with the coefficient of tail dependence.
The alternative of different lower tail copulas is detected with reasonable power, where both tests yield rather similar results with slight advantages for the pdm-bootstrap. An investigation of the impact of the choice of the parameter $k$ under the alternative shows some advantages for $k = 200$. This may again be explained by the fact that bias terms cancel out if one calculates the difference $\hat{\Lambda}_{L,X} - \hat{\Lambda}_{L,Y}$.

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

In this section we are interested in estimating the tail copula of $X$ under the additional assumption that it is an element of some parametric class, say $\mathcal{L} = \{\Lambda_L(\cdot; \theta) \mid \theta \in \Theta\}$. Recently, estimates for parametric classes of tail copulas and stable tail dependence functions have been investigated by de Haan et al. (2008) and Einmahl et al. (2008) who proposed a censored likelihood and a moment based estimator, respectively.

In the present section we investigate a further estimate, which is based on the minimum distance method. To be precise let $\Lambda_L$ denote an arbitrary lower tail copula and $\Lambda_L(\cdot; \theta)$ 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)

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

where $\varrho$ is defined in (4.2). We call $\hat{\theta}_n^{MD} = T(\hat{\Lambda}_L)$ 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 the null hypothesis is satisfied.

Throughout this subsection let $X_1, \ldots, X_n$ denote independent identically distributed bivariate random variables with cumulative distribution function $F = C(F_1, F_2)$ and existing lower tail copula $\Lambda_L$. Furthermore, assume that the standard conditions of minimum distance estimation are satisfied. For a precise formulation of these conditions and a proof of the following result we refer to Bücher (2011).
Theorem 4.1. If the true tail copula $\Lambda_L$ satisfies the first order condition \(^{[2.13]}\) of Theorem \(^{2.2}\), then the minimum distance estimator $\hat{\Lambda}_n^{MD}$ is consistent for the parameter $\theta_B$ corresponding to the best approximation with respect to the distance $\varrho$. Moreover,

$$\Theta_n^{MD} := \sqrt{k}(\hat{\Lambda}_n^{MD} - \theta_B) = \sqrt{k} \int \gamma_{\theta_B}(\varphi) \left( \hat{\Lambda}_n^{\varphi}(\varphi) - \Lambda_L(\varphi) \right) d\varphi + o_P(1)$$

$$\rightsquigarrow \int \gamma_{\theta_B}(\varphi) G_{\Lambda_L}(\varphi) d\varphi =: \Theta^{MD},$$

where $\Lambda_L^{\varphi}(\varphi) = \Lambda_L(\cos \varphi, \sin \varphi)$, $\hat{\Lambda}_n^{\varphi}(\varphi) = \hat{\Lambda}_n(\cos \varphi, \sin \varphi)$, $\gamma_{\theta_B}(\varphi) = A_{\theta_B}^{-1} \delta_{\theta_B}(\varphi)$, $\delta_{\theta}(\varphi) = \partial_\theta \Lambda_L(\cos \varphi, \sin \varphi, \theta)$, $G_{\Lambda_L}(\varphi) = G_{\Lambda_L}(\cos \varphi, \sin \varphi)$ and

$$A_{\theta_B} := \int \delta_{\theta_B}(\varphi) \delta_{\theta_B}(\varphi') d\varphi =: \Theta^{MD},$$

with $\Lambda_L^{\varphi}(\varphi; \theta) = \Lambda_L(\cos \varphi, \sin \varphi; \theta)$. The limiting variable $\Theta^{MD}$ 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 $G_{\Lambda_L}$ defined in \(^{[2.10]}\).

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

Theorem 4.2. If the assumptions of the Theorems \(^{3.3}\) and \(^{3.4}\) 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

$$\Theta_n^{MD,m} := \int \gamma_{\hat{\theta}_n^{MD}}(\varphi) \Gamma_n^\varphi(\varphi) d\varphi \xrightarrow{\xi} \Theta^{MD},$$

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

$$\hat{\Lambda}_n^{MD} := \int \delta_{\hat{\theta}_n^{MD}}(\varphi) \delta_{\hat{\theta}_n^{MD}}(\varphi') d\varphi =: \Theta^{MD},$$

On the basis of this result it is possible to construct asymptotic confidence regions for the parameter $\theta$ as well as to test point hypotheses regarding the parameter. In Table \(^{4}\) we present 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 $K_{1-\alpha} = [\hat{\theta}_n^{MD} - k^{-1/2} \hat{q}_{1-\alpha/2}, \hat{\theta}_n^{MD} - k^{-1/2} \hat{q}_{\alpha/2}]$, 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 the calculation of 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 \({1/4, 2/4, 3/4}\).
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 \( G_{\Lambda_L} \).

In order to investigate the impact of the choice of \( k \) we chose three different scenarios: \( k = 50 \), \( k = 200 \) and for the last setting we used 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 is 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 \( k = 50 \) followed by the case of two different values of \( k \) [these findings may be compared to the results of Section 3.2]. For \( k = 200 \) the large bias of \( \hat{\theta}_n^{MD} \) (compare the left-hand side of Figure 1) entails that the true parameter does not lie in the estimated confidence interval for more than 95% of the repetitions. For stronger tail dependence \( \lambda_L = 0.75 \) the choice \( k = 200 \) yields better results, with almost perfect coverage probabilities for \( \lambda_L = 0.75 \). It is also of interest to note that the choice \( k = 50 \) in the estimator \( \hat{\Lambda}_L \) and \( k = 200 \) in the corresponding bootstrap statistic does not yield an improvement with respect to the approximation of the coverage probability compared to the case \( k = 50 \).

