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

Nonparametric estimation of tail dependence

RAFAEL SCHMIDT[1] and ULRICH STADTMÜLLER[2]

ABSTRACT. Dependencies between extreme events (extremal dependencies) are attracting an increasing attention in modern risk management. In practice, the concept of tail dependence represents the current standard to describe the amount of extremal dependence. In theory, multivariate extreme-value theory (EVT) turns out to be the natural choice to model the latter dependencies. The present paper embeds tail dependence into the concept of tail copulae which describes the dependence structure in the tail of multivariate distributions but works more generally. Various non-parametric estimators for tail copulae and tail dependence are discussed, and weak convergence, asymptotic normality, and strong consistency of these estimators are shown by means of a functional delta method. Further, weak convergence of a general upper-order rank-statistics for extreme events is investigated and the relationship to tail dependence is provided. A simulation study compares the introduced estimators and two financial data sets are analyzed with our methods.

Key words: asymptotic normality, copula, empirical copula, nonparametric estimation, strong consistency, tail copula, tail dependence, tail-dependence coefficient.

1 Introduction.

Dependencies between (extreme) financial asset-returns have significantly increased during recent time periods in almost all international markets. This phenomenon is a direct consequence of globalization and relaxed market regulation in finance and insurance industry. Especially during bear markets many empirical surveys like Karolyi & Stulz (1996), Longin & Solnik (2001), and Campbell, Koedijk & Kofman (2002) show evidence of increasing dependencies between (extreme) asset-returns. However, increasing extremal dependencies strongly impact the companies' profit contributions and may weaken the financial stability of entire industrial sectors. Typically, risk managers pursue diversification strategies by analyzing and utilizing positive and negative correlations between various asset-returns in order to cut one's losses due to *market* or *credit risk* and to increase the (risk-adjusted) returns. However, diversification strategies become less effective or may break down if the financial markets fall simultaneously during bear markets or market crashes. According to Ong (1999), the primary issue risk managers have always been interested in, is assessing the *size* - more than the *frequency* - of losses. For example, the presumable most well-known risk measure called the Value-at-Risk (VaR) (describes the amount of extreme portfolio loss which is exceeded only with a certain small probability) depends strongly on the dependence structure of extreme events which makes it important to model and analyze extremal dependence.

In practice, the current standard of studying extremal dependencies is to use the concept of *tail dependence*, cf. Embrechts et al. (2003), Malevergne & Sornette (2004). The aim of the present paper is to study the *estimation* of the so-called tail-dependence coefficient in a *nonparametric* context. Therefore, tail dependence is embedded into the general framework of *tail copulae* which refers to the dependence structure of extreme events of multivariate distributions independently of their marginal distributions and hence is of interest in extreme value theory as well. Our definition of tail copulae and tail dependence is based on the well-known concept of copulae (referring

[1]Department of Economic and Social Statistics, University of Cologne, Germany and Department of Statistics, London School of Economics, UK

[2]Department of Number Theory and Probability Theory, University of Ulm, Germany.

to the underlying distribution). The considered nonparametric estimator of the tail copula, for which we obtain an estimator of the tail-dependence coefficient as a special case, will be named *empirical tail copula.* The reason for this is its close relationship to the *empirical copula (process)* which has been investigated by many authors in the context of process convergence; we mention Deheuvels (1979, 1981) , Stute (1984), Van der Vaart & Wellner (1996), and Fermanian, Radulović & Wegkamp (2004). The last two references establish weak convergence of the empirical copula process under independent and dependent marginal distributions. The methods we utilize to derive the asymptotics of the empirical tail copula process have been essentially inspired by the methods used in Van der Vaart & Wellner (1996), Chapter 3.9.4.4, and Fermanian et al. (2004). However, the limiting results are different and the proofs are more delicate, since we work in the tails of the distribution. From the viewpoint of EVT, the empirical tail copula, although different, is closely related to a nonparametric estimator introduced and studied by Huang (1992), Chapter 2, (see also Peng (1998), pp.96, and Drees & Huang (1998)) in the context of so-called stable tail-dependence functions. The techniques used in Huang (1992) cannot be applied to the empirical tail copula as we explain later. We will see that the empirical tail copula and the resulting estimate for the tail dependence coefficient have smaller finite sample bias than the estimates based on the concept described above. Another type of estimators for the tail copula can be established via the so-called spectral measure. The nonparametric estimation of the latter has been investigated by several authors, e.g. Abdous et al. (1999) and Einmahl, de Haan & Piterbarg (2001). For the estimation of the tail-dependence coefficient, the estimator in Einmahl et al. (2001) coincides with the estimator proposed by Huang (1992). However, the general limiting process turns out to be complicated and difficult to estimate, which makes it unattractive especially for (financial) applications. Beside the derivation of the limiting process of the empirical tail copula, we will propose a procedure how to estimate (asymptotic) confidence intervals regarding the estimation of the tail-dependence coefficient.

Similarly to the well-known copula-concept, cf. Joe (1997), Nelsen (1999), we may construct multivariate extreme-value distributions with a given tail copula. Copulae itself have become quite prominent in theory and applications, see e.g. Sklar (1996), Song (2000), Cuculescu & Theodorescu (2003), Embrechts et al. (2003). Important applications of tail copulae in actuarial sciences and finance concern the modelling of dependencies between extreme insurance claims and large default events in credit portfolios, and Value-at-Risk considerations of asset portfolios.

We start with the definition of *tail copulae* and *tail dependence*, and derive several analytical properties which justify the name *tail copula*, even though it is not a copula in the usual sense. In Section 3, various non-parametric estimators for the tail copula are discussed and Section 4 provides the main results on weak convergence, asymptotic normality, and strong consistency by means of a functional delta method (as provided in Van der Vaart & Wellner (1996)). The subsequent section generalizes the asymptotic results for functionals of general upper-order rank statistics. A simulation study and a real data analysis complement the theoretical results. Some mathematics can be found in the Appendix.

2 Copulae, tail copulae and tail dependence

The theory of copulae investigates the dependence structure of multidimensional random vectors. Copulae are functions that join or "couple" multivariate distribution functions to their corresponding marginal distribution functions. A *copula* function $C: [0, 1]^n \to [0, 1]$ is a multivariate distribution function with uniformly distributed margins on the interval $[0, 1]$. For notational convenience, all further definitions and results are provided for the bivariate case only. Various multidimensional extensions are possible. We consider a random vector $(X, Y)'$ with joint distribution function F and continuous marginal distribution functions G (for X) and H (for Y). Then the bivariate distribution function F can be written in the form

$$F(x, y) = C(G(x), H(y)) \quad \text{or} \quad C(u, v) = F(G^{-1}(u), H^{-1}(v)),\ 0 \le u, v \le 1, \tag{1}$$

where G^{-1}, H^{-1} denote the generalized inverse distribution functions of G and H, i.e., for all $u \in [0, 1] : G^{-1}(u) := \inf\{x \in I\!R \mid G(x) \ge u\}$ with $\inf\{\emptyset\} = \infty$. Conversely, if C is a copula and G, H are distribution functions, then the function F defined via (1) is a bivariate distribution function with margins G, H. Further, with $U := G(X)$ and $V := H(Y)$, the random variables U and V are uniform on $[0,1]$ and $C(u,v) = \mathbb{P}(U \le u, V \le v)$. The *survival copula* associated with the survival function $\bar{F}(x,y) = \mathbb{P}(X > x, Y > y)$ is defined by

$$\bar{F}(x,y) = \bar{C}(\bar{G}(x), \bar{H}(y)) \quad \text{or} \quad \bar{C}(u,v) = \bar{F}(\bar{G}^{-1}(u), \bar{H}^{-1}(v)),\ 0 \le u, v \le 1, \tag{2}$$

where $\bar{G} := 1 - G$ and $\bar{G}^{-1}(u) := G^{-1}(1-u)$ (analogously for H). Hence, we have

$$\bar{C}(u,v) = \mathbb{P}\Big(X > G^{-1}(1-u), Y > H^{-1}(1-v)\Big) = \mathbb{P}(U > 1-u, V > 1-v)$$

and thus $\bar{C}(u,v) = u + v - 1 + C(1-u, 1-v)$. For more details regarding the theory of copulae we refer the reader to the monograph of Nelsen (1999) or Joe (1997).

Tail copulae are functions that describe the dependence structure of joint distributions in the tail and are defined as follows. Throughout this paper we denote by $\bar{\mathbb{R}}_+^2 := [0,\infty]^2 \backslash \{(\infty,\infty)\}$.

Definition 1 (Tail copulae) *Let F be a distribution function with corresponding copula C. If the following limit exists everywhere on $\bar{\mathbb{R}}_+^2$*

$$\Lambda_L(x,y) := \lim_{t\to\infty} tC(x/t, y/t) \tag{3}$$

then the function $\Lambda_L : \bar{\mathbb{R}}_+^2 \to \mathbb{R}$ is called a lower tail copula associated with F.
The corresponding upper tail copula is defined by

$$\Lambda_U(x,y) := \lim_{t\to\infty} t\bar{C}(x/t, y/t) \tag{4}$$

provided the limit exists.

The following relationships between tail copulae and joint/conditional distributions are worth mentioning:

$$\Lambda_L(x,y) = y \lim_{t\to\infty} \mathbb{P}(U \le x/t \mid V \le y/t) = \lim_{t\to\infty} t\mathbb{P}\big(X \le G^{-1}(x/t), Y \le H^{-1}(y/t)\big)$$

$$\text{and} \quad \Lambda_U(x,y) = \lim_{t\to\infty} t\mathbb{P}\big(X > G^{-1}(1-x/t), Y > H^{-1}(1-y/t)\big).$$

The next definition embeds the well-known concept of tail dependence (Sibuya 1960) within the framework of tail copulae. The tail-dependence coefficient has attracted a lot of attention among practitioners as a simple and intuitive measure for dependence between extreme events (such as extreme asset returns or extreme credit losses), cf. Embrechts et al. (2003), Malevergne & Sornette (2004). However, the estimation of tail dependence is a nontrivial task, especially for nonstandard distributions (For an account on tail dependence for elliptically contoured distributions we refer to Schmidt (2002).). It is precisely this reason that motivates us to consider the tail-dependence coefficient, as a special case of the tail copula, in more detail.

Definition 2 (Tail dependence) *A random vector $(X,Y)'$ is said to be upper tail-dependent if $\Lambda_U(1,1)$ exists and*

$$\lambda_U := \Lambda_U(1,1) = \lim_{t\to\infty} t\bar{C}(1/t, 1/t) > 0. \tag{5}$$

Consequently, $(X,Y)'$ is called upper tail-independent if λ_U equals 0. Further, λ_U is referred to as the upper tail-dependence coefficient. Similarly, the lower tail-dependence coefficient is defined by $\lambda_L := \Lambda_L(1,1)$ if existent and lower tail-dependence (independence) is present if $\lambda_L > 0$ $(= 0)$.

It is well known that the multivariate normal distributions, the multivariate generalized hyperbolic distributions (cf. Barndorff-Nielsen (1978)), and the multivariate logistic distributions are upper and lower tail-independent whereas the multivariate t-distributions and the α-stable distributions are upper and lower tail-dependent.

For our purpose, tail copulae are of primary interest because of the following four reasons:

1. As an intuitive generalization of the tail-dependence coefficient via a function describing the dependence structure in the tail of a distribution.
2. To derive explicit weak convergence results for the estimators presented below of the lower and upper tail-dependence coefficient in a completely nonparametric distribution model.
3. As a counterpart of the so-called stable tail-dependence function, cf. Huang (1992), Chapter 2, and Peng (1998), pp.96, which we describe later.
4. As another starting point to construct multidimensional extreme-value distributions.

Estimating the tail copula can be coped with techniques from EVT. It can be shown, see Resnick (1987), Chapter 5, that the upper tail copula exists on $\mathbb{R}_+^2$ and $\Lambda_U \neq 0$ if the associated distribution function F lies in the domain of attraction of an (max-stable) extreme-value distribution with dependent margins. A similar result holds for the lower tail copula. However, the latter is only a sufficient condition as the marginal distributions are not necessarily in the domain of attraction of an extreme-value distribution. Further, within the concept of tail dependence we do not require the existence of the entire tail copula. As a consequence our estimate will not necessarily provide an extreme value distribution.

If F lies in the domain of attraction of some (max-stable) extreme-value distribution and if we normalize the margins to Fréchet distributions then the corresponding extreme-value distribution G_E follows

$$G_E(x,y) = \exp\{-1/x - 1/y + \Lambda_U(1/x, 1/y)\} \quad \text{for } x, y > 0\,.$$

Note that if $\Lambda_U \equiv 0$ on $\mathbb{R}_+^2$ then we are in the independent extremal situation. Obviously the function Λ_U describes the dependence structure of the extreme-value distribution, this is one reason why we call it a tail copula even though it does not posses all copula properties.

In bivariate EVT the major interest concerns the probability

$$\mathbb{P}(X > G^{-1}(1-x) \text{ or } Y > H^{-1}(1-y)), \tag{6}$$

whereas in the context of the (upper) tail copulae the probability under consideration closely relates to

$$\mathbb{P}(X > G^{-1}(1-x) \text{ and } Y > H^{-1}(1-y)). \tag{7}$$

In case of tail dependence, the mapping $t \mapsto \mathbb{P}(X > G^{-1}(1-x/t) \text{ and/or } Y > H^{-1}(1-y/t))$ is regularly varying of order -1, and consequently a homogeneity property holds for large t (see the next section for more details).

At this point we would like to mention that the nonparametric estimators we propose later are based on the empirical counterparts of the probabilities (6) and (7) and utilize the above homogeneity property. Notice, in case $(X,Y)'$ is tail independent, the latter property does not hold for (7). Here, an adjusted homogeneity property can sometimes be obtained namely if the limit $\lim_{t\to\infty} t^{\eta}\mathbb{P}(X > G^{-1}(1-x/t), Y > H^{-1}(1-y/t))$, $\eta < 1$, exists and does not vanish. The parameter η was introduced by Ledford & Tawn (1997, 1996, 1998) as the coefficient of asymptotic dependence given tail independence. Several estimators for η and related tests for tail independence were introduced by Coles, Heffernan & Tawn (1999), Peng (1999), and Draisma, Drees, Ferreira & de Haan (2004). However, according to the latter paper the tests on tail dependence or tail independence show a disappointing behavior. In contrast to these approaches we concentrate on tail dependence (i.e. the case $\eta = 1$).

3 Tail-copula properties

The name *tail copula* is justified by the results of the present section. Many properties of the tail copula are closely related to copula properties, cf. Nelsen (1999), Chapter 2.