Remark 4.3. As pointed out at the beginning of this section there exist two alternative estimators for parametric classes of tail copulas. [de Haan et al. (2008)] 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. (2008)] proposed a method of moment type estimator and proved a similar statement as given in Theorem 4.1 for the minimum distance estimate. In Table 5 we compare the asymptotic variances of the method of moment and the minimum distance estimator for the parameter \( \theta \) in the Clayton family chosen such that the coefficient of tail dependence \( \lambda \) varies in the set \( \{0.1, \ldots, 0.9\} \). The calculated values \( E_\lambda \) are defined as

\[
E_\lambda = \frac{\text{Asymptotic variance of the minimum distant estimate}}{\text{Asymptotic variance of the moment type estimate}}
\]

(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 moment estimator requires the specification of a function \( g \), which was chosen as in [Einmahl et al. (2008)] as the indicator of the set \( \{x \in [0,1]^2 : x_1 + x_2 \leq 1\} \). We observe that none of the two estimates is globally preferable to the other. For small amounts of tail dependence the minimum distance estimate performs slightly better while for increasing tail dependence the moment type estimator is more qualified from an asymptotic point of view.

It is also notable that the dm- and pdm-bootstrap can be used to construct a consistent approximation of the asymptotic distribution of the censored likelihood and moment estimator investigated in [de Haan]...
Table 5: Relative efficiency of the minimum distant estimate to the moment type 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 | 1.14 | 1.38 | 2.11 |

et al. (2008) and Einmahl et al. (2008). The main argument for proving consistency is that the limiting distribution of the method of moments and the minimum distance estimator can be represented in the form $\Phi(\mathcal{G}_{\hat{\lambda}L}, \lambda_L, \partial \lambda_L)$ 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 the functional $\Phi$ is suitable smooth the bootstrap approximation is then obtained by $\Phi(\alpha_n, \hat{\lambda}_L, \hat{\partial \lambda}_L)$ where $\alpha_n$ is $\alpha_{n}^{pdm}$ of $\alpha_{n}^{dm}$ and $\hat{\partial \lambda}_L$ is a consistent estimate of $\partial \lambda_L$.

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

$$
\mathcal{H}_0 : \lambda_L \in \mathcal{L} = \{\lambda_L(\cdot, \theta) \mid \theta \in \Theta\}, \quad \mathcal{H}_1 : \lambda_L / \in \mathcal{L}.
$$

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

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

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

where $\hat{\theta}_n^{MD}$ denotes the minimum distance estimate. If the standard assumptions of minimum distance estimation are satisfied [see Bücher (2011) for details] we obtain for the process $H_n = \sqrt{k} \left( \hat{\lambda}_L - \lambda_L(\cdot; \hat{\theta}_n^{MD}) \right)$ under the null hypothesis $H_0 : \lambda_L = \lambda_L(\cdot; \theta_B)$

$$
H_n = \sqrt{k} \left( \hat{\lambda}_L - \lambda_L - \delta_\theta(\hat{\theta}_n^{MD} - \theta) \right) + o_P(1)
$$

$$
= \sqrt{k} \left( \hat{\lambda}_L - \lambda_L - \delta_\theta \int \gamma_\theta(\varphi) (\hat{\lambda}_L^\gamma(\varphi) - \lambda_L^\gamma(\varphi) d\varphi) + o_P(1)
$$

$$
\rightsquigarrow \mathcal{G}_{\lambda_L} - \delta_\theta \int \gamma_\theta(\varphi) \mathcal{G}_{\lambda_L}^\gamma(\varphi) d\varphi = \mathcal{G}_{\lambda_L} - \delta_\theta \Theta^{MD}.
$$

Under the alternative hypothesis we get an additional summand

$$
H_n = \sqrt{k} \left( \hat{\lambda}_L - \lambda_L - \delta_\theta(\hat{\theta}_n^{MD} - \theta) - (\lambda_L(\cdot; \theta_B) - \lambda_L) \right) + o_P(1),
$$

17
which converges to either plus or minus infinity whenever $\Lambda_L(x, \theta_B) \neq \Lambda_L(x)$. 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

$$GOF_n = \int \{H_n^\angle(\varphi)\}^2 d\varphi \rightarrow Z := \int \left( G_{\hat{\Lambda}_L}(\varphi) - \delta_{\hat{\theta}}(\varphi) \Theta^{MD} \right)^2 d\varphi,$$

(4.8)

while under the alternative $GOF_n = \int \{H_n^\angle(\varphi)\}^2 d\varphi \overset{p}{\rightarrow} \infty$.

The critical values of the test, which rejects the null hypothesis for large values of $GOF_n$ can be calculated on the basis of the following theorem. For a proof see Bücher (2011).

**Theorem 4.5.** If the assumptions of the Theorems 3.3 and 3.4 hold and $\Gamma_n$ denotes either the process $\alpha_{n\text{pdm}}$ (Theorem 3.3) or $\alpha_{n\text{dm}}$ (Theorem 3.4) obtained by the pdm- or dm-bootstrap, respectively, then it holds independently of the hypotheses that

$$H_n^m := \Gamma_n - \delta_{\hat{\theta}M}_n \int \gamma_{\hat{\theta}M}_n(\varphi) \Gamma_n^\angle(\varphi) d\varphi \overset{p}{\rightarrow} \xi G_{\hat{\Lambda}_L} - \delta_{\theta_B} \Theta^{MD}.$$

Therefore $GOF_n^m = \int \{H_n^{m\angle}(\varphi)\}^2 d\varphi \overset{p}{\rightarrow} \xi Z$, where $Z$ is defined in (4.8).