Theorem 1 *If the limit functions $\Lambda_U(x,y)$ and $\Lambda_L(x,y)$, $(x,y)' \in \bar{\mathbb{R}}_+^2$, exist, they have the following properties.*

i) **(Groundedness)** $\Lambda_U(x,0) = \Lambda_U(0,y) = \Lambda_L(x,0) = \Lambda_L(0,y) = 0$ *for all* $x,y \in \bar{\mathbb{R}}_+$, *and* $\Lambda_U(x,\infty) = \Lambda_L(x,\infty) = x$ *and* $\Lambda_U(\infty,y) = \Lambda_L(\infty,y) = y$ *for all* $x,y \in \mathbb{R}_+$.

ii) **(Homogeneity)** $\Lambda_U(tx,ty) = t\Lambda_U(x,y)$ *and* $\Lambda_L(tx,ty) = t\Lambda_L(x,y)$ *for all* $t > 0$ *and* $(x,y)' \in \bar{\mathbb{R}}_+^2$.

iii) **(Monotonicity)** $\Lambda_U(x,y)$ *and* $\Lambda_L(x,y)$ *are nondecreasing and Lipschitz continuous.*

iv) $\Lambda_U(x,y)$ *and* $\Lambda_L(x,y)$ *are nonzero everywhere if they do not vanish in a single point* $(x,y)' \in \mathbb{R}_+^2$. *Hence* $\Lambda_U(x,y) = 0$ $(\Lambda_L(x,y) = 0)$ *for all* $(x,y)' \in \mathbb{R}_+^2$ *in case of upper (lower) tail-independence.*

v) **(Uniformity)** *The limit relations for* $\Lambda_U(x,y)$ *and* $\Lambda_L(x,y)$ *are locally uniform in* $(x,y)' \in \mathbb{R}_+^2$.

Proof. Properties i) and ii) follow immediately from Definition 1. Note that the limit of a regular varying function with index -1 is homogeneous.

iii) Consider e.g. $\Lambda := \Lambda_L$ and let C denote the corresponding copula. As the limit of nondecreasing functions, Λ_L is nondecreasing. Further, for $(x,y)', (\bar{x},\bar{y})' \in \bar{\mathbb{R}}_+^2$ we have

$$
\begin{aligned}
|\Lambda(x,y) - \Lambda(\bar{x},\bar{y})| &= \lim_{t\to\infty} t|C(x/t,y/t) - C(\bar{x}/t,\bar{y}/t)| \\
&\le |x-\bar{x}| + |y-\bar{y}| \le K\left|\left|\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}\right|\right|_2
\end{aligned}
\tag{8}
$$

for some constant $K > 0$ because C is a distribution function with uniform margins.

iv) The following inequalities hold for a, $b > 0$

$$
\min\{a,\, b\}\Lambda(x,y) \le \Lambda(a\,x,\, b\,y) \le \max\{a,\, b\}\Lambda(x,\, y)\,.
$$

To verify this, note that in case $a \le b$, using $\tau = t/a$ we find

$$
\begin{aligned}
\Lambda(a\,x,\, b\,y) &= \lim_{t\to\infty} t\,C(ax/t,\, by/t) \\
&= \lim_{\tau\to\infty} a\,\tau\,C(x/\tau,\, (b/a)y/\tau) = a\,\Lambda(x,\, (b/a)\,y) \ge a\,\Lambda(x,\, y)
\end{aligned}
$$

and the upper inequality follows similarly. Notice that this result also implies homogeneity. Next, if $\Lambda(x_0,\, y_0) > 0$ for some $x_0,\, y_0 > 0$, then we get

$$
\Lambda(x,\, y) \ge \min\{x/x_0,\, y/y_0\}\Lambda(x_0,\, y_0) > 0\,.
$$

v) Finally, uniform convergence is obtained from the fact that for $x_n \to x_0$, $y_n \to y_0$ and $t_n \to \infty$, putting $\tau_n = t_n/\min\{x_n/x_0,\, y_n/y_0\}$ and $\xi_n = t_n/\max\{x_n/x_0,\, y_n/y_0\}$ we have

$$
\min\{x_n/x_0,\, y_n/y_0\}\tau_n\, C(x_0/\tau_n,\, y_0/\tau_n) \le t_n\, C(x_n/t_n,\, y_n/t_n)
$$

$$
\le \max\{x_n/x_0,\, y_n/y_0\}\xi_n\, C(x_0/\xi_n,\, y_0/\xi_n).
$$

This implies that $t_n\, C(x_n/t_n,\, y_n/t_n) \to \Lambda(x_0,\, y_0)$ as $t_n \to \infty$.

The next properties are given for the lower tail copula only. However, analogous properties hold for the upper pendant.

Theorem 2 *Suppose the limit function $\Lambda_L(x,y),\ (x,y)' \in \bar{I\!R}_+^2$, exists. Then for all $(x,y)', (\bar{x},\bar{y})' \in \bar{I\!R}_+^2$ such that $x \le \bar{x}, y \le \bar{y}$ the following properties hold.*

i) **("Fréchet-Hoeffding bounds")** $0 \le \Lambda_L(x,y) \le \min\{x,y\}$.

ii) For $a,\, b > 0$: $\min\{a,\, b\}\Lambda_L(x,y) \le \Lambda_L(a\,x,\, b\,y) \le \max\{a,\, b\}\Lambda_L(x,\, y)$.

iii) **(2-increasing)** $\Lambda_L(\bar{x},\bar{y}) - \Lambda_L(\bar{x},y) - \Lambda_L(x,\bar{y}) + \Lambda_L(x,y) \ge 0$.

iv) **(Strict monotonicity)** *For $\Lambda_L \not\equiv 0 : \Lambda_L(x,y) < \Lambda_L(\bar{x},\bar{y})$ if $x < \bar{x}$ and $y < \bar{y}$.*

v) If Λ_L exist for $x,y \in I\!R_+^2$ with $x^2+y^2=1$ then it exists everywhere on $\bar{I\!R}_+^2$.

Proof. i) The upper bound arises from the upper Fréchet-Hoeffding bound for copulae, cf. Nelsen (1999), Theorem 2.2.3. In particular, for every copula function C we have $C(u,v) \le \min(u,v)$. Part ii) has already been shown in the proof of Theorem 1, part iv). Part iii) is deduced from the fact that every copula is 2-increasing. Finally, part iv) is implied by part ii) and the last part follows directly from

$$\lim_{t\to\infty} t\,C(x/t,y/t) = (\sqrt{x^2+y^2}\lim_{\tau\to\infty}\tau C(x/(\sqrt{x^2+y^2}\tau), y/(\sqrt{x^2+y^2}\tau))\,.$$

Theorem 3 *Suppose the limit function $\Lambda_L(x,y),\ (x,y)' \in \bar{I\!R}_+^2$, exists. Then, for any $y \in \bar{R}_+$ the derivative $\partial\Lambda_L/\partial x$ exists for almost all $x \in I\!R_+$, and for such x and y*

$$0 \le \frac{\partial}{\partial x}\Lambda_L(x,y) \le 1. \tag{9}$$

Similarly, for any $x \in \bar{R}_+$ the partial derivative $\partial\Lambda_L/\partial y$ exists for almost all $y \in I\!R_+$, and for such x and y

$$0 \le \frac{\partial}{\partial y}\Lambda_L(x,y) \le 1. \tag{10}$$

Furthermore, the functions $x \mapsto \partial\Lambda_L(x,y)/\partial y$ and $y \mapsto \partial\Lambda_L(x,y)/\partial x$ are defined and nondecreasing almost everywhere on $\bar{I\!R}_+$.

Proof. The partial derivatives $\partial\Lambda_L/\partial x$ and $\partial\Lambda_L/\partial y$ exist because monotone functions are differentiable almost everywhere (c.f Theorem 7.2.1 in Wheeden & Zygmund (1977)). Inequalities (9) and (10) are implied by the Lipschitz condition (8). Further, for fixed $y \le \bar{y}$ the function $y \mapsto \Lambda_L(x,y) - \Lambda_L(x,\bar{y})$ is nondecreasing according to part iii) in Theorem 2. Thus $\partial(\Lambda_L(x,y) - \Lambda_L(x,\bar{y}))/\partial x$ is defined and nonnegative almost everywhere. The final assertion is now immediate.

4 Nonparametric estimators

Suppose $(X,Y)',\ (X^{(1)},Y^{(1)})', \ldots, (X^{(m)}, Y^{(m)})'$ are iid random vectors with distribution function F having marginal distribution functions G, H and copula C.

Below we propose a set of nonparametric estimators for the upper and lower tail copula $\Lambda_U(x,y)$ and $\Lambda_L(x,y),\ (x,y)' \in \bar{R}_+^2$. Note that nonparametric estimation turns out to be appropriate for unknown tail copulae as no general finite-dimensional parametrization of tail copulae exists (in contrast to the one-dimensional EVT). Let C_m denote the *empirical copula* defined by

$$C_m(u,v) = F_m(G_m^{-1}(u), H_m^{-1}(v)),\ (u,v)' \in [0,1]^2 \tag{11}$$

with F_m, G_m, H_m being the empirical distribution functions corresponding to F, G, H. Analogously we define the *empirical survival copula* by $\bar{C}_m(u,v) = \bar{F}_m(\bar{G}_m^{-1}(u), \bar{H}_m^{-1}(v)),\ (u,v)' \in [0,1]^2$ with

$$\bar{F}_m(x,y) = \frac{1}{m}\sum_{j=1}^{m} \mathbf{1}_{\{X^{(j)}>x, Y^{(j)}>y\}}$$

and $\bar{G}_m = 1 - G_m$, $\bar{H}_m = 1 - H_m$. Note that a slightly modified version of the generalized inverse distribution function of $\bar{G}_m$ and $\bar{H}_m$ is necessary, i.e., we define $\bar{G}^{-1}(u) := \sup\{x \in \mathbb{R} \mid \bar{G}(x) \geq u\}$ with $\sup\{\emptyset\} = -\infty$. The choice of the empirical distribution function to model the marginal distributions avoids any misidentification due to a wrong parametrical fit of the marginal distributions. Empirical investigations regarding such misidentifications and misinterpretations of the corresponding (extremal) dependence structure are provided in Frahm, Junker & Schmidt (2005).

Let $R_{m1}^{(j)}$ and $R_{m2}^{(j)}$ denote the rank of $X^{(j)}$ and $Y^{(j)}$, $j = 1, \ldots, m$, respectively. The first set of estimators are based on formulae (3) and (4):

$$\hat{\Lambda}_{L,m}(x, y) := \frac{m}{k} C_m\Big(\frac{kx}{m}, \frac{ky}{m}\Big) \approx \frac{1}{k} \sum_{j=1}^{m} \mathbf{1}_{\{R_{m1}^{(j)} \leq kx \text{ and } R_{m2}^{(j)} \leq ky\}} \tag{12}$$

and

$$\hat{\Lambda}_{U,m}(x, y) \quad := \quad \frac{m}{k} \bar{C}_m\Big(\frac{kx}{m}, \frac{ky}{m}\Big) \approx \frac{1}{k} \sum_{j=1}^{m} \mathbf{1}_{\{R_{m1}^{(j)} > m-kx \text{ and } R_{m2}^{(j)} > m-ky\}} \tag{13}$$

with some parameter $k \in \{1, \ldots, m\}$ to be chosen by the statistician. For the asymptotic results we assume throughout this paper that $k = k(m) \to \infty$ and $k/m \to 0$ as $m \to \infty$. The estimators $\hat{\Lambda}_{U,m}(x, y)$ and $\hat{\Lambda}_{L,m}(x, y)$ are referred to as *empirical tail copulae.*

The far right sides in equations (13) and (12) provide two approximative rank order statistics which are based on a slightly modified representation of the empirical tail copula. Such a representation was proposed by Genest, Ghoudi & Rivest (1995):

$$\check{C}_m(u, v) = \frac{1}{m} \sum_{j=1}^{m} \mathbf{1}_{\{G_m(X^{(j)}) \leq u \text{ and } H_m(Y^{(j)}) \leq v\}}, \quad (u, v)' \in [0, 1]^2. \tag{14}$$

The present paper establishes results of weak convergence and strong consistency for the tail-copula estimators $\hat{\Lambda}_{U,m}$ and $\hat{\Lambda}_{L,m}$. However, the following reasoning shows that all results hold also for the related rank order statistics. Note that C_m, $\bar{C}_m$, and the corresponding empirical tail copulae depend also on the ranks of $X^{(j)}$ and $Y^{(j)}$, $j = 1, \ldots, m$.

The empirical tail copulae and the rank order statistics coincide on the grid $\{(i/k, j/k),\ 1 \leq i, j \leq m\}$. Otherwise the pointwise differences are at most $2/k$. Consider e.g. the lower empirical tail copula $\hat{\Lambda}_{L,m}(x, y)$ which is left-continuous whereas the corresponding rank order statistics is right-continuous. The difference between the latter estimators is bounded by

$$\sup_{(x,y)' \in \bar{\mathbb{R}}_+^2} |\hat{\Lambda}_{L,m}(x, y) - \frac{1}{k} \sum_{j=1}^{m} \mathbf{1}_{\{R_{m1}^{(j)} \leq kx \text{ and } R_{m2}^{(j)} \leq ky\}}|$$
$$\leq \max_{1 \leq i,j \leq m} \Big| \frac{m}{k} C_m\Big(\frac{i}{m}, \frac{j}{m}\Big) - \frac{m}{k} C_m\Big(\frac{i-1}{m}, \frac{j-1}{m}\Big) \Big| \leq \frac{m}{k} \frac{2}{m} = \frac{2}{k}.$$

Alternatively, we could define the empirical copula C_m in (11) by considering a right-continuous version of the inverse function, i.e., $\tilde{G}^{-1}(u) = \inf\{x \in \mathbb{R} \mid G(x) > u\}$. This choice would yield a right continuous version of the empirical copula. The asymptotic behavior, however, does not change because of the same reasoning as above.

The following related estimator was introduced and investigated by Huang (1992), Chapter 2, (see also Peng (1998), pp.96, and Drees & Huang (1998)) in the context of so-called stable tail-dependence functions. The relationship between the upper tail copula and the stable tail-dependence function l is given by $\Lambda_U(x, y) = x + y - l(x, y)$. The latter authors discuss the estimation of the function l with respect to questions arising from EVT. Recall that $\Lambda_U(x, y) = x + y -$

$\lim_{t\to\infty} t(1-C(1-x/t,1-y/t))$. The corresponding estimator for $\Lambda_U(x,y)$ on $I\!R_+^2$ is

$$\begin{aligned}\hat{\Lambda}^{EVT}_{U,m}(x,y) &:= x+y-\frac{m}{k}\Big(1-C_m\Big(1-\frac{kx}{m},1-\frac{ky}{m}\Big)\Big) \\ &\approx x+y-\frac{1}{k}\sum_{j=1}^{m}\mathbf{1}_{\{R^{(j)}_{m1}>m-kx \text{ or } R^{(j)}_{m2}>m-ky\}}, \quad (x,y)\in I\!R_+^2,\end{aligned} \tag{15}$$

with $k=k(m)\to\infty$ and $k/m\to 0$ as $m\to\infty$. The estimator $\hat{\Lambda}^{EVT}_{L,m}$ could be similarly defined by substituting the empirical survival copula for the empirical copula in (15). An important practical problem for all estimators arises in the optimal choice of the parameter k which relates to the usual variance-bias problem. Some methods of choosing an optimal k are described below.