In order to investigate the finite sample properties of a goodness-of-fit test on the basis of the multiplier bootstrap we show in Table 6 the simulated rejection probabilities of the pdm-bootstrap test

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

(4.9)

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

1. A convex combination between the independence copula $\Pi(u) = u_1 u_2$ and a Clayton copula [with convex parameter 1/3], such that the tail copula is given by $\Lambda_L(x) = 1/3 (x_1^{-\theta} + x_2^{-\theta})^{-1/\theta}$. 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 (1990)], defined by

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

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 (1988)], 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.
Table 6: Simulated rejection probabilities of the pdm-bootstrap test (4.9) for the hypothesis (4.7). The first three lines are models from the null hypothesis, whereas the last eight lines correspond to alternatives. The sample size is \( n = 1000 \) and \( B = 500 \) Bootstrap replications have been performed. \( \lambda_L \) denotes the lower tail dependence coefficient.

The results are based on 1000 simulation runs, while the sample size is \( n = 1000 \) and two cases \( k = 50, 200 \) are investigated for the choice of the parameter \( k \). For each scenario the critical values have been calculated by \( B = 500 \) bootstrap replications with \( U(\{0,2\}) \)-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 \( \theta = 1 \)] the model is exactly the same as the Clayton model with parameter 1, i.e. we get close to the null hypothesis with increasing tail dependence. Finally, we note that a larger choice of the parameter \( k \) results in substantial better power properties, while we do not observe 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 \( H_n = \sqrt{k} \left( \hat{\Lambda}_L - \Lambda_L(\theta; \hat{\theta}_n^{MD}) \right) \). Therefore, we propose to use rather large values for \( k \) in applications of the goodness-of-fit test.

Acknowledgements The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise and to two unknown referees and the Associate editor for their constructive comments on an earlier version of this manuscript. This work has been supported by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823) of the German Research Foundation (DFG). The authors would also like to thank John Einmahl for pointing out important references on the subject and Johan Segers for discussing this subject with us in much detail.
A Proofs

A.1 Proof of Theorem 2.2

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

$$
\begin{aligned}
\gamma \ld (\gamma, \cdot, \cdot) \quad &\text{if} \quad x, y \neq \infty \\
\gamma(\gamma, x, \infty) \ld \gamma(\gamma, \infty, \cdot) \quad &\text{if} \quad y = \infty \\
\gamma(\infty, \gamma, \cdot) &\quad \text{if} \quad x = \infty 
\end{aligned}
$$

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

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

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

see also Schmidt and Stadtmüller (2006). Observing that $\Lambda_L \in \mathcal{B}_\infty^{l,0}(\mathbb{R}_+^2)$ and that the adjusted generalized inverse of $\Lambda_L(x, \infty)$ is given by $\frac{2}{z} F_1(F_{n1}(kx/n))$, one can conclude that $\Phi(\Lambda_L) = \Lambda_L$ and $\Phi(\hat{\Lambda}_L) = \hat{\Lambda}_L$ (P-almost surely) and the proof of Theorem 2.2 follows from the functional delta method (Theorem 3.9.4 in Van der Vaart and Wellner (1996)) and the following Lemma, which is an extension of the result in the proof of Theorem 5 in Schmidt and Stadtmüller (2006).

**Lemma A.1.** Let $\Lambda_L$ be a lower tail copula whose partial derivatives satisfy the first order property (2.13) for $p = 1, 2$. Then $\Phi$ is Hadamard-differentiable at $\Lambda_L$ tangentially to the set

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

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

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

(A.1)

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

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

$$
\begin{align*}
\Phi_1 : &\mathcal{B}_\infty^{l,0}(\mathbb{R}_+^2) \to \mathcal{B}_\infty^{l,0}(\mathbb{R}_+^2) \times \mathcal{B}_\infty^{l,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{l,0}(\mathbb{R}_+) \\
\gamma &\mapsto (\gamma(\cdot, \cdot), \gamma(\infty, \cdot), \gamma(\cdot, \cdot)) \\
\Phi_2 : &\mathcal{B}_\infty^{l,0}(\mathbb{R}_+^2) \times \mathcal{B}_\infty^{l,0}(\mathbb{R}_+) \times \mathcal{B}_\infty^{l,0}(\mathbb{R}_+) \to \mathcal{B}_\infty^{l,0}(\mathbb{R}_+^2) \times \mathcal{B}_\infty^{l,0,-}(\mathbb{R}_+) \times \mathcal{B}_\infty^{l,0,-}(\mathbb{R}_+) \\
(\gamma, f, g) &\mapsto (\gamma, f^-, g^-)
\end{align*}
$$