The main purpose of the paper concerns the study of the asymptotic behavior of the empirical tail copulae $\hat{\Lambda}_{U,m}$ and $\hat{\Lambda}_{L,m}$ stated in (13) and (12). These estimators, although different, are related to the estimator in (15) proposed in Huang (1992) and an asymptotic result for (15) is proven in the latter reference. However, instead of utilizing the Skorokhod representation theorem, we apply a general Delta method to prove the asymptotic results for $\hat{\Lambda}_{U,m}(x,y)$ and $\hat{\Lambda}_{L,m}(x,y)$. We must chose a different approach since the method of proof, as it is used in Huang (1992), cannot be applied in our case. In particular, the estimator in (15) is only defined on $I\!R_+^2$ and the fluctuation of the empirical tail copula at infinity cannot be coped with using the techniques in the latter reference. The evaluation of the tail copula at the point $x=\infty$ or $y=\infty$ is very useful since by doing this we obtain immediately the lower-dimensional tail copula (similarly to setting arguments equal to 1 in a multivariate distribution function). We remark that the asymptotic results for $\hat{\Lambda}^{EVT}_{U,m}$ can be shown with the same techniques. The estimator $\hat{\Lambda}^{EVT}_{U,m}$ has an additional finite sample bias since the uncertainty of the copula's margins enter into the estimation (cf. formula (15)). The absolute difference between $\hat{\Lambda}_{U,m}$ and $\hat{\Lambda}^{EVT}_{U,m}$ is bounded by $2/k$ for $x\le m/k$ and $x-m/k$ for $x>m/k$. Although at the first glance this difference for $x\le m/k$ appears to be small, one has to keep in mind that for sample sizes $m=1000$ (2000) the threshold is usually around $k=50$ (100). The impact of the additional (finite sample) bias cannot be disregarded which we illustrate in the empirical study in Section 7. Based on the above estimators for the lower and upper tail copula, we propose

$$\hat{\lambda}_{U,m} := \hat{\Lambda}_{U,m}(1,1), \text{ and } \hat{\lambda}^{EVT}_{U,m} := \hat{\Lambda}^{EVT}_{U,m}(1,1) \tag{16}$$

as nonparametric estimators for the upper tail-dependence coefficient and analogous estimates for the lower tail-dependence coefficient.

Another type of estimators utilizes a representation of tail copulae via the *spectral measure* with respect to the $||\cdot||_\infty$ norm which is well known in EVT. In particular the following relationship holds if the underlying distribution is in the domain of attraction of an extreme value distribution:

$$\Lambda_U(x,y) = x+y-\int_0^{\pi/2}\Big(\frac{x}{1\vee\cot\theta}\vee\frac{y}{1\vee\tan\theta}\Big)\Phi(d\theta), \quad (x,y)\in I\!R_+^2,$$

where the finite measure Φ, which lives on $[0,\pi/2]$, denotes the spectral measure of Λ_U. Einmahl et al. (2001) propose a nonparametric estimator for the above spectral measure Φ :

$$\hat{\Phi}_m(\theta) = \frac{1}{k}\sum_{j=1}^{m}\mathbf{1}_{\Big\{R^{(j)}_{m1}\vee R^{(j)}_{m2}\ge m+1-k,\, \arctan\frac{m+1-R^{(j)}_{m2}}{m+1-R^{(j)}_{m1}}\le\theta\Big\}}$$

for $\theta\in[0,\pi/2]$ and discuss the related asymptotic properties. Thus, a natural estimator for the upper tail copula is defined by

$$\hat{\Lambda}^{S}_{U,m}(x,y) := x+y-\int_0^{\pi/2}\Big(\frac{x}{1\vee\cot\theta}\vee\frac{y}{1\vee\tan\theta}\Big)\hat{\Phi}_m(d\theta),$$

with $k = k(m) \to \infty$ and $k/m \to 0$ as $m \to \infty$. Note that $\hat{\lambda}^S_{U,m} := \hat{\Lambda}^S_{U,m}(1,1)$ degenerates to $\hat{\Phi}_m(\pi/2)$ and therefore $\hat{\lambda}^S_{U,m}$ corresponds to the rank order statistics in (15) which is approximately $\hat{\Lambda}^{EVT}_{U,m}$. Another representation of the spectral measure, called the Pickands representation (Pickands 1981), yields similar estimators. In this context, nonparametric estimators have been proposed, e.g., in Abdous et al. (1999), Hall & Tajvidi (2000), and Falk & Reiss (2003). Further nonparametric estimators for the lower tail-dependence coefficient are introduced in Dobrić & Schmid (2005).

5 Asymptotic normality

The proof of asymptotic normality for the estimators $\hat{\Lambda}_{U,m}(x,y)$ and $\hat{\Lambda}_{L,m}(x,y)$ is accomplished in two steps. In the first step we assume that the margins G and H are known, and we provide the asymptotic results. In the second step we assume that the marginal distribution functions G and H are unknown, and prove asymptotic normality by utilizing a Delta method (see Theorem 9). The techniques to convey this can be found in Van der Vaart & Wellner (1996). Some important tools and the underlying space $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$, where weak convergence takes place, are provided in the Appendix A.1. In the case of known marginal distribution functions G and H we consider the following estimator for $\Lambda_U(x,y)$ and $\Lambda_L(x,y)$:

$$\hat{\Lambda}^*_{U,m}(x,y) := \frac{1}{k}\sum_{j=1}^{m} \mathbf{1}_{\left\{G(X^{(j)}) > 1 - \frac{kx}{m} \text{ and } H(Y^{(j)}) > 1 - \frac{ky}{m}\right\}} \tag{17}$$

and

$$\hat{\Lambda}^*_{L,m}(x,y) := \frac{1}{k}\sum_{j=1}^{m} \mathbf{1}_{\left\{G(X^{(j)}) \leq \frac{kx}{m} \text{ and } H(Y^{(j)}) \leq \frac{ky}{m}\right\}}\,. \tag{18}$$

Condition 1 (Second order condition) *The lower tail copula $\Lambda_L(x,y)$ is said to satisfy a second order condition if a function $A : \mathbb{R}_+ \to \mathbb{R}_+$ exists such that $A(t) \to 0$ as $t \to \infty$ and*

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

locally uniformly for $(x,y)' \in \bar{\mathbb{R}}^2_+$ and some nonconstant function g. The second order condition for the upper tail copula is defined analogously.

Note that $A(t)$ is regularly varying at infinity so this is just a second order condition on regular variation, cf. de Haan & Stadtmüller (1996).

Theorem 4 (Asymptotic normality under known margins G and H) *Let F be a distribution function with continuous marginal distribution functions G and H. Suppose the tail copulae $\Lambda_U \not\equiv 0$ and $\Lambda_L \not\equiv 0$ exist and the Second order Condition 1 with*

$$\sqrt{k}\, A(m/k) \to 0 \quad as \quad m \to \infty \tag{19}$$

holds for some sequence $k = k(m) \to \infty$ and $k/m \to 0$. Then

$$\sqrt{k}\left(\hat{\Lambda}^*_{U,m}(x,y) - \Lambda_U(x,y)\right) \xrightarrow{w} \mathbb{G}_{\hat{\Lambda}^*_U}(x,y) \tag{20}$$

and

$$\sqrt{k}\left(\hat{\Lambda}^*_{L,m}(x,y) - \Lambda_L(x,y)\right) \xrightarrow{w} \mathbb{G}_{\hat{\Lambda}^*_L}(x,y), \tag{21}$$

where $\mathbb{G}_{\hat\Lambda^*_U}(x,y)$ *and* $\mathbb{G}_{\hat\Lambda^*_L}(x,y)$ *are centered tight continuous Gaussian random fields. Weak convergence takes place in* $\mathcal{B}_\infty(\bar{I\!R}^2_+)$ *and the covariance structure of* $\mathbb{G}_{\hat\Lambda^*_L}(x,y)$ *is given by*

$$I\!E\big(\mathbb{G}_{\hat\Lambda^*_L}(x,y)\cdot\mathbb{G}_{\hat\Lambda^*_L}(\bar x,\bar y)\big) \quad = \quad \Lambda_L\big(\min\{x,\bar x\},\min\{y,\bar y\}\big) \tag{22}$$

for $(x,y)',(\bar x,\bar y)'\in\bar{I\!R}^2_+$. *The same holds for* $\mathbb{G}_{\hat\Lambda^*_U}$ *substituting* Λ_L *by* Λ_U.

Remark. Suppose $A(t)=t^{-\rho},\ \rho>0$, then $k=o(m^{2\rho/(1+2\rho)})$.

Intuition. Note that for $y=\bar y=\infty$ the covariance structure (22) degenerates to $\min\{x,\bar x\}$ which equals the covariance structure of a Brownian motion. At the first glance this seems to be odd since the empirical distribution function converges weakly to a Brownian bridge. However, the fluctuation at infinity cannot decrease for the limiting process of an empirical tail copula since we move into the tail with the threshold k with increasing sample size (Thus, solely taking extreme events into account of our estimation.).

Proof. The claim of weak convergence is proven for the estimator $\hat\Lambda^*_{L,m}(x,y)$. The upper counterpart is treated analogously. Because of (19) it suffices to prove

$$\alpha_m(x,y) := \sqrt{k}\left(\hat\Lambda^*_{L,m}(x,y)-\frac{m}{k}C\Big(\frac{kx}{m},\frac{ky}{m}\Big)\right)\xrightarrow{w}\mathbb{G}_{\hat\Lambda^*_L}(x,y)\quad\text{as } m\to\infty$$

with $k=k(m)\to\infty$ and $k/m\to0$ as $m\to\infty$, and $\mathbb{G}_{\hat\Lambda^*_L}(x,y)$ being a centered tight Gaussian random field or process. Further, weak convergence takes place in $\mathcal{B}_\infty(\bar{I\!R}^2_+)$. We have to verify: Finite dimensional convergence and tightness of the process $\alpha_m(x,y)$.

i) (**Finite dimensional convergence**) We show that the finite-dimensional projections of $\alpha_m(x,y)$ converge in distribution to a normal random vector, i.e., for each finite subset $\{(x_1,y_1),\dots,(x_t,y_t)\}$ of $\bar{I\!R}^2_+$ there exists a centered normal random vector $(\alpha(x_1,y_1),\dots,\alpha(x_t,y_t))$ with appropriate covariance structure such that

$$(\alpha_m(x_1,y_1),\dots,\alpha_m(x_t,y_t))\xrightarrow{d}(\alpha(x_1,y_1),\dots,\alpha(x_t,y_t)).$$

The latter is shown by a multivariate version of the Lindeberg-Feller theorem for triangular arrays, see Durrett (1996), p.116, or Araujo & Giné (1980), p.41. Let $\{(x_1,y_1),\dots,(x_t,y_t)\}$ be an arbitrary but fixed finite subset of $\bar{I\!R}^2_+$. Put

$$Z^{(j)}_{i,m} := \frac{1}{\sqrt{k}}\mathbf{1}_{\big\{G(X^{(j)})\le\frac{kx_i}{m}\text{ and }H(Y^{(j)})\le\frac{ky_i}{m}\big\}}-\frac{1}{\sqrt{k}}C\big(\frac{kx_i}{m},\frac{ky_i}{m}\big)$$

for all $i=1,\dots,t$. Then $I\!E(Z^{(j)}_{i,m})=0$ for all $i=1,\dots,t$. For every $r,s\in\{1,\dots,t\}$

$$\begin{aligned}\sum_{j=1}^m I\!E\big(Z^{(j)}_{r,m}\cdot Z^{(j)}_{s,m}\big)&=\frac{m}{k}\Big\{I\!P\Big(G(X^{(1)})\le\frac{k}{m}\min\{x_r,x_s\}\text{ and }H(Y^{(1)})\le\frac{k}{m}\min\{y_r,y_s\}\Big)\\&\quad-C\left(\frac{kx_r}{m},\frac{ky_r}{m}\right)C\left(\frac{kx_s}{m},\frac{ky_s}{m}\right)\Big\}\to\Lambda_L(\min\{x_r,x_s\},\min\{y_r,y_s\})=:a_{r,s}\text{ as } m\to\infty.\end{aligned}$$

Notice that $t\,C(x_r/t,y_r/t)\,C(x_s/t,y_s/t)\to0$ as $t\to\infty$. The matrix $A=(a_{r,s})_{r,s=1,\dots,t}$ is nonzero if $\Lambda_L\not\equiv0$ according to Theorem 1. Further, for $Z^{(j)}_m=(Z^{(j)}_{1,m},\dots,Z^{(j)}_{t,m})$ and the Euclidian norm $||.||_2$ we have $||Z^{(j)}_m||_2^2=\sum_{i=1}^t\big(Z^{(j)}_{i,m}\big)^2\le t/k$ and thus

$$\sum_{j=1}^m\int_{\{||Z^{(j)}_m||_2>\varepsilon\}}||Z^{(j)}_m||_2^2\,dI\!P\to0\quad\text{as } m\to\infty$$

for every $\varepsilon > 0$. Therefore $(\alpha_m(x_1, y_1), \ldots, \alpha_m(x_t, y_t)) \xrightarrow{d} (\alpha(x_1, y_1), \ldots, \alpha(x_t, y_t)) \sim N(0, A)$ with $A = (a_{r,s})$.

ii) (**Tightness**) First we prove tightness on $[0, M]^2$ for every fixed $M \in \mathbb{N}$ via asymptotic uniform equicontinuity in probability of α_m, i.e, for each $\xi > 0$ and $\eta > 0$ there exist $\delta \in (0, 1)$ and $m_0 \in \mathbb{N}$ such that

$$\mathbb{P}\Big(\sup_{\substack{|x_1 - x_2|^2 + |y_1 - y_2|^2 < \delta \\ x_i, y_i \in [0, M], i = 1, 2}} |\alpha_m(x_1, y_1) - \alpha_m(x_2, y_2)| \geq \xi \Big) \leq \eta \quad \forall m \geq m_0.$$

Note that α_m belongs to the space of cádlág functions $D(\bar{\mathbb{R}}_+^2)$. It can be shown that the ball-σ-field (with respect to the uniform metric on compacta) coincides with the projection σ-field in the space $D(\bar{\mathbb{R}}_+^2)$ and therefore α_m is measurable with respect to the ball-σ-field. This justifies to take probability instead of outer probability (see the Appendix for more details) in the above expression. Tightness is now shown by the following reasoning. Consider a partition of $[0, M]^2$ into equally sized cubes $I_{l,L}$ with partition points $(M \cdot l_1/L, M \cdot l_2/L)$, $l_i \in \{0, \ldots, L\}, L \in \mathbb{N}$, $i = 1, 2$. Then for arbitrary but fixed $\xi > 0$ and $\delta \in (0, 1)$ such that $1/L \geq \delta$ we obtain

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{\substack{|x_1 - x_2|^2 + |y_1 - y_2|^2 < \delta \\ x_i, y_i \in [0, M], i = 1, 2}} |\alpha_m(x_1, y_1) - \alpha_m(x_2, y_2)| \geq \xi \Big) \\
&\quad \leq \mathbb{P}\big(3 \max_{\substack{1 \leq l_i < L \\ i = 1, 2}} \sup_{\binom{x_1}{y_1}, \binom{x_2}{y_2} \in I_{l,L}} |\alpha_m(x_1, y_1) - \alpha_m(x_2, y_2)| \geq \xi \big) =: I_1.
\end{aligned}$$

Without loss of generality we assume $x_1 < x_2$ and $y_1 < y_2$. Then,

$$\begin{aligned}
I_1 \leq & \sum_{\substack{1 \leq l_i < L \\ i = 1, 2}} \Big\{ \mathbb{P}\Big(\sup_{\binom{x_1}{y_1}, \binom{x_2}{y_2} \in I_{l,L}} \sqrt{m} \Big| \frac{1}{m} \sum_{j=1}^{m} \mathbf{1}_{\{ \frac{kx_1}{m} < G(X^{(j)}) \leq \frac{kx_2}{m} \text{ and } H(Y^{(j)}) \leq \frac{ky_2}{m} \}} \\
& \qquad - C\Big(\Big(\frac{kx_1}{m}, \frac{kx_2}{m} \Big] \times \Big[0, \frac{ky_2}{m} \Big] \Big) \Big| \geq \frac{1}{2} \frac{\sqrt{k}}{\sqrt{m}} \frac{\xi}{3} \Big) \\
& \qquad + \mathbb{P}\Big(\sup_{\binom{x_1}{y_1}, \binom{x_2}{y_2} \in I_{l,L}} \sqrt{m} \Big| \frac{1}{m} \sum_{j=1}^{m} \mathbf{1}_{\{ G(X^{(j)}) \leq \frac{kx_1}{m} \text{ and } \frac{ky_1}{m} < H(Y^{(j)}) \leq \frac{ky_2}{m} \}} \\
& \qquad - C\Big(\Big[0, \frac{kx_1}{m} \Big] \times \Big(\frac{ky_1}{m}, \frac{ky_2}{m} \Big] \Big) \Big| \geq \frac{1}{2} \frac{\sqrt{k}}{\sqrt{m}} \frac{\xi}{3} \Big) \Big\} \\
\leq & \sum_{n=1}^{2} \sum_{\substack{1 \leq l_i < L \\ i = 1, 2}} c \cdot \exp \left(- \frac{\eta^2 k \xi^2}{36 m C(A_{n,m}^{l_i,L})} \cdot \psi\Big(\frac{\sqrt{k}\eta}{m C(A_{n,m}^{l_i,L})} \Big) \right) =: I_2,
\end{aligned}$$

where the constants $c, \eta > 0$ are independent of the other parameters, and

$$\begin{aligned}
A_{1,m}^{l_i,L} &:= \Big(M \frac{k(l_1 - 1)}{mL}, M \frac{kl_1}{mL} \Big] \times \Big[0, M \frac{kl_2}{mL} \Big], \text{ and} \\
A_{2,m}^{l_i,L} &:= \Big[0, M \frac{k(l_1 - 1)}{mL} \Big] \times \Big(M \frac{k(l_2 - 1)}{mL}, M \frac{kl_2}{mL} \Big].
\end{aligned}$$

The last inequality is due to Ruymgaart & Wellner (1982), Inequality 1.1. In particular the function $\psi : [-1, \infty) \to \mathbb{R}$ satisfies $\psi(0) = 1$, $\psi(x) \sim (2 \log x)/x \to 0$ as $x \to \infty$, ψ is decreasing and continuous, and $(\cdot)\psi(\cdot)$ is increasing. Observe that $C(A_{n,m}^{l_i,L}) \leq \frac{k}{Lm}$ for all $l_i \in \{1, \ldots, L\}$, $i = 1, 2$.

Distinguish two cases: Either for $m, L \in \mathbb{N}$, $n = 1, 2$, and $l_i \in \{1, \ldots, L\}$, $i = 1, 2$,

$$\frac{\sqrt{k}\eta}{m C(A_{n,m}^{l_i,L})} \leq 1 \quad \text{or} \quad \frac{\sqrt{k}\eta}{m C(A_{n,m}^{l_i,L})} > 1.$$

In the first case an upper bound is provided by

$$I_2 \leq 2L^2 c \cdot \exp\Big(-\frac{\eta^2\xi^2}{36} L\psi(1)\Big), \quad L \in I\!N,$$

whereas in the second case we utilize the upper bound

$$I_2 \leq 2L^2 c \cdot \exp\Big(-\frac{\eta\xi^2}{36} \sqrt{k}\psi(1)\Big) \to 0 \quad \text{as } m \to \infty.$$

This immediately yields tightness on $[0, M]^2$ for every fixed $M \in I\!N$. Tightness on $[0, M] \times \{\infty\}$ and $\{\infty\} \times [0, M]$ is shown along the same lines.

iii) According to part ii), Theorem 1.5.7 and Lemma 1.3.8 in Van der Vaart & Wellner (1996) the sequence of restrictions $\alpha_m|_{T_i}$ with T_i as defined in Definition 4 is asymptotically tight. This follows because the limiting process $\mathbb{G}_{\hat{\Lambda}_L^*}(x,y)|_{T_i}$ is tight in $I\!R$ for every $(x,y)' \in T_i$ as its law is a Borel probability-measure on a Polish space. Hence, the sequence of restrictions $\alpha_m|_{T_i}$ weakly converges to the tight limit $\mathbb{G}_{\hat{\Lambda}_L^*}(x,y)|_{T_i}$ due to part i) and Theorem 1.5.4 in Van der Vaart & Wellner (1996). Finally, weak convergence of α_m in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ is provided by Theorem 8. Continuity of the sample paths in $\bar{I\!R}_+^2$ follows according to the Addendum 1.5.8 in the latter reference.

The covariance structure (22) has been explicitly derived in part i).

Remark. If the tail copula is only defined on some subinterval of $\bar{I\!R}_+^2$, the latter results hold only on this subinterval of $\bar{I\!R}_+^2$.

Theorem 5 (Asymptotic normality under unknown margins G and H) *Let F be a distribution function with continuous marginal distribution functions G and H. If the tail copulae $\Lambda_U \not\equiv 0$ and $\Lambda_L \not\equiv 0$ exist, possess continuous partial derivatives, and the Second order Condition* 1 *holds, then for $\sqrt{k}A(m/k) \to 0$ as $m \to \infty$*

$$\sqrt{k}\big\{\hat{\Lambda}_{U,m}(x,y) - \Lambda_U(x,y)\big\} \xrightarrow{w} \mathbb{G}_{\hat{\Lambda}_U}(x,y) \quad \text{and} \quad \sqrt{k}\big\{\hat{\Lambda}_{L,m}(x,y) - \Lambda_L(x,y)\big\} \xrightarrow{w} \mathbb{G}_{\hat{\Lambda}_L}(x,y),$$

where $\mathbb{G}_{\hat{\Lambda}_U}(x,y)$ and $\mathbb{G}_{\hat{\Lambda}_L}(x,y)$ are centered tight continuous Gaussian random fields. Weak convergence takes place in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ and the limiting process $\mathbb{G}_{\hat{\Lambda}_L}(x,y)$ can be expressed by

$$\mathbb{G}_{\hat{\Lambda}_L}(x,y) \;=\; \mathbb{G}_{\hat{\Lambda}_L^*}(x,y) - \frac{\partial}{\partial x}\Lambda_L(x,y)\mathbb{G}_{\hat{\Lambda}_L^*}(x,\infty) - \frac{\partial}{\partial y}\Lambda_L(x,y)\mathbb{G}_{\hat{\Lambda}_L^*}(\infty,y) \tag{23}$$

with $\mathbb{G}_{\hat{\Lambda}_L^*}$ established in Theorem 4. The same holds for $\mathbb{G}_{\hat{\Lambda}_U}$ replacing Λ_L (and $\hat{\Lambda}_L^*$) by Λ_U (and $\hat{\Lambda}_U^*$).

Proof. The proof is given for the lower empirical tail copula as the upper empirical tail copula can be treated similarly.

First we note that, in contrast to the estimator $\hat{\Lambda}_{L,m}^*(x,y)$, the estimator $\hat{\Lambda}_{L,m}(x,y)$ contains additional fluctuation which is due to the empirical marginal distributions. At the end of the proof, however, we will see that $\hat{\Lambda}_{L,m}(x,y)$ can be expressed as the value of a functional map ϕ evaluated at the 'point' $\hat{\Lambda}_{L,m}^*(x,y)$. Since we have already derived the asymptotic behavior of the estimator $\hat{\Lambda}_{L,m}^*(x,y)$ in Theorem 4, we might thus apply a functional Delta method in order to derive the asymptotic behavior of $\hat{\Lambda}_{L,m}(x,y)$. In other words, a functional Delta method combined with the appropriate mapping ϕ will yield the asymptotic results. The limiting process $\mathbb{G}_{\hat{\Lambda}_L^*}(x,y)$ from Theorem 4 will play a central role in this context.

The space of locally uniformly bounded real functions on compact sets of $I\!R_+$ is denoted by $\mathcal{B}_\infty(I\!R_+)$; the appropriate metric is defined analogously to (31) given in the Appendix A.1. Let $B^I(I\!R_+) \subset \mathcal{B}_\infty(I\!R_+)$ denote the set of all nondecreasing functions $\zeta : I\!R_+ \mapsto I\!R_+$. Define the set

$$\mathcal{B}_\infty^I(\bar{I\!R}_+^2) := \{\gamma \in \mathcal{B}_\infty(\bar{I\!R}_+^2) \mid \gamma(\cdot,\infty) \in B^I(I\!R_+) \text{ and } \gamma(\infty,\cdot) \in B^I(I\!R_+)\}.$$

We apply the Delta method as stated in Theorem 9 in Appendix A.1 (with $r_m = \sqrt{k(m)} = \sqrt{k}$) to the following map

$$\phi : \mathcal{B}_\infty^I(\bar{I\!R}_+^2) \mapsto \mathcal{B}_\infty(\bar{I\!R}_+^2).$$

The Delta method involves a first-order Taylor approximation (in our particular functions space $\mathcal{B}_\infty(\bar{I\!R}_+^2)$) evaluated at the (function) point $\Lambda_L(x,y)$ in "direction" $\mathbb{G}_{\hat{\Lambda}_L^*}(x,y)$.

For the precise definition of ϕ we need some additional notation: Let ζ^- denote the adjusted generalized inverse function of $\zeta \in B^I(I\!R_+)$ defined by

$$\zeta^-(p) := \begin{cases} \zeta^{-1}(p) & \text{if} \quad \zeta^{-1}(p) < \infty, \\ \lim_{z\to\infty} \zeta(z) & \text{if} \quad \zeta^{-1}(p) = \infty, \end{cases}$$

where ζ^{-1} refers to the generalized inverse function. Split the set $\bar{I\!R}_+^2$ into three subsets $S_1 := I\!R_+^2$, $S_2 := [0,\infty) \times \{\infty\}$, and $S_3 := \{\infty\} \times [0,\infty)$. For some arbitrary function $\gamma \in \mathcal{B}_\infty^I(\bar{I\!R}_+^2)$ the map ϕ is defined for $(x,y)' \in S_1$ by

$$\phi : \gamma(x,y) \stackrel{\phi_1}{\mapsto} (\gamma(x,y), \gamma(x,\infty), \gamma(\infty,y))$$

$$\stackrel{\phi_2}{\mapsto} (\gamma(x,y), \gamma^-(x,\infty), \gamma^-(\infty,y)) \stackrel{\phi_3}{\mapsto} \gamma \circ (\gamma^-(x,\infty), \gamma^-(\infty,y)),$$

for $(x,y)' \in S_2$ by

$$\phi : \gamma(x,y) \stackrel{\phi_1}{\mapsto} (\gamma(x,y), \gamma(x,\infty), \gamma(x,\infty))$$

$$\stackrel{\phi_2}{\mapsto} (\gamma(x,y), \gamma^-(x,\infty), \gamma^-(x,\infty)) \stackrel{\phi_3}{\mapsto} \gamma \circ (\gamma^-(x,\infty), \infty),$$

and for $(x,y)' \in S_3$ by

$$\phi : \gamma(x,y) \stackrel{\phi_1}{\mapsto} (\gamma(x,y), \gamma(\infty,y), \gamma(\infty,y))$$

$$\stackrel{\phi_2}{\mapsto} (\gamma(x,y), \gamma^-(\infty,y), \gamma^-(\infty,y)) \stackrel{\phi_3}{\mapsto} \gamma \circ (\infty, \gamma^-(\infty,y)).$$

The spaces $C(\bar{I\!R}_+^2) \subset \mathcal{B}_\infty(\bar{I\!R}_+^2)$ and $C(I\!R_+) \subset \mathcal{B}_\infty(I\!R_+)$ consist of all continuous functions in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ and $\mathcal{B}_\infty(I\!R_+)$, respectively. In order to apply the Delta method we have to show that the map ϕ is Hadamard-differentiable on $\mathcal{B}_\infty^I(\bar{I\!R}_+^2)$ at $\gamma_0 = \Lambda_L$ tangentially to $C(\bar{I\!R}_+^2) \subset \mathcal{B}_\infty(\bar{I\!R}_+^2)$. Note that the Delta method (Theorem 9) involves a directional derivative on the function space $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ where the "direction" is a member of the function space $C(\bar{I\!R}_+^2) \subset \mathcal{B}_\infty(\bar{I\!R}_+^2)$.