20
\[ \Phi_3: B_{\infty}^{1,0}(\mathbb{R}_+^2) \times B_{\infty}^{1,0,-}(\mathbb{R}_+) \times B_{\infty}^{1,0,-}(\mathbb{R}_+) \to B_{\infty}(\mathbb{R}_+^2) \]

\[(\gamma, f, g) \mapsto \begin{cases} 
\gamma(f(x), g(y)) & \text{if } x, y \neq \infty \\
\gamma(f(x), \infty) & \text{if } y = \infty \\
\gamma(\infty, g(y)) & \text{if } x = \infty,
\end{cases} \]

where \( B_{\infty}^{1,0,-}(\mathbb{R}_+) \) denotes the set of all adjusted generalized inverse functions \( f^- \) with \( f \in B_{\infty}^{1,0}(\mathbb{R}_+) \). Now \( \Phi_1 \) is Hadamard-differentiable at \( \Lambda_L \) tangentially to \( C^0(\mathbb{R}_+^2) \) since it is linear and continuous. The second map \( \Phi_2 \) is Hadamard-differentiable at \( (\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+}) \) tangentially to \( C^0(\mathbb{R}_+^2) \times C^0(\mathbb{R}_+) \times C^0(\mathbb{R}_+) \) where \( C^0(\mathbb{R}_+) \) consists of all continuous functions \( f \) on \( \mathbb{R}_+ \) with \( f(0) = 0 \) and its derivative at \( (\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+}) \) in \( (\gamma, f, g) \) is given by \( \Phi_2'((\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+}), (\gamma, f, g)) = (\gamma, -f, -g) \). The proof follows along similar as the one of Theorem 5 in Schmidt and Stadtmüller (2006, p. 321) and is therefore omitted, we just note that \( (\text{id}_{\mathbb{R}_+^2} + t_n f_n)(x) > 0 \) for all \( x > 0 \) is implied by the additional assumption of continuity in 0 for functions in the set \( B_{\infty}^{1,0}(\mathbb{R}_+) \). Some more efforts are necessary to show that \( \Phi_3 \) is Hadamard-differentiable at \( (\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+}) \) tangentially to \( C^0(\mathbb{R}_+^2) \times C^0(\mathbb{R}_+) \times C^0(\mathbb{R}_+) \) with derivative

\[ \Phi_3'((\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+}), (\gamma, f, g))(x) = \gamma(x) + \partial_1 \Lambda_L(x)f(x_1) + \partial_2 \Lambda_L(x)g(x_2). \]

To see this let \( t_n \to 0 \), \( (\gamma_n, f_n, g_n) \in B_{\infty}^{1,0}(\mathbb{R}_+^2) \times B_{\infty}(\mathbb{R}_+) \times B_{\infty}(\mathbb{R}_+) \) with \( (\gamma_n, f_n, g_n) \to (\gamma, f, g) \in C^0(\mathbb{R}_+^2) \times C^0(\mathbb{R}_+) \) such that \( (\Lambda_L + t_n \gamma_n, \text{id}_{\mathbb{R}_+^2} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n) \in B_{\infty}^{1,0}(\mathbb{R}_+^2) \times B_{\infty}(\mathbb{R}_+) \times B_{\infty}(\mathbb{R}_+) \). Now \( \Phi_3 \) is linear in its first argument and we introduce the decomposition

\[ t_n^{-1}\left\{ \Phi_3((\Lambda_L + t_n \gamma_n, \text{id}_{\mathbb{R}_+^2} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n)) - \Phi_3((\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+})) \right\} = L_{n1} + L_{n2}, \]

where

\[ L_{n1} = t_n^{-1}\left\{ \Phi_3((\Lambda_L, \text{id}_{\mathbb{R}_+^2} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n)) - \Phi_3((\Lambda_L, \text{id}_{\mathbb{R}_+^2}, \text{id}_{\mathbb{R}_+})) \right\}, \]

\[ L_{n2} = \Phi_3((\gamma_n, \text{id}_{\mathbb{R}_+^2} + t_n f_n, \text{id}_{\mathbb{R}_+} + t_n g_n)). \]

By the definition of \( d \) it suffices to show uniform convergence on sets \( T \) of the form \( T = [0, M_1] \times \{ \infty \} \cup \{ \infty \} \times [0, M_2] \cup [0, M_3]^2 \), where \( M_1, M_2, M_3 \in \mathbb{N} \). Since \( T \subseteq \mathbb{R}_+^2 \) is compact \((f_n, g_n)\) converges uniformly and \( \gamma \) is uniformly continuous; hence \( L_{n1} \) uniformly converges to \( \gamma \).

Considering \( L_{n1} \) we split the investigation into six different cases. First, let \( x \in (0, M_3]^2 \). A series expansion at \( x \) yields

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

where the error term \( r_n \) can be written as

\[ r_n(x) = (\partial_1 \Lambda_L(y) - \partial_1 \Lambda_L(x))f_n(x_1) + (\partial_2 \Lambda_L(y) - \partial_2 \Lambda_L(x))g_n(x_2) \]

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

Now consider the case \( x \in (0, M_3] \times \{0\} \). By Lipschitz-continuity of \( \Lambda_L \) on \( \mathbb{R}^2_+ \) we get

\[
|L_n(0, x)| = t_n^{-1} |\Lambda_L(x + t_n f_n(x), t_n g_n(0))| \\
= t_n^{-1} |\Lambda_L(x + t_n f_n(x), t_n g_n(0)) - \Lambda_L(x + t_n f_n(x), 0)| \\
\leq |g_n(0)| \rightarrow g(0) = 0.
\]

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

\[
L_n(0, x, \infty) = t_n^{-1} (\Lambda_L(x + t_n f_n(x, \infty) - \Lambda_L(x, \infty)) = f_n(x) \rightarrow f(x).
\]

By \( \partial_1 \Lambda_L(x, \infty) = 1 \) and \( \partial_2 \Lambda_L(x, \infty) = 0 \) this yields the assertion. 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 (1996)) completes the proof of the Lemma. \( \square \)

### A.2 Proof of Theorem 3.1

Due to Lemma B.2 in the Appendix B [which is an analogue of Theorem 1.6.1 in Van der Vaart and Wellner (1996) 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 \( l^{\infty}(T_i) \) separately. To this end we note that every \( T_i \) can be written in the form \( T = [0, M_1] \times \{\infty\} \cup \{\infty\} \times [0, M_2] \cup [0, M_3] \), and show conditional weak convergence in \( l^{\infty}(T) \). Recalling the notation of \( f_{n,x}(U_i) \) in (3.3) we can express \( \tilde{\alpha}_n^m \) as

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

and the assertion now follows by an application of Theorem 11.23 in Kosorok (2008). For this purpose we show that the assumptions for this result are satisfied. Let \( \mathcal{F}_n = \{ f_{n,x} : x \in T \} \) be a class of functions changing with \( n \) and denote by

\[
F_n(u) = \sqrt{\frac{n}{k}} \{ u_1 \leq k M/n \text{ or } u_2 \leq k M/n \},
\]

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

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

\[
\limsup_{n \rightarrow \infty} \sup_{Q} \int_0^1 \sqrt{\log N(\varepsilon, \|F_n\|_{Q,2}, F_n, L_2(Q))} d\varepsilon < \infty, \tag{A.2}
\]

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 \( H(x, y) = \lim_{n \rightarrow \infty} E[\tilde{\alpha}_n(x) \tilde{\alpha}_n(y)] \) exists for every \( x \) and \( y \) in \( T \).
\( \iii \limsup_{n \to \infty} \mathbb{E} F_n^2(U_1) < \infty \)

\( \iv \lim_{n \to \infty} \mathbb{E} F_n^2(U_1) \mathbb{I}\{F_n(U_1) > \varepsilon \sqrt{n}\} = 0 \) for all \( \varepsilon > 0 \).

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

\[
\rho_n(x, y) = \left( \mathbb{E}(f_{n,x}(U_1)) - f_{n,y}(U_1) \right)^2 / 2.
\]

Furthermore, for all sequences \( (x_n)_n, (y_n)_n \) in \( T \) the convergence \( \rho_n(x_n, y_n) \to 0 \) holds, provided \( \rho(x_n, y_n) \to 0 \).

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

(a) \( \mathbb{P}^*(\sup_{x \in T_n} \inf_{y \in T_n} |f_{n,x}(U_1) - f_{n,y}(U_1)| > 0) = 0 \),

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

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

\[
\begin{align*}
    f_{n,x}^{(1)}(U_1) &= \sqrt{\frac{n}{k}} \mathbb{I}\{U_{i1} \leq kx_1/n\} \mathbb{I}\{x_2 = \infty\}, \\
    f_{n,x}^{(2)}(U_1) &= \sqrt{\frac{n}{k}} \mathbb{I}\{U_{i2} \leq kx_2/n\} \mathbb{I}\{x_1 = \infty\}, \\
    f_{n,x}^{(3)}(U_1) &= \sqrt{\frac{n}{k}} \mathbb{I}\{U_{i1} \leq kx_1/n, U_{i2} \leq kx_2/n\} \mathbb{I}\{x_1 < \infty, x_2 < \infty\}.
\end{align*}
\]

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

\[
\begin{align*}
    F_n^{(1)}(U_1) &= \sqrt{\frac{n}{k}} \mathbb{I}\{U_{i1} \leq kM/n\}, \\
    F_n^{(2)}(U_1) &= \sqrt{\frac{n}{k}} \mathbb{I}\{U_{i2} \leq kM/n\}, \\
    F_n^{(3)}(U_1) &= \sqrt{\frac{n}{k}} \mathbb{I}\{U_{i1} \leq kM/n, U_{i2} \leq kM/n\},
\end{align*}
\]

so that \( F_n(U_1) = \max_{i=1}^3 \{F_n^{(i)}(U_1)\} \). If we prove that the sequences \( (\mathcal{F}_n^{(i)}, F_n^{(i)}) \) satisfy the bounded uniform integral entropy condition given in \( \text{(A.2)} \), then the condition holds also for \( (\mathcal{F}_n, F_n) \) by Lemma \( \text{C.1} \) in the appendix and thus the assertion in (i) is proved. We only consider the (hardest) case of \( \mathcal{F}_n^{(3)} \).

Note that \( \mathcal{F}_n^{(3)} = \{f_{n,x} : x \in [0, M]^2\} = G_{n,1}^{(3)} \cdot G_{n,2}^{(3)} \), where

\[
\begin{align*}
    f_{n,x} &= (n/k)^{1/2} \|U_{i1} \leq kx_1/n, U_{i2} \leq kx_2/n\|, \\
    G_{n,j} &= \{g_{n,t} = (n/k)^{1/4} \|U_{ij} \leq kt/n\| \colon t \in [0, M]\},
\end{align*}
\]