Let us first outline the next steps: We will show that the map ϕ is Hadamard-differentiable by proving that each intermediate map ϕ_1, ϕ_2 and ϕ_3 is Hadamard-differentiable and finally applying the chain rule for Hadamard-derivatives. Differentiability of ϕ_1 and ϕ_3 follows easily, but for ϕ_2 we have to make an effort. Essentially we have to show that the adjusted generalized inverse map $\gamma(\cdot,\infty) \to \gamma^-(\cdot,\infty)$ is Hadamard-differentiable at the point $\Lambda_L(\cdot,\infty)$ and $\Lambda_L(\infty,\cdot)$, respectively. Differentiability will be proven directly by verifying the existence of the corresponding limit (cf. (32) in Definition 5). The calculation depends crucially on the fact that differentiability is established tangentially or directionally to a set of continuous functions, since we will need to work out a certain form of uniform convergence/boundedness on compacta (convergence of functions).

i) The first map ϕ_1 is Hadamard-differentiable on $\mathcal{B}_\infty^I(\bar{I\!R}_+^2)$ at Λ_L tangentially to $C(\bar{I\!R}_+^2)$ as it is linear and continuous.

ii) The second map ϕ_2 is Hadamard-differentiable on $\mathcal{B}_\infty^I(\bar{I\!R}_+^2) \times B^I(I\!R_+) \times B^I(I\!R_+)$ at $(\Lambda_L, \mathrm{id}_{I\!R_+}, \mathrm{id}_{I\!R_+})$ tangentially to $C(\bar{I\!R}_+^2) \times C(I\!R_+) \times C(I\!R_+)$. Note that Hadamard differentiability of ϕ_2 is equivalent to Hadamard differentiability of the respective (vector) components of ϕ_2. The first (vector) component of ϕ_2 is Hadamard-differentiable as it represents the identity map on $\mathcal{B}_\infty^I(\bar{I\!R}_+^2)$. The second and third (vector) components are Hadamard-differentiable because the adjusted generalized inverse maps

$$\gamma(\cdot,\infty) \to \gamma^-(\cdot,\infty) \quad \text{and} \quad \gamma(\infty,\cdot) \to \gamma^-(\infty,\cdot)$$

are Hadamard-differentiable on $B^I(I\!R_+)$ at $\mathrm{id}_{I\!R_+}$ tangentially to $C(I\!R_+)$ as we will show next. Note that

$$\gamma_0(\cdot,\infty) = \gamma_0(\infty,\cdot) = \Lambda_L(\cdot,\infty) = \Lambda_L(\infty,\cdot) = \mathrm{id}_{I\!R_+}$$

corresponds to the identity function on R_+.

We restrict ourselves to the map $\gamma(\cdot,\infty) \to \gamma^-(\cdot,\infty)$. Set $\bar{\gamma}_0(\cdot) := \gamma_0(\cdot,\infty) = \mathrm{id}_{I\!R_+}$. Consider the sequence $h_t \to h$ as $t \to 0$ where $h \in C(I\!R_+)$ (i.e., h is a continuous function on $I\!R_+$ and $h_t \in \mathcal{B}_\infty(I\!R_+)$) such that $\bar{\gamma}_0 + th_t \in B^I(I\!R_+)$ for every t . Let $p \in I\!R_+$ then $\bar{\gamma}_0^-(p) = p$. Abbreviate $(\bar{\gamma}_0 + th_t)^-(p)$ to ξ_{pt} and notice that $\xi_{pt} \in I\!R_+$. Setting $\varepsilon_{pt} := t^2 \wedge \xi_{pt} \geq 0$ yields for $p \in [0, M_t)$ with $M_t := \lim_{z\to\infty}(\bar{\gamma}_0 + th_t)(z)$

$$(\bar{\gamma}_0 + th_t)(\xi_{pt} - \varepsilon_{pt}) \leq p \leq (\bar{\gamma}_0 + th_t)(\xi_{pt}).$$

Further $\bar{\gamma}_0(\xi_{pt}) = \xi_{pt}$ and $\bar{\gamma}_0(\xi_{pt} - \varepsilon_{pt}) = \xi_{pt} - \varepsilon_{pt}$ for all $p \in I\!R_+$. Thus it follows that

$$-th(\xi_{pt}) + o(t) \leq \xi_{pt} - p \leq -th(\xi_{pt} - \varepsilon_{pt}) + o(t), \tag{24}$$

where the $o(t)$-terms are uniform in $p \in [0, M_t)$. Note that $M_t \to \infty$ as $t \to 0$ because $h_t \to h$ with $h \in C(I\!R_+)$.

Finally, $h(\xi_{pt}) \to h(p)$ and $h(\xi_{pt} - \varepsilon_{pt}) \to h(p)$ uniformly in $p \in I\!R_+$ because h is continuous on R_+ and $\xi_{pt} \to p$ uniformly in $p \in I\!R_+$ on $\mathcal{B}_\infty(I\!R_+)$. The last claim is proven if we show that $|\xi_{pt} - p| = O(t)$ uniformly in p on the interval $[0, T]$ for arbitrary but fixed T. According to (24), it suffices to show that $h(\xi_{pt})$ is uniformly bounded on $[0, T]$. However, the function h is continuous on $I\!R_+$, therefore, we must prove that ξ_{pt} is uniformly bounded on $[0, T]$. This follows by the definition of the adjusted generalized inverse function and the fact that $h_t \to h$ as $t \to 0$ (h is continuous) on $\mathcal{B}_\infty(I\!R_+)$.

Hence Hadamard differentiability of $\gamma(\cdot,\infty) \to \gamma^-(\cdot,\infty)$ holds and its derivative at $\mathrm{id}_{I\!R_+}$ is given by the linear map $h \mapsto -h$.

iii) The third map ϕ_3 (composition map) is Hadamard-differentiable on $\mathcal{B}^I_\infty(\bar{I\!R}^2_+) \times B^I(I\!R_+) \times B^I(I\!R_+)$ at $(\Lambda_L, \mathrm{id}_{I\!R_+}, \mathrm{id}_{I\!R_+})$ tangentially to $C(\bar{I\!R}^2_+) \times C(I\!R_+) \times C(I\!R_+)$ according to Lemma 1 stated in the Appendix. Uniform Fréchet differentiability in Lemma 1 is implied by the continuous partial derivatives of Λ_L which yield (uniformly) continuous differentiability of Λ_L with respect to the metric (31), cf. Heuser (2000) , Satz 164.4, and Van der Vaart & Wellner (1996), Problem 1, p.397.

iv) Hadamard differentiability of ϕ on $\mathcal{B}^I_\infty(\bar{I\!R}^2_+)$ at $\gamma_0 = \Lambda_L$ tangentially to $C(\bar{I\!R}^2_+) \subset \mathcal{B}_\infty(\bar{I\!R}^2_+)$ follows now with the chain rule (Lemma 3.9.3. in Van der Vaart & Wellner (1996)).

v) The final steps link the Delta method to the desired weak convergence result. Note that $\gamma_0(x, y) = \Lambda_L(x, y)$ and the paths of $\hat{\Lambda}_{L,m}(x, y) \in \mathcal{B}^I_\infty(\bar{I\!R}^2_+)$ can be (almost surely) decomposed into

$$\hat{\Lambda}^*_{L,m}\Big(\frac{m}{k} G\big(G_m^{-1}\big(\frac{k}{m}x\big)\big), \frac{m}{k} H\big(H_m^{-1}\big(\frac{k}{m}y\big)\big)\Big)$$

with $\frac{m}{k} G\big(G_m^{-1}\big(\frac{k}{m}x\big)\big)$ and $\frac{m}{k} H\big(H_m^{-1}\big(\frac{k}{m}y\big)\big)$ being the adjusted generalized inverse functions (empirical quantile functions) of the margins $\hat{\Lambda}^*_{L,m}(x,\infty)$ and $\hat{\Lambda}^*_{L,m}(\infty,\, y)$. In other words, we can apply the Delta method as stated in Theorem 9 with the above defined map ϕ, $r_m = \sqrt{k(m)} = \sqrt{k}$, $\theta = \Lambda_L(x, y)$, and $X_m = \hat{\Lambda}^*_{L,m}(x, y)$ (consult the latter theorem for the meaning of the respective variables). At this point we also use the fact that the limiting process X is known and corresponds to $\mathbb{G}_{\hat{\Lambda}^*_L}$ in Theorem 4.

The structure of the limiting process (23) follows from the Delta method (see Theorem 9 in the Appendix). Recall that the Delta method involves a first-order Taylor approximation (in the functions space $\mathcal{B}_\infty(\bar{I\!R}^2_+)$) evaluated at the (function) point $\Lambda_L(x, y)$ in "direction" $\mathbb{G}_{\hat{\Lambda}^*_L}(x, y)$. In other words, $\mathbb{G}_{\hat{\Lambda}^*_L}(x, y)$ serves as the direction of the directional derivatives which appears in the

Delta method leading to the limiting process $\mathbb{G}_{\hat{\Lambda}_L}(x,y)$. The partial derivatives in formula (23) belong to the gradient of $\Lambda_L(x,y)$ which arises from Lemma 1 stated in the Appendix (cf. proof below). This Lemma basically transfers the chain rule for directional derivatives of the Euclidian space to our function space (in the context of Hadamard differentiability tangentially or directionally to a specific set, cf. Definition 5).

After some calculation we obtain that the derivative ϕ'_{Λ_L} of the map ϕ (at the point Λ_L) is of the form

$$\phi'_{\Lambda_L}(\gamma)(x,y) = \gamma(x,y) - \frac{\partial}{\partial x}\Lambda_L(x,y)\gamma(x,\infty) - \frac{\partial}{\partial y}\Lambda_L(x,y)\gamma(\infty,y).$$

The latter formula utilizes the specific form of the Hadamard-derivative of the inverse operator $\gamma(\cdot,\infty) \to \gamma^{-}(\cdot,\infty)$ at $\mathrm{id}_{\mathbb{R}_+}$ which is the linear map $h \to -h$ (cf. the derivative of the map ϕ_2). Further, the derivative of the map ϕ_3 is derived via Lemma 1 with the following notation: $\alpha(x,y) = (\gamma(x,\infty), \gamma(\infty,y))$ and $\beta(x,y) = \gamma(x,y)$. The latter derivative is then evaluated at the points $B(x,y) = \Lambda(x,y)$ and $A(x,y) = (\Lambda^{-}(x,\infty), \Lambda^{-}(\infty,y)) = (\mathrm{id}_{\mathbb{R}_+}, \mathrm{id}_{\mathbb{R}_+})$. Finally we apply the chain rule as stated in Lemma 3.9.3 in Van der Vaart & Wellner (1996). The evaluation of the derivative ϕ'_{Λ_L} in "direction" of the process $\mathbb{G}_{\hat{\Lambda}^*_L}$ finalizes the proof.

Theorem 6 (Strong consistency) *Let F be a distribution function with continuous marginal distribution functions G and H. If the tail copulae $\Lambda_U \neq 0$ and $\Lambda_L \neq 0$ exist and $k/\log\log m \to \infty$ as $m \to \infty$ then $\hat{\Lambda}_{U,m}$ converges almost surely to Λ_U and $\hat{\Lambda}_{L,m}$ converges almost surely to Λ_L in the space $\mathbb{B}_\infty(\bar{\mathbb{R}}^2_+)$ (equipped with the metric d as in* (31)*) In particular*

$$\mathbb{P}\big(\lim_{m\to\infty} d(\hat{\Lambda}_{U,m}, \Lambda_U) = 0\big) = 1 \quad \textit{and} \quad \mathbb{P}\big(\lim_{m\to\infty} d(\hat{\Lambda}_{L,m}, \Lambda_L) = 0\big) = 1. \tag{25}$$

Proof. The proof is provided for the upper tail copula. Recall that a sequence converges in the space $\mathbb{B}_\infty(\bar{\mathbb{R}}^2_+)$ with respect to the metric d (cf. (31)) if the sequence converges uniformly on each compact subset T_i introduced in Definition 4. Let $T > 0$ be an arbitrary but fixed constant. The conclusion follows now with the strong consistency result for empirical stable tail-dependence functions given in Theorem 1.1 in Qi (1997) and the relationship $\Lambda_U(x,y) = x+y-l(x,y)$. Further, we utilize the fact that

$$|\hat{\Lambda}_{U,m}(x,y) - \Lambda_U(x,y)| = \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m1} > m-kx \text{ and } R^{(j)}_{m2} > m-ky\}} - \Lambda_U(x,y)\Big| \leq$$

$$\begin{aligned}
&\leq \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m1} > m-kx \text{ or } R^{(j)}_{m2} > m-ky\}} - l(x,y)\Big| \\
&+ \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m1} > m-kx\}} - x\Big| + \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m2} > m-ky\}} - y\Big| \\
&\leq \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m1} > m-kx \text{ or } R^{(j)}_{m2} > m-ky\}} - l(x,y)\Big| + \frac{2}{k}.
\end{aligned}$$

The proof for the lower tail copula is similar.

Corollary 1 (Asymptotic normality of $\hat{\lambda}_{U,m}$ and $\hat{\lambda}_{L,m}$) *With the prerequisites of Theorem* 5

$$\sqrt{k}\{\hat{\lambda}_{U,m} - \lambda_U\} \xrightarrow{d} \mathcal{N}_{0,\sigma^2_U} \quad \textit{and} \quad \sqrt{k}\{\hat{\lambda}_{L,m} - \lambda_L\} \xrightarrow{d} \mathcal{N}_{0,\sigma^2_L},$$

where $\mathcal{N}_{0,\sigma^2_L}$ and $\mathcal{N}_{0,\sigma^2_U}$ are centered Gaussian random variables with variances

$$\begin{aligned}
\sigma^2_L &= \lambda_L + \Big(\frac{\partial}{\partial x}\Lambda_L(1,1)\Big)^2 + \Big(\frac{\partial}{\partial y}\Lambda_L(1,1)\Big)^2 \\
&+ 2\lambda_L\Big(\Big(\frac{\partial}{\partial x}\Lambda_L(1,1) - 1\Big)\Big(\frac{\partial}{\partial y}\Lambda_L(1,1) - 1\Big) - 1\Big)
\end{aligned} \tag{26}$$

and analogue variance σ_U^2.

Proof. Note that e.g. for the lower tail copula Λ_L we know from Theorem 5 that

$$\begin{aligned} I\!E\mathbb{G}^2_{\hat{\Lambda}_L}(x,y) &= \Lambda_L(x,y) + \Big(\frac{\partial}{\partial x}\Lambda_L(x,y)\Big)^2 x + \Big(\frac{\partial}{\partial y}\Lambda_L(x,y)\Big)^2 y \\ &+ 2\Lambda_L(x,y)\Big(\Big(\frac{\partial}{\partial x}\Lambda_L(x,y)-1\Big)\Big(\frac{\partial}{\partial y}\Lambda_L(x,y)-1\Big)-1\Big). \end{aligned}$$

Remarks. i) From formula (26) it becomes clear that the local behavior of the tail copula at the point $(1,1)$ determines the asymptotic variance of $\hat{\lambda}_{U,m}$ and $\hat{\lambda}_{L,m}$. Thus, the embedding of the tail-dependence coefficients into the concept of tail copulae was necessary in order to obtain the above corollary.

ii) A sole consideration of the tail copula at the point $(1,1)$, as it is the case for the tail-dependence coefficient by Definition 2, should be considered with care. Although the (tail) copula is invariant with respect to scaling or any strictly increasing transformation of the marginal distributions of X and Y, the point $\Lambda(1,1)$ only determines the tail copula on its diagonal (by the homogeneity property).