for \( j = 1, 2 \). Since the functions \( g_{n,t} \) are increasing in \( t \) the \( G_{n,j} \) are VC-classes with VC-index 2. Thus by Lemma 11.21 in \cite{Kosorok2008} both classes satisfy the bounded uniform integral entropy condition \( \text{(A.2)} \). Proposition 11.22 in \cite{Kosorok2008} 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 \( \mathbb{E}[\tilde{\alpha}_n(x)\tilde{\alpha}_n(y)] = n/k \left( C\left(\frac{x \land y}{n}\right) - C\left(\frac{x}{n}\right)C\left(\frac{y}{n}\right)\right) \), which converges to \( \Lambda_L(x \land y) =: H(x, y) \), since \( \frac{n}{k} C\left(\frac{x}{n}\right)C\left(\frac{y}{n}\right) \to 0 \).
Regarding (iii) and (iv) we note that \( \mathbb{E} F_n(U_1)^2 = 2M - \frac{n}{K} C(Mk/n, Mk/n) \), which converges to \( 2M - \Lambda_L(M, M) \). Further,
\[
\mathbb{E} F_n^2(U_1) \mathbb{I}\{F_n(U_1) > \varepsilon \sqrt{n}\} = \int_{\{F_n(U_1) > \varepsilon \sqrt{n}\}} F_n^2(U_1) \, d\mathbb{P} \\
\leq \frac{n}{K} \mathbb{P}\left(\frac{1}{K} \sum_{i=1}^{n} U_{11} \leq kM/n \text{ or } U_{12} \leq kM/n > \varepsilon\right) = 0
\]
for sufficiently large \( n \), such that \( k > 1/\varepsilon \). For (v) we note that
\[
\rho_n(x, y) = \left(\mathbb{E}(f_{n, x}(U_1) - f_{n, y}(U_1))^2\right)^{1/2} = \sqrt{\frac{n}{K}} \left(C(xk/n) - 2C((x \wedge y)k/n) + C(yk/n)\right)^{1/2}
\]
\[
\rightarrow (\Lambda_L(x) - 2\Lambda_L(x \wedge y) + \Lambda_L(y))^{1/2} =: \rho(x, y).
\]
Due to Theorem 1 in Schmidt and Stadtmüller (2006) we have locally uniform convergence in the latter expression, which yields the second condition stated in (v).

For the proof of condition (vi) we use Lemma 11.15 and the discussion on page 224 in Kosorok (2008) and show separability of \( F_n \), i.e. for every \( n \geq 1 \) there exists a countable subset \( T_n \subset T \) such that
\[
\mathbb{P}^*(\sup_{x \in T} \inf_{y \in T_n} |f_{n, y}(U_1) - f_{n, x}(U_1)| > 0) = 0.
\]
Choose \( T_n = (\mathbb{Q} \cap [0, M_1] \times \{\infty\}) \cup \{\infty\} \times (\mathbb{Q} \cap [0, M_2]) \cup (\mathbb{Q}^2 \cap [0, M_3]^2) \), then we have (note that the functions \( f_{n, x} \) are built by indicators) that for every \( \omega \) and every \( x \in T \) there is an \( y \in T_n \) with \( |f_{n, x}(U_1(\omega)) - f_{n, y}(U_1(\omega))| = 0 \). This yields the assertion and thus the proof of Lemma 2.1 is finished.

**Remark A.2.** Observing that
\[
\bar{\alpha}_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f_{n, x}(U_i) - \mathbb{E} f_{n, x}(U_i))
\]
and that in Section A.2 we showed the sufficient conditions for an application of Theorem 11.20 in Kosorok (2008), 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 Theorem 3.2 and 3.3. The proof is essentially a consequence of a bootstrap version of the functional delta method, see Theorem 12.1 in Kosorok (2008). Since this result only holds for Banach space valued stochastic processes some adjustments have to be made. Note that the space \( B_\infty(\mathbb{R}_+^d) \) is a complete topological vector space with a metric \( d \) and some care is necessary whenever technical results depending on the norm are used.

Due to Lemma 2.1 and Theorem 3.1 we have
\[
\sqrt{k}(\hat{\Lambda}_L - \Lambda_L) \rightsquigarrow \mathcal{G}_{\hat{\Lambda}_L}, \quad \sqrt{k} \frac{\mu}{\tau} \left(\hat{\Lambda}_L^\xi - \hat{\Lambda}_L\right) \rightsquigarrow \xi \mathcal{G}_{\hat{\Lambda}_L}
\]
in \( B_\infty(\mathbb{R}_+^d) \). Observing that the generalized inverses of \( \hat{\Lambda}_L(x, \infty) \) and \( \hat{\Lambda}_L^\xi(x, \infty) \) are (\( \mathbb{P} \)-almost surely) given by \( \frac{\mu}{\tau} F_1(F_{n,1}(kx/n)) \) and \( \frac{\mu}{k} F_1(F_{n,1}^\xi(kx/n)) \), respectively, one can conclude that \( \Phi(\Lambda_L) = \Lambda_L, \Phi(\hat{\Lambda}_L) = \hat{\Lambda}_L \) and \( \Phi(\hat{\Lambda}_L^\xi) = \hat{\Lambda}_L^\xi \) (\( \mathbb{P} \)-almost surely). By Lemma A.1 \( \Phi \) is Hadamard-differentiable on \( B_\infty^I(\mathbb{R}_+^d) \) at \( \gamma_0 = \Lambda_L \).
tangentially to $\mathcal{C}^0(\mathbb{R}^2_+) \subset \mathcal{B}_\infty(\mathbb{R}^2_+)$. Therefore it remains to argue why Theorem 12.1 in Kosorok (2008) can be applied in the present context.