Example. To illustrate the previous results we calculate the tail copula Λ_L and the asymptotic variance σ_L^2 in (26) for the well-known Pareto copula.

The Pareto copula $C(u,v)$ is given by

$$C(u,v) = \max\Big([u^{-\theta}+v^{-\theta}-1]^{-1/\theta}, 0\Big), \quad \theta \in [-1,\infty)\backslash\{0\}.$$

It can be shown that the Pareto copula is lower tail-dependent with lower tail-dependence coefficient $\lambda_L = 2^{-1/\theta}$ for $\theta > 0$. Further, the lower tail copula exists for $\theta > 0$ and can be expressed by

$$\Lambda_L(x,y) = \big(x^{-\theta}+y^{-\theta}\big)^{-1/\theta}.$$

Thus, the partial derivatives are

$$\frac{\partial}{\partial x}\Lambda_L(x,y) = \big(x^{-\theta}+y^{-\theta}\big)^{-((1/\theta)+1)}x^{-(\theta+1)} \text{ and } \frac{\partial}{\partial y}\Lambda_L(x,y) = \big(x^{-\theta}+y^{-\theta}\big)^{-((1/\theta)+1)}y^{-(\theta+1)}.$$

Consequently, the asymptotic variance σ_L^2 in (26) is given by (see also Figure 1)

$$\sigma_L^2(\theta) = 2^{-1/\theta} - \frac{3}{2}4^{-1/\theta} + \frac{1}{2}8^{-1/\theta}. \tag{27}$$

As often in nonparametric statistics the asymptotic variance of e.g. $\hat{\lambda}_{L,m}$ depends on the derivative of an unknown function and has to be estimated. In our case one could estimate the derivatives of the tail copula by some smoothing method, but since we do not have too many data in the tails, we do not recommend this. An alternative and appealing method is to find a simple but flexible parametric copula as the Pareto copula, calculate its tail copula, and utilize the corresponding variance functional $\sigma_L^2(\theta)$ as an approximation for the unknown asymptotic variance. In the simulation study we applied the proposed method and estimated $\sigma_L^2(\theta) = 2^{-1/\theta} - \frac{3}{2}4^{-1/\theta} + \frac{1}{2}8^{-1/\theta}$ via $\sigma_L^2(\hat{\theta})$, where we replaced θ by the MLE $\hat{\theta}$. Several simulation results are explained in Section 7.

6 General rank order statistics for extreme events

In the present section we extent the main asymptotic results (stated in Theorem 5) in the framework of general rank order statistics of extreme events. We restrict ourselves to the lower tail copula Λ_L.

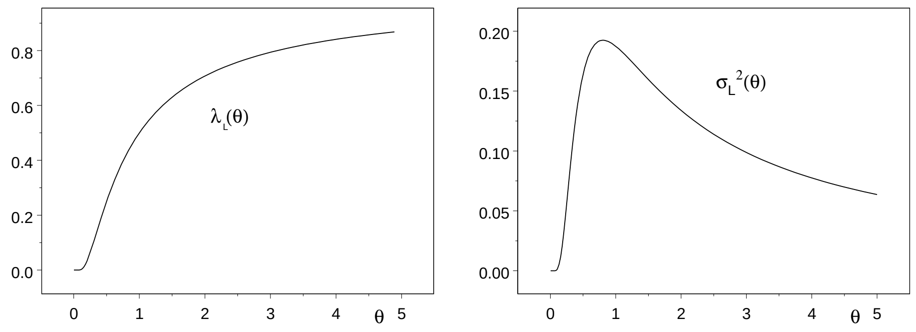


Figure 1: Lower tail-dependence coefficient $\lambda_L(\theta)$ (left plot) and corresponding asymptotic variance $\sigma_L^2(\theta)$ as in formula (26) (right plot) for the Pareto copula.

Rank order statistics of the type

$$\frac{1}{m}\sum_{j=1}^{m} J\big(R_{m1}^{(j)}/m, R_{m2}^{(j)}/m\big) = \frac{1}{m}\sum_{j=1}^{m} J(G_m(X^{(j)}), H_m(Y^{(j)})) \tag{28}$$

have been investigated, for example, by Ruymgaart, Shorack & van Zwet (1972), Ruymgaart (1974) and Rüschendorf (1976). Recently, Fermanian et al. (2004) considered general rank order statistics in the framework of empirical copula processes.

In the context of (lower) tail copulae, a similar family of multivariate (lower) rank order statistics can be investigated:

$$\mathcal{R}_m := \frac{1}{k}\sum_{j=1}^{m} J\Big(R_{m1}^{(j)}/k, R_{m2}^{(j)}/k\Big).$$

In this formula the function J is defined on $\bar{I\!R}_+^2$, whereas in formula (28) the function J has domain $[0,1]^2$. The next theorem establishes asymptotic normality of $\mathcal{R}_m$ under certain regularity assumptions on J.

Theorem 7 *Let F be a distribution function with continuous marginal distribution functions G and H. Suppose that the (lower) tail copula $\Lambda_L \not\equiv 0$ exists and possesses continuous partial derivatives. Assume that $J : \bar{I\!R}_+^2 \to I\!R$ is of bounded variation, continuous from above with discontinuities of the first kind Neuhaus (1971), and bounded on $\bar{I\!R}_+^2$. Then*

$$\frac{1}{\sqrt{k}}\sum_{j=1}^{m}\left(J\Big(\frac{R_{m1}^{(j)}}{k}, \frac{R_{m2}^{(j)}}{k}\Big) - I\!E J\Big(\frac{m}{k}G(X^{(j)}), \frac{m}{k}H(Y^{(j)})\Big)\right) \xrightarrow{w} \int_{\bar{I\!R}_+^2} \mathbb{G}_{\Lambda_L}(x,y)\,dJ(x,y), \tag{29}$$

where $\mathbb{G}_{\Lambda_L}$ equals the limiting process in Theorem 5 and weak convergence takes place in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$. Moreover, the limiting process is also a centered Gaussian field.

Proof. With the stated prerequisites we have

$$\begin{aligned}&\frac{1}{\sqrt{k}}\sum_{j=1}^{m}\left(J\Big(\frac{R_{m1}^{(j)}}{k}, \frac{R_{m2}^{(j)}}{k}\Big) - I\!E J\Big(\frac{m}{k}G(X^{(j)}), \frac{m}{k}H(Y^{(j)})\Big)\right)\\ &\quad = \frac{m}{\sqrt{k}}\int_{[0,1]^2} J\Big(\frac{m}{k}u, \frac{m}{k}v\Big)d(\check{C}_m - C)(u,v) =: I_1,\end{aligned}$$

where $\check{C}_m(u,v) = \frac{1}{m}\sum_{j=1}^m \mathbf{1}_{\{G_m(X^{(j)})\le u \text{ and } H_m(Y^{(j)})\le v\}}$ denotes the modified empirical copula process. Utilizing the integration by parts formula, given in Baron, Liflyand & Stadtmüller (2000), yields

$$\begin{aligned} I_1 &= \int_{[0,1]^2} \frac{m}{\sqrt{k}}\Big((\check{C}_m - C)(u-,v-)\Big)dJ\Big(\frac{m}{k}u,\frac{m}{k}v\Big) \\ &- \int_{[0,1]} \frac{m}{\sqrt{k}}((\check{C}_m - C)(u-,1))dJ\Big(\frac{m}{k}u,1\Big) \\ &- \int_{[0,1]} \frac{m}{\sqrt{k}}((\check{C}_m - C)(1,v-))dJ\Big(1,\frac{m}{k}v\Big) =: I_2. \end{aligned}$$

Substituting $x = \dfrac{m}{k}v$ and $y = \dfrac{m}{k}u$ provides

$$\begin{aligned} I_2 &= \int_{[0,m/k]^2} \frac{m}{\sqrt{k}}\Big((\check{C}_m - C)\Big(\frac{k}{m}x-,\frac{k}{m}y-\Big)\Big)dJ(x,y) \\ &- \int_{[0,m/k]} \frac{m}{\sqrt{k}}\Big((\check{C}_m - C)\Big(\frac{k}{m}x-,1\Big)\Big)dJ(x,1) \\ &- \int_{[0,m/k]} \frac{m}{\sqrt{k}}\Big((\check{C}_m - C)\Big(1,\frac{k}{m}y-\Big)\Big)dJ(1,y) =: I_3. \end{aligned}$$

Notice that

$$\begin{aligned} \frac{m}{\sqrt{k}}\big((\check{C}_m - C)(\frac{k}{m}x-,1)\big) &= \sqrt{k}\big(\hat{\Lambda}_{L,m}(x-,\infty) - x\big) \\ &= \sqrt{k}\big(\frac{1}{k}[kx-] - x\big) \in [0,1/\sqrt{k}]. \end{aligned}$$

Thus,

$$\frac{m}{\sqrt{k}}\big((\check{C}_m - C)(\frac{k}{m}x-,1)\big) = O\Big(\frac{1}{\sqrt{k}}\Big).$$

The expression $\frac{m}{\sqrt{k}}\big((\check{C}_m - C)(1,\frac{k}{m}y-)\big)$ possesses the same property. Therefore, the continuous mapping theorem (Van der Vaart & Wellner (1996), Theorem 1.3.6) leads to

$$I_3 \xrightarrow{w} \int_{\bar{\mathbb{R}}_+^2} \mathbb{G}_{\Lambda_L}(x-,y-)dJ(x,y) = \int_{\bar{\mathbb{R}}_+^2} \mathbb{G}_{\Lambda_L}(x,y)dJ(x,y)$$

which is centered Gaussian.

7 Simulation and empirical study

The simulation study aims at three different types of questions. First, we analyze the finite-sample behavior of the nonparametric estimators for the (lower or upper) tail-dependence coefficient (in short: TDC) which have been considered in Section 4. Second, we construct asymptotic confidence intervals related to one particular TDC estimator and discuss the applicability of the proposed approximation method. Finally, we investigate the finite-sample behavior of the nonparametric estimator for the tail copula, i.e., the empirical tail copula. Moreover, the TDCs and the tail copulae are estimated for two financial time series.

7.1 Comparison of nonparametric TDC estimators

Consider 1000 independent copies of $m = 500, 1000,\ 2000$ iid pseudo-random vectors which are generated from a bivariate standard t-distribution with $\nu = 1.5, 2, 3$ degrees of freedom, i.e., a spherically contoured t-distribution with density generator $g(u) = c\,(1+u/\nu)^{-(1+\nu/2)}$ (for a detailed discussion of the latter and various applications to finance we refer to Bingham, Kiesel & Schmidt (2003)). We restrict ourselves to the estimation of the upper TDC. The empirical bias and mean-squared error (MSE) for all implemented TDC estimations are derived and presented in Tables 1 and 2.

Table 1: Sample-bias and MSE for the nonparametric upper TDC estimator $\hat{\lambda}_U$ (For notational convenience we drop the index m representing the sample length).

Original parameters	$\nu = 1.5$ $\lambda_U = 0.2296$	$\nu = 2$ $\lambda_U = 0.1817$	$\nu = 3$ $\lambda_U = 0.1161$
Estimator	$\hat{\lambda}_U$ Bias (MSE)	$\hat{\lambda}_U$ Bias (MSE)	$\hat{\lambda}_U$ Bias (MSE)
$m = 500$	0.0255 (0.00369)	0.0434 (0.00530)	0.0718 (0.00858)
$m = 1000$	0.0151 (0.00223)	0.0287 (0.00306)	0.0518 (0.00466)
$m = 2000$	0.0082 (0.00149)	0.0191 (0.00169)	0.0369 (0.00270)

Table 2: Sample-bias and MSE for the nonparametric upper TDC estimator $\hat{\lambda}_U^{EVT}$ (For notational convenience we drop the index m representing the sample length).

Original parameters	$\nu = 1.5$ $\lambda_U = 0.2296$	$\nu = 2$ $\lambda_U = 0.1817$	$\nu = 3$ $\lambda_U = 0.1161$
Estimator	$\hat{\lambda}_U^{EVT}$ Bias (MSE)	$\hat{\lambda}_U^{EVT}$ Bias (MSE)	$\hat{\lambda}_U^{EVT}$ Bias (MSE)
$m = 500$	0.0539 (0.00564)	0.0703 (0.00777)	0.1031 (0.01354)
$m = 1000$	0.0333 (0.00301)	0.0491 (0.00437)	0.0748 (0.00744)
$m = 2000$	0.0224 (0.00173)	0.0329 (0.00228)	0.0569 (0.00436)

We conclude that the TDC estimator $\hat{\lambda}_U$ outperforms the estimator $\hat{\lambda}_U^{EVT}$ with respect to the sample-bias and MSE. For example, for $m = 2000$ the bias of $\hat{\lambda}_U^{EVT}$ is two times larger than the bias of $\hat{\lambda}_U$ whereas the MSE is one and a half times larger. The larger bias of $\hat{\lambda}_U^{EVT}$ reflects the additional uncertainty induced by the unknown marginal distribution functions. Regarding the finite-sample variances, both types of estimators behave similarly. Further we observe that the empirical bias and variance decrease with increasing sample size m. Moreover, the bias of the latter estimators increases with growing correlation.

Choosing the threshold k. The algorithm to choose the threshold k utilizes the homogeneity property of tail copulae as stated in Theorem 1 part ii) which corresponds to a balancing of the variance-bias problem. For sufficiently large data sets, this homogeneity property transfers to the nonparametric estimators yielding a characteristic plateau while plotting the estimates for successive k (cf. the well-known Hill plot for the Hill estimator). The optimal threshold k is now estimated via a simple plateau-finding algorithm after smoothing the latter plot by some box kernel. The results of the proposed algorithm are quite satisfying according to our simulation study.

7.2 Confidence intervals

Recall the discussion at the end of Section 5: The estimation of the asymptotic variance or standard deviation of $\hat{\lambda}_L$ (and $\hat{\lambda}_U$) depends on the parameters λ_L (and λ_U) itself and involves certain derivatives of the tail copula. We restrict ourselves to the lower TDC λ_L. Unfortunately the estimation of the latter derivatives turns out to be quite sensitive. Therefore, for arbitrary copulae we proposed to construct an approximation $\sigma_L(\theta)$ of the true asymptotic standard deviation σ_L via formula (27). The parameter θ is received from an ML-fitted Pareto copula, i.e., we utilize $\sigma_L(\hat{\theta})$ where $\hat{\theta}$ denotes the ML-estimate of θ. The quality of the approximation is investigated for the t-copula.