A careful inspection of the proof of Theorem 12.1 in Kosorok (2008) shows that properties going beyond our specific assumptions (i.e. the complete topological vector space $(\mathcal{B}_\infty(\mathbb{R}^2_+), d)$) are used only three times. First of all the mapping $\Phi'_{\Lambda_L}$ needs to be extended to the whole space $\mathcal{B}_\infty(\mathbb{R}^2_+)$, which is possible using equation (A.1) as the defining identity. Secondly, the proof of Theorem 12.1 in Kosorok (2008) uses the usual functional delta method as stated in Theorem 2.8 in the same reference, but this result can be replaced by Theorem 3.9.4 in Van der Vaart and Wellner (1996), which provides a functional delta method holding in general metrizable topological vector spaces. Finally, the proof of Theorem 12.1 in Kosorok (2008) makes use of a bootstrap continuous mapping theorem, see Theorem 10.8 in Kosorok (2008), which would yield that

$$\Phi'_{\Lambda_L}(\sqrt{\mathcal{H}_\theta} (\tilde{\Lambda}_L - \Lambda_L)) \xrightarrow{P} \Phi'_{\Lambda_L}(\mathcal{G}_{\tilde{\Lambda}_L}).$$

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 $\Psi : \mathcal{B}_\infty^I(\mathbb{R}_+^2) \times \mathcal{B}_\infty^I(\mathbb{R}_+^2) \rightarrow \mathcal{B}_\infty(\mathbb{R}^2_+)$ 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

$$\Psi_1 : \mathcal{B}_\infty^I(\mathbb{R}_+^2) \times \mathcal{B}_\infty^I(\mathbb{R}_+^2) \rightarrow \mathcal{B}_\infty^I(\mathbb{R}_+^2) \times \mathcal{B}_\infty^I(\mathbb{R}_+) \times \mathcal{B}_\infty^I(\mathbb{R}_+)

(\beta, \gamma) \mapsto (\beta, \gamma(\cdot, \infty), \gamma(\infty, \cdot)).$$

Note that we obtain the representations $\Psi(\Lambda_L, \Lambda_L) = \Lambda_L$, $\Psi(\tilde{\Lambda}_L, \tilde{\Lambda}_L) = \tilde{\Lambda}_L$ and $\Psi(\tilde{\Lambda}_L, \tilde{\Lambda}_L) = \tilde{\Lambda}_L$ (P-almost surely). Clearly, $\Psi_1$ is Hadamard-differentiable at $(\Lambda_L, \Lambda_L)$ since 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 Van der Vaart and Wellner (1996), we can conclude that $\Psi$ is Hadamard-differentiable $(\Lambda_L, \Lambda_L)$ tangentially to $\mathcal{C}^0(\mathbb{R}^2_+) \times \mathcal{C}^0(\mathbb{R}^2_+)$ with derivative

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

Note that, unlike in the previous proof, we do not have weak convergence (resp. weak conditional convergence) of $\sqrt{\mathcal{H}_\theta} \left( (\Lambda_L, \tilde{\Lambda}_L) - (\Lambda_L, \Lambda_L) \right)$ and $\frac{1}{\sqrt{\mathcal{H}_\theta}} \left( (\tilde{\Lambda}_L, \tilde{\Lambda}_L) - (\Lambda_L, \Lambda_L) \right)$ towards the same limiting field, which would be necessary for an application of the functional delta method for the bootstrap [see for example Theorem 12.1 in Kosorok (2008)]. Nevertheless, we can mimic certain steps in the proof of this theorem to conclude the result. To be precise, note that we obtain by analogous arguments as on page 236 of Kosorok (2008) that

$$\sqrt{\mathcal{K}} \left( \frac{\tilde{\Lambda}_L - \Lambda_L}{\Lambda_L - \Lambda_L} \right) \xrightarrow{\text{unconditionally}} \left( c^{-1} \mathbb{G}_1 + \mathbb{G}_2 \right) \text{G}_2,$$

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 $(\beta, \gamma) \mapsto (\Psi(\beta, \gamma), \Phi(\gamma, \gamma), (\beta, \gamma), (\gamma, \gamma))$ and the usual functional delta method [Theorem...
Continuity of the map (T \text{ Let} A.5 Proof of Theorem 3.3. 3.9.4 in Van der Vaart and Wellner (1996) yields for any \( \delta \) in outer probability and thus by boundedness of the metric assertion of the Theorem we set proceed similar as in the proof of Lemma 4.1 in Segers (2011). The details are omitted. Regarding the \( T \) in outer probability. For a proof of (A.5) split \( \hat{\Lambda} \) into three subsets as indicated by its definition and then proceed similar as in the proof of Lemma 4.1 in Segers (2011). The details are omitted. Regarding the assertion of the Theorem we set

\[
\bar{\partial}_n^{\text{pdm}}(x) = \beta_n(x) - \partial_1 \Lambda_n(x) \beta_n(x_1, \infty) - \partial_2 \Lambda_n(x) \beta_n(\infty, x_2).
\]