Consider 200 independent copies of $m = 500$ iid pseudo-random vectors which are generated from a bivariate t-copula with various parameters ν and correlation coefficient $\rho = 0.25$. The corresponding sample-means of the nonparametric estimator $\hat{\lambda}_L$ and the approximated confidence intervals are presented in Figure 2. The threshold k is chosen according to the plateau algorithm described above. Note that there is an increasing bias for larger values of the parameters ν which results in an asymmetric confidence band. The estimation improves (in the sense of a smaller sample bias) with increasing sample size. The results are characteristic for all other estimations which we have not listed in this work.

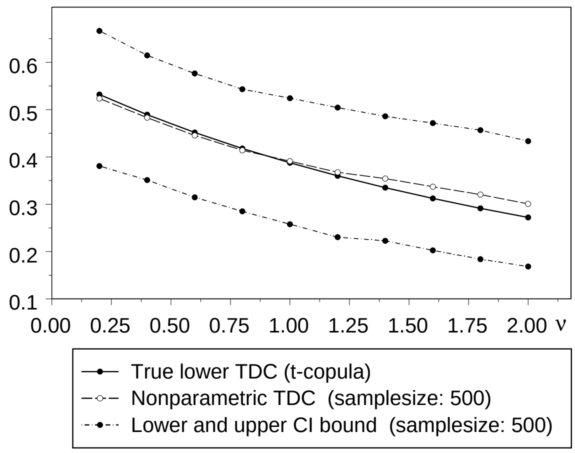


Figure 2: True lower TDC for the t-copula for various parameters ν and correlation coefficient $\rho = 0.25$, the corresponding sample-means of the nonparametric estimator $\hat{\lambda}_L$, and the sample-means of the approximated confidence intervals (confidence level $\alpha = 0.05$) for sample size $m = 500$.

The quality of the approximation of the true confidence interval is illustrated in Table 3. For almost all parameter constellations ν, the sample-means of the approximated asymptotic standard deviations are below the corresponding sample standard deviations $\hat{\sigma}_L$ which are disturbed by the sample-bias. The last column of Table 3, which represents the percentage of the approximated confidence intervals containing the real TDC, shows very satisfying results. These results justify the usage of the approximated asymptotic standard deviation $\sigma_L(\theta)$ even if the copula is not the Pareto copula. Note that an increasing correlation ρ deteriorates the latter results because of the increasing bias of the nonparametric TDC estimator. However, for most applications (especially in finance) a correlation of $\rho = 0.25$ is quite common. A sensitivity analysis regarding different choices of the threshold k has been also implemented. However, it has been omitted in order to shorten the presentation and can be obtained from the authors on request.

7.3 Estimation of the tail copula

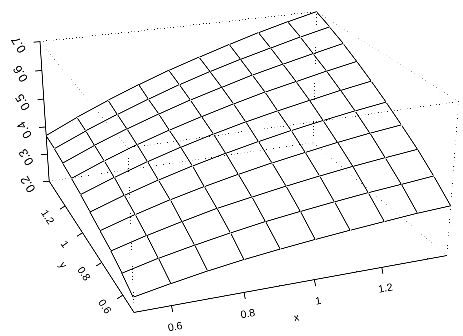
So far we have concentrated on estimating the TDC via the nonparametric estimators $\hat{\lambda}_L$ and $\hat{\lambda}_U$ given in (16). Now we turn to the estimation of the entire tail copula utilizing the proposed non-parametric estimators. More precisely, we consider the estimation of the lower tail copula belonging to a Pareto copula via the estimator $\hat{\Lambda}_L(x, y)$ stated in (12). For the simulation experiment we uti-

Table 3: Various estimation results using simulated data generated from a bivariate t-copula with various parameters ν and correlation coefficient $\rho = 0.25$ (the confidence level is $\alpha = 0.05$).

ν	λ_L	$mean(\hat{\lambda}_L)$	$\hat{\sigma}_L$	$mean(\sigma(\hat{\theta}))$	% of $\lambda_L \in [\hat{\lambda}_L \pm 1.64 \frac{\sigma_L(\hat{\theta})}{\sqrt{k}}]$
0.2	0.532	0.523	0.43	0.433	1
0.3	0.51	0.506	0.444	0.416	0.974
0.4	0.489	0.483	0.466	0.371	0.954
0.6	0.452	0.44	0.427	0.349	0.966
0.8	0.418	0.411	0.47	0.343	0.921
1	0.388	0.384	0.458	0.34	0.923
1.2	0.36	0.366	0.428	0.328	0.968
1.4	0.335	0.351	0.421	0.339	0.964
1.6	0.312	0.327	0.423	0.33	0.976
1.8	0.291	0.318	0.44	0.313	0.939
2	0.272	0.305	0.424	0.311	0.948

lize 200 independent copies of 500 iid pseudo-random vectors which are generated from a bivariate Pareto copula with parameter $\theta = 1$. The estimation results are presented in Figure 3.

For reasons of comparability we chose a fixed threshold k for the tail-copula estimation (which we estimated for $x = y = 1$ according to the described plateau algorithm). The above estimation of the nonparametric tail copula yields very satisfying results for $x, y \leq 1.2$. For larger arguments x and y, the increasing bias results from the fixed threshold choice. A more flexible threshold choice improves these results.



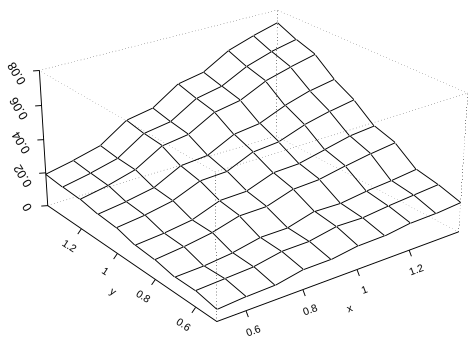


Figure 3: **Left:** True lower tail copula of the Pareto copula with parameter $\theta = 1$. **Right:** Mean-bias of nonparametric estimates $\hat{\Lambda}_L(x, y)$ for 200 samples of a Pareto copula with parameter $\theta = 1$ and sample size 500.

A particular point of interest concerns the estimation of the tail copula in case the data result from a tail-independent random vector. Below we consider data which are generated from a bivariate normal distribution (with correlation coefficients $\rho = 0$ and $\rho = 0.25$) which is known to be tail independent, cf. Schmidt (2002). The left picture in Figure 4 reveals that in case of the standard normal distribution (copula) the tail-copula estimates are performing well. Usually the tail-copula estimates are more volatile in the case of tail independence in comparison to tail dependence. Consider, for example, the single realization of the tail-copula estimator for the normal distribution with $\rho = 0.25$ in the right picture of Figure 4.

Thus our simulations point out that in addition to solely glancing on the TDC the estimation of the tail copula helps a lot to decide whether data are tail dependent or not.

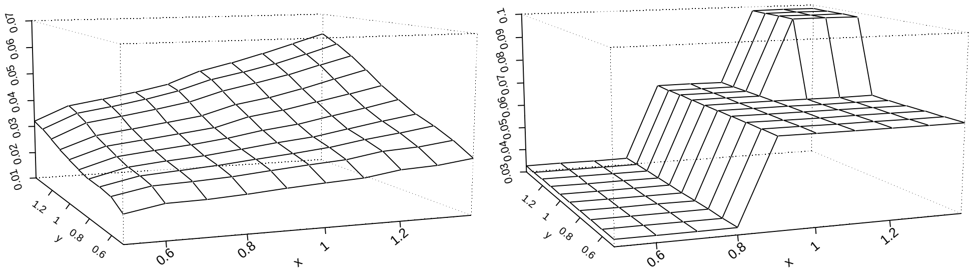


Figure 4: **Left:** Mean of nonparametric estimates $\hat{\Lambda}_L(x, y)$ for 200 samples of a data set of sample size 500 generated from a bivariate standard normal distribution. **Right:** Specific nonparametric estimate $\hat{\Lambda}_L(x, y)$ for a data set of sample size 500 generated from a normal distribution with parameter $\rho = 0.25$.

7.4 Application to financial data

The present section reveals that tail dependence is indeed often found in financial data. Provided are two bivariate series of daily negative log-returns of financial securities and the corresponding upper TDC-estimate $\hat{\lambda}_U$ for various k (again, for notational convenience we drop the index m representing the sample size). Data set D_1 contains negative daily stock log-returns of BMW and Deutsche Bank for the time period 1993-2002 and data set D_2 consists of negative daily exchange-rate log-returns of DM-USD and Yen-USD for the time period 1989-2002. For modelling reasons we assume that the daily log-returns are iid observations (which usually cannot be rejected for extreme returns). Plot 5 shows the presence of tail dependence and the order of magnitude of the tail-dependence coefficient. Moreover, the typical variance-bias problem for various threshold values k can be observed, too. In particular, a small k comes along with a large variance of the TDC estimate, whereas an increasing k results in a strong bias. The threshold k is chosen according to the plateau-finding algorithm described in Section 7.1. Thus for the data set D_1 the algorithm takes k between 80 and 110 which provides a TDC estimate of $\hat{\lambda}_{U,D_1} = 0.31$, whereas for D_2 we obtain $\hat{\lambda}_{U,D_2} = 0.14$.

One application of TDC estimations is given within the Value at Risk (VaR) framework of asset portfolios. VaR calculations relate to high quantiles of portfolio-loss distributions and asset return distributions, respectively. In particular, VaR estimations are highly sensitive towards the tail behavior and the tail dependence of the portfolio's asset-return distribution. Fitting the asset-return random vector towards a multidimensional distribution while utilizing a TDC estimation leads to more accurate VaR estimates (Schmidt 2005). Observe that upper tail-dependence for a random vector $(X_1, X_2)'$ is equivalent to

$$\lambda_U = \lim_{\alpha \to 0} \mathbb{P}(X_2 > \mathrm{VaR}_{1-\alpha}(X_2) | X_1 > \mathrm{VaR}_{1-\alpha}(X_1)) > 0. \tag{30}$$

Finally, in Figure 6 we provide the estimation of the tail copula related to both financial data sets.

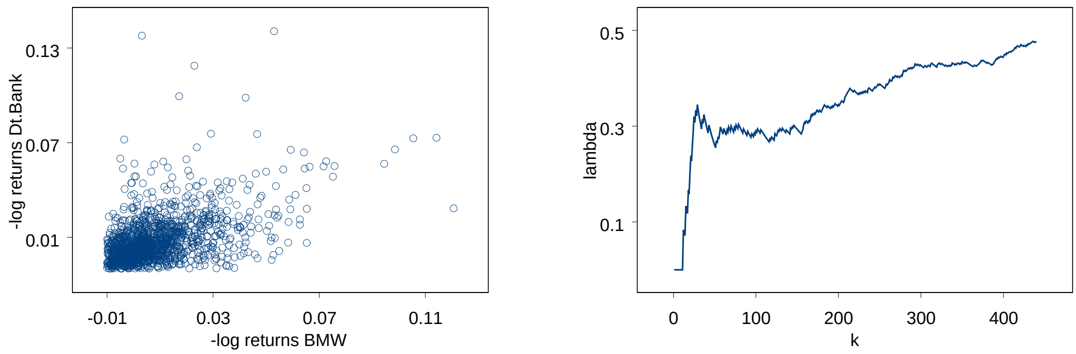


Figure 5: Scatter plot of BMW versus Dt. Bank negative daily stock log-returns (2325 data points) and the corresponding TDC estimates $\hat{\lambda}_U$ for various k.

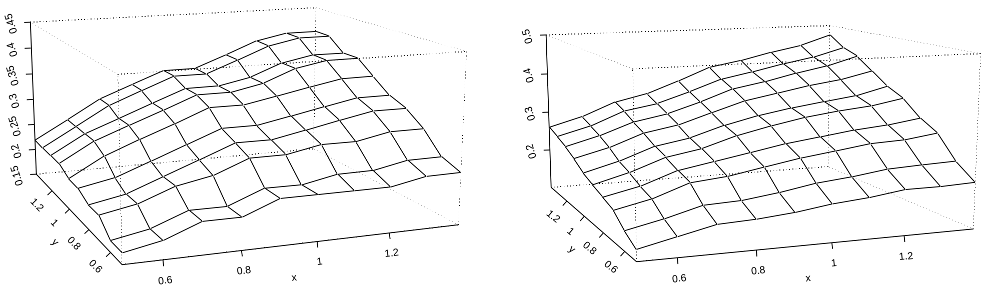


Figure 6: **Left:** Nonparametric tail-copula estimate $\hat{\Lambda}_L(x,y)$ for negative daily stock log-returns of BMW versus Dt. Bank. **Right:** Nonparametric tail-copula estimate $\hat{\Lambda}_L(x,y)$ for negative daily exchange-rate log-returns of DM-USD versus Yen-USD.

8 Conclusions

Summarizing the results, we presented the concept of tail copulae to model extremal dependencies between random variables. It was shown that tail copulae have several analogies to ordinary copulae. Further, we provided nonparametric estimators for the tail copula and the tail-dependence coefficient, and we established the asymptotic normality and strong consistency. A simulation study illustrated that the finite sample behavior of these estimators is satisfying. Among other results, our simulations pointed out that, in addition to solely exploring the tail-dependence coefficient, the estimation of the tail copula helps a lot to decide whether data are tail dependent or not.

Acknowledgment

Both authors gratefully acknowledge the financial support of DaimlerChrysler AG Research and Technologies. This work was partly supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD).
Furthermore, we thank the Editor and the referee for their constructive criticism which helped us to improve the style of the paper.

References

Araujo, A. & Giné, E. (1980). *The central limit theorem for real and Banach valued random variables.* John Wiley & Sons, New York.

Abdous, B., Ghoudi, K. & Khoudraji, A. (1999). Non-parametric estimation of the limit dependence function of multivariate extremes. *Extremes* **2**(3), 245–268.

Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Statist.* **5**, 151–157.

Baron, S., Liflyand, E. & Stadtmüller, U. (2000). Complementary spaces and multipliers of double fourier series for functions of bounded variation. *J. Math. Anal. Appl.* **250**, 706–721.

Billingsley, P. (1968). *Convergence of probability measures.* John Wiley & Sons, New York.

Bingham, N. H., Kiesel, R. & Schmidt, R. (2003). Semi-parametric modelling in finance: Econometric applications, *Quant. Finance* **3**(6), 426–441.

Campbell, R., Koedijk, K. & Kofman, P. (2002). Increased correlation in bear markets. *Financial Analysts Journal* **Jan-Feb**, 87–94.

Casella, G. & Berger, R. L. (2002). *Statistical inference.* Duxbury Advanced Series, 2. edn, Duxbury Press, Pacific Grove.

Coles, S., Heffernan, J. & Tawn, J. (1999). Dependence measures for extreme value analysis. *Extremes* **2**, 339–366.

Cuculescu, I. & Theodorescu, R. (2003). Are copulas unimodal? *J. Multivariate Anal.* **86**, 48-71.

de Haan, L. & Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Aust. Math. Soc.* **61**, 381–395.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg., Bull. C1 Sci. 5ième sér* **65**, 274–292.