Under consideration of Lemma C.4 it suffices to prove that \( d(\alpha_n^{\text{pdm}}, \hat{\alpha}_n^{\text{pdm}}) \) converges to 0 in outer probability. By the definition of \( d \) we have to show uniform convergence on the set \( T \). Since \( |\alpha_n^{\text{pdm}} - \hat{\alpha}_n^{\text{pdm}}| \leq D_n1 + D_n2 \), where

\[
D_n1 = |\partial_1 \hat{\Lambda}_L - \partial_1 \Lambda_n| |\beta_n(\cdot, \infty)|, \quad D_n2 = |\partial_2 \hat{\Lambda}_L - \partial_2 \Lambda_n| |\beta_n(\infty, \cdot)|
\]

we can consider both summands \( D_{np} \) separately and deal with \( D_n1 \) exemplarily. First consider the case \( x \in [0, M_3]^2 \), then for arbitrary \( \varepsilon > 0 \) and \( \delta \in (0, 1) \)

\[
\mathbb{P}^*(\sup_{x \in [0, M_3]^2} D_n1(x) > \varepsilon) \leq \mathbb{P}^*(\sup_{x \in [0, M_3]^2, x_1 \geq \delta} D_n1(x) > \varepsilon/2) + \mathbb{P}^*(\sup_{x \in [0, M_3]^2, x_1 < \delta} D_n1(x) > \varepsilon/2).
\]

(A.6)
Since $\hat{\partial}_1 \Lambda_L$ is uniformly consistent on $\{x \in [0, M_3]^2 \mid x_1 \geq \delta\}$ and since $\beta_n$ is asymptotically tight in $l^\infty(T)$ $[\beta_n$ converges unconditionally by the results in Chapter 10 of Kosorok (2008)] the first probability on the right-hand side converges to zero.

Regarding the second summand note that $F^{-1}_n(kx/n) = \lfloor kx/n \rfloor$ so that

$$\sup_{x \in [0, M_3]^2, x_1 \geq \delta} \left| \hat{\partial}_1 \Lambda_L(x) \right| \leq \sup_{x \in [0, M_3]^2, x_1 \geq \delta} \frac{\lfloor k(x_1 + h) \rfloor - \lfloor k(x_1 - h) \rfloor}{2h} \leq 1 + \frac{M_3}{2kh} \leq 2$$

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

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

eventually. As $\beta_n \rightsquigarrow G_{\Lambda_L}$ (unconditionally) the lim sup of this outer probability is bounded by

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

Since $G_{\Lambda_L}$ has continuous trajectories and $G_{\Lambda_L}(0, x_2) = 0$ (almost surely) this probability can be made arbitrary small by choosing $\delta$ sufficiently small. The case $x \in [0, M_3]^2$ is finished. For $x \in [0, M_1] \times \{\infty\}$ the arguments are similar, while for $x \in \{\infty\} \times [0, M_2]$ we have $D_{n1} = 0$ and nothing has to be shown. To conclude, $\sup_{x \in T} D_{n1}(x)$ converges to zero in outer probability and because the term $\sup_{x \in T} D_{n2}$ can be treated similarly the proof is finished.

B Partial derivatives of tail copulas

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

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

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

Proof. By groundedness and homogeneity of $\Lambda$, see Theorem 1 in Schmidt and Stadtmüller (2006), we have

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

for all $x \in (0, \infty)$. Similarly, $\partial_1 \Lambda(0, 0) = 0$. The addendum follows by Theorem 1 (iv) in Schmidt and Stadtmüller (2006).

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

In the last section we present several technical details. We omit the proofs of the assertions and refer the reader to the thesis B"ucher (2011).

Lemma C.1. Suppose $G_n$ and $H_n$ are sequences of measurable functions with envelopes $G_n$ and $H_n$, so that $(G_n, G_n)$ and $(H_n, H_n)$ satisfy the bounded uniform integral entropy condition as stated in (A.2). Then the bounded uniform entropy integral condition (A.2) holds also for $F_n = G_n \cup H_n$, with envelopes $F_n = G_n \lor H_n$.

Lemma C.2. Suppose $G_n = G_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n)$ is some statistic taking values in $B_\infty(\bar{\mathbb{R}}^2)$. Then a conditional version of Theorem 1.6.1 in Van der Vaart and Wellner (1996) holds, namely $G_n \xrightarrow{\mathbb{P}} G$ in $B_\infty(\bar{\mathbb{R}}^2)$ is equivalent to $G_n \xrightarrow{\xi \mathbb{P}} G$ in $l^\infty(T_i)$ 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 $G_n = G_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n) \xrightarrow{\mathbb{P}} G$ in $\mathbb{D}_1$, where $G$ is tight, then $g(G_n) \xrightarrow{\xi \mathbb{P}} g(G)$ in $\mathbb{D}_2$.

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

References

B"ucher, A. (2011). Statistical inference for copulas and extremes. PhD thesis, Ruhr-University Bochum, Germany.

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

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

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

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

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

Drees, H. and Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. Stochastic Process. Appl., 75(2):149–172.

Einmahl, H. J., de Haan, L., and Li, D. (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. Annals of Statistics, 34:1987–2014.
Einmahl, H. J., de Haan, L., and Xin, H. (1993). Estimating a multidimensional extreme-value distribution. *Journal of Multivariate Analysis*, 47:35–47.

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

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

Gomes, M. I. and Oliveira, O. (2001). The bootstrap methodology in statistics of extremes—choice of the optimal sample fraction. *Extremes*, 4(4):331–358 (2002).

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

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

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

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

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

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

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

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

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

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

Segers, J. (2011). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. *arXiv:1012.2133*.

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

Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika*, 75:397–415.
Tiago de Oliveira, J. (1980). Bivariate extremes: Foundations and statistics. In Krishnaiah, P. R., editor, *Multivariate Analysis 5*, pages 349–366. North-Holland, Amsterdam.

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

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