Deheuvels, P. (1981). A nonparametric test for independence. *Pub. Inst. Stat. Univ. Paris* **26**(2), 29–50.

Dobrić, J. & Schmid, F. (2005). Nonparametric estimation of the lower tail dependence λ_L in Bivariate Couplas. *J. Appl. Stat.* **32**(4), 387–407.

Draisma, G., Drees, H., Ferreira, A. & de Haan, L. (2004). Bivariate tail estimation: Dependence in asymptotic independence. *Bernoulli* **10**, 251–280.

Drees, H. & Huang, X. (1998). Best attainable rates of convergence for estimates of the stable tail dependence function. *J. Multivariate Anal.* **64**, 25–47.

Dudley, R. M. (1966). Weak convergence of measures on nonseparable metric spaces and empirical measures on euclidean spaces. *Illinois J. Math.* **10**, 109–126.

Dudley, R. M. (1967). Measures on nonseparable metric spaces. *Illinois J. Math.* **11**, 449–453.

Durrett, R. (1996). *Probability theory and examples.* 2. edn, Duxbury Press.

Einmahl, J., de Haan, L. & Piterbarg, V. (2001). Multivariate extremes estimation. *Ann. Statist.* **29**, 1401–1423.

Embrechts, P., Lindskog, F. & McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. *In: Handbook of heavy tailed distibutions in finance*, ed. S. Rachev, Elsevier, Chapter 8, 329–384.

Falk, M. & Reiss, R. (2003). Efficient estimation of the canonical dependence function. *Extremes* **6**, 61–82.

Fermanian, J.-D., Radulović, D. & Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10**(5), 847-860.

Frahm, G., Junker, M. & Schmidt, R. (2005). Estimating the tail-dependence coefficient: Properties and Pitfalls. *Forthcoming in Insurance Math. Econom.* **37**, preprint retrievable from http://www.uni-koeln.de/wiso-fak/wisostatsem/autoren/schmidt.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Hall, P. & Tajvidi, N. (2000). Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli* **6**(5), 835–844.

Heuser, H. (2000). *Lehrbuch der Analysis.* Vol. 2, 11 edn, B.G. Teubner Verlag, Stuttgart.

Huang, X. (1992). *Statistics of bivariate extreme values.* PhD thesis, Tinbergen Institute Research Series Vol. 22, Thesis Publishers and Tinbergen Institute.

Joe, H. (1997). *Multivariate models and dependence concepts.* Monogr. Statist. Appl. Probab. 73, Chapman & Hall, London.

Karolyi, G. & Stulz, R. (1996). Why do markets move together? An investigation of U.S.-Japan stock return comovements. *The Journal of Finance* **51**(3), 951–989.

Ledford, A. & Tawn, J. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.

Ledford, A. & Tawn, J. (1997a). Concomitant tail behavior for extremes. *Adv. in Appl. Probab.* **30**, 197–215.

Ledford, A. & Tawn, J. (1997b). Modelling dependence within joint tail regions. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **59**, 475–499.

Longin, F. & Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance* **56**(2), 649–676.

Nelsen, R. (1999). *An introduction to copulas.* Lecture Notes in Statist. 139, Springer Verlag, New York.

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

Malevergne, Y. & Sornette, D. (2004). How to account for extreme co-movements between individual stocks and the market. *The Journal of Risk* **6**(3), 71–116.

Ong, M. (1999). *Internal credit risk models.* Risk Books, Haymarket, London.

Peng, L. (1998). *Second order condition and extreme value theory.* PhD thesis, Tinbergen Institute Research Series Vol. 178, Thesis Publishers and Tinbergen Institute.

Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.* **43**, 399–409.

Pickands, J. (1981). Multivariate extreme value distributions. *Proc. 43th Session ISI (Buenos Aires)*, 859–878.

Pollard, D. (1984). *Convergence of stochastic processes.* Springer Verlag, New York.

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.

Resnick, S. (1987). *Extreme values, regular variation, and point processes.* Springer Verlag, New York.

Rüschendorf, L. (1976). Asymptotic normality of multivariate rank order statistics. *Ann. Statist.* **4**, 912–923.

Ruymgaart, F. & Wellner, J. (1982). *Growth properties of multivariate empirical processes.* Report 8202, Mathematisch Instituut Katholieke Universiteit, Nijmegen.

Ruymgaart, F. H. (1974). Asymptotic normality of nonparametric test for independence. *Ann. Statist.* **2**, 892–910.

Ruymgaart, F. H., Shorack, G. R. & van Zwet, W. R. (1972). Asymptotic normality of nonparametric test for independence. *Annals of Mathematical Statistics* **43**, 1122–1135.

Schmidt, R. (2002). Tail dependence for elliptically contoured distributions. *Math. Methods Oper. Res.* **55**(2), 301–327.

Schmidt, R. (2005). Tail dependence, *In: Statistical tools in finance and insurance.* eds. P. Čížek, W. Härdle & R. Weron, Springer Verlag, Heidelberg, 65–91.

Sibuya, M., 1960. Bivariate extreme statistics. *Ann. Inst. Statist. Math.* **11**, 195–210.

Sklar, A (1996). Random variables, distribution functions, and copulas - a personal look backward and forward. *In: Distributions with fixed marginals and related topics*, ed. by L. Rueschendorff, B. Schweizer & M.D. Taylor, Institute of Mathematical Statistics, Hayward, 1–14.

Skorokhod, A. V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1**, 261–290.

Song, P. X.-K. (2000). Multivariate dispertion models generated from gaussian copula. *Scand. J. Statist.* **27**(2), 305–320.

Stute, W. (1984). The oscillation behavior of empirical processes: The multivariate case. *Ann. Probab.* **12**(2), 361–379.

Van der Vaart, A. W. & Wellner, J. A. (1996). *Weak convergence and empirical processes.* Springer Ser. Statist., Springer Verlag, New York.

Wheeden, R. & Zygmund, A. (1977). *Measure and integral - An introduction to real analysis.* A series of monographs and textbooks: Pure and Applied Mathematics, Marcel Dekker Inc.

Rafael Schmidt, Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln, 50923 Köln, Germany. E-mail: rafael.schmidt@uni-koeln.de (corresponding author)

Ulrich Stadtmüller, Abteilung für Zahlen- und Wahrscheinlichkeitstheorie, Universität Ulm, 89069 Ulm, Germany. E-mail: stamue@mathematik.uni-ulm.de

A Appendix

A.1 The space $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ and the Delta method

The present section defines the function space $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ and introduces the appropriate concept of weak convergence. We equip the space with some metric and establish necessary and sufficient conditions for weak convergence. The latter space is very suitable for empirical tail copulae. Further, a general *Delta method* is formulated which we utilize in order to prove the main asymptotic results.

Consider the metric spaces $(\mathbb{D}, d)$ and $(\mathbb{E}, e)$. Concepts of weak convergence and almost-sure convergence are traditionally applied to Borel probability measures defined on some space $(\mathbb{D}, \mathcal{D})$ with $\mathcal{D}$ denoting the Borel σ-field of $\mathbb{D}$; see for example Billingsley (1968), p.68. In particular, $\mathcal{D}$ is the smallest σ-field generated by the open sets. However, in the context of empirical tail copulae which are defined in some space $(\mathbb{D}, \mathcal{D})$, these concepts have to be modified as no probability measures can be established on the corresponding Borel σ-field $\mathcal{D}$. Loosely speaking, the Borel σ-field turns out to be too large. This problem arises also for general empirical processes, and several solutions have been proposed in the literature. First, one could restrict to a smaller σ-field like the ball σ-field $\mathcal{D}_B$, and define weak convergence on the new space $(\mathbb{D}, \mathcal{D}_B)$; see for instance Dudley (1966), Dudley (1967) and Pollard (1984). Second, the metric d could be adjusted in such a way that the classical theory is still applicable. A famous example represents the Skorokhod metric on the càdlàg space $D[0,1]$; see Skorokhod (1956). For our purpose the concepts of weak convergence and almost-sure convergence defined by outer expectations are appropriate. A good reference for this theory is the book by Van der Vaart & Wellner (1996).

Definition 3 (Weak convergence with respect to outer expectations)
Let Y be an arbitrary (not necessarily measurable) map from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to the extended real line $\bar{\mathbb{R}}$. The outer integral of Y with respect to the probability measure $\mathbb{P}$ is defined as

$$\mathbb{E}^* Y = \inf\{\mathbb{E}U : U \geq Y, U : \Omega \to \bar{\mathbb{R}} \text{ measurable and } \mathbb{E}U \text{ exists}\}.$$

For each $m \geq 1$, let X_m be an arbitrary (not necessarily measurable) map from a probability space $(\Omega_m, \mathcal{A}_m, \mathbb{P}_m)$ to a metric space $(\mathbb{D}, d)$. Then, X_m is said to converge weakly ($\stackrel{w}{\to}$) to a Borel-measurable map X, if

$$\mathbb{E}^* f(X_m) \to \mathbb{E} f(X) \text{ for every } f \in C_b(\mathbb{D}),$$

where $C_b(\mathbb{D})$ denotes the set of all bounded, continuous and real functions on $\mathbb{D}$.

The main advantage of Definition 3 in general empirical processes theory arises from the fact that classical theorems like the continuous mapping theorem and Prohorov's theorem can be established in the new setting; see Van der Vaart & Wellner (1996). Once the latter theorems are established, the convergence theory becomes less technical like for instance a multidimensional Skorokhod construction, as in Neuhaus (1971).

In order to define the space $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ together with an appropriate metric we need some more notation: The space $l^\infty(T)$ for an arbitrary set T is defined as the set of all uniformly bounded, real functions on T, i.e. all functions $f : T \to I\!R$ such that

$$||f||_T := \sup_{t \in T} |f(t)| < \infty.$$

Consequently, the uniform distance on $l^\infty(T)$ is defined by

$$d(f_1, f_2) = ||f_1 - f_2||_T.$$

The stochastic processes $\{X_m(t) : t \in T\}$ considered below will have their sample paths in $l^\infty(T)$ if T is a compact subset of $\bar{I\!R}_+^2$. If not stated otherwise, T will always denote a compact subset of $\bar{I\!R}_+^2$.

The space $l^\infty(T)$ will be equipped with the corresponding Borel σ-field. The only measurability we require for X_m is the measurability of the maps $X_m(t) : \Omega_m \to I\!R$, $t \in T$ (this means that $X_m(t)$ is a random variable for each fixed $t \in T$, cf. Pollard (1984)); a rather weak condition. However, the limiting process X which turns out to be a continuous Gaussian process is a Borel measurable map $X|_T : \Omega \to C(T)$ as the space $C(T)$ of all continuous real function is a separable and complete subspace of $l^\infty(T)$ with respect to the uniform metric. Moreover, it can be shown that the Borel σ-field of $C(T)$ correspond to the related projection σ-field.

We are ready to define the metric space $\mathcal{B}_\infty(\bar{I\!R}_+^2)$.

Definition 4 *The space $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ is defined as the family of all functions $f : \bar{I\!R}_+^2 \to I\!R$ which are locally uniformly-bounded on every compact subset of $\bar{I\!R}_+^2$ (but not necessarily on $\bar{I\!R}_+^2$). Then, $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ is a complete metric space under the metric*

$$d(f_1, f_2) = \sum_{i=1}^{\infty} 2^{-i}(||f_1 - f_2||_{T_i} \wedge 1) \tag{31}$$

with $T_{3i} = T_{3i-1} \cup [0, i]^2$, $T_{3i-1} = T_{3i-2} \cup ([0, i] \times \{\infty\})$, $T_{3i-2} = T_{3(i-1)} \cup (\{\infty\} \times [0, i]) \subset \bar{I\!R}_+^2$, $i \in I\!N$, and $T_0 = \emptyset$. Thus a sequence of elements in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ converges in this metric if it converges uniformly on each T_i.

Remark. The space $\mathcal{B}_\infty(\bar{I\!R}_+^n)$ is defined analogously to $\mathcal{B}_\infty(\bar{I\!R}_+^2)$.

The following theorem is fundamental for our purposes. A *proof* can be found in Van der Vaart & Wellner (1996), Theorem 1.6.1.

Theorem 8 *For each $m \geq 1$, let $X_m : \Omega_m \to \mathcal{B}_\infty(\bar{I\!R}_+^2)$ be an arbitrary map. Then the sequence X_m converges weakly to a tight limit if and only if every sequence of restrictions $X_m|_{T_i} : \Omega_m \to l^\infty(T_i)$ converges weakly to a tight limit.*

The Delta method, cf. Casella & Berger (2002), Section 5.5.4, is a well-known technique in statistics to prove results about the asymptotic normality of functionals of estimators. In the context of tail copulae we need a quite general version of the Delta method. For this, the notion of Hadamard differentiability is useful. Let $(I\!D, d)$ and $(I\!E, e)$ be metrizable or topological vector spaces, in particular vector addition and scalar multiplication are continuous operations.

Definition 5 (Hadamard differentiability) *A map $\phi : I\!D_\phi \subset I\!D \to I\!E$ is called Hadamard-differentiable at $\theta \in I\!D_\phi$ if there exists a continuous linear map $\phi'_\theta : I\!D \to I\!E$ such that*

$$\frac{\phi(\theta + t_m h_m) - \phi(\theta)}{t_m} \to \phi'_\theta(h), \quad \text{as } m \to \infty, \tag{32}$$

for all converging sequences $t_m \to 0$ and $h_m \to h$ such that $\theta + t_m h_m \in I\!D_\phi$ for all m. Further, $\phi : I\!D_\phi \subset I\!D \to I\!E$ is called Hadamard-differentiable tangentially to a set $I\!D_0 \subset I\!D$ by requiring that $h_m \to h$ with $h \in I\!D_0$. In that case the derivative ϕ'_θ needs only be defined on $I\!D_0$.

Note that $\mathbb{D}_\phi$ is allowed to be any arbitrary subset of $\mathbb{D}$; this fact turned out to be important in our elaborations.

Theorem 9 (Delta method) *Let $\phi : \mathbb{D}_\phi \subset \mathbb{D} \to \mathbb{E}$ be Hadamard-differentiable at θ tangentially to $\mathbb{D}_0$. Suppose $X_m : \Omega_m \to \mathbb{D}_\phi$ are (not necessarily measurable) maps with $r_m(X_m - \theta) \stackrel{w}{\to} X$ for some sequence of constants $r_m \to \infty$, where $X : \Omega \to \mathbb{D}_\phi$ is separable. Then*

$$r_m\big(\phi(X_m) - \phi(\theta)\big) \stackrel{w}{\to} \phi'_\theta(X).$$

For details we refer the reader to Van der Vaart & Wellner (1996), p.374.

A.2 Hadamard differentiability

The proof of Theorem 5 (asymptotic normality of the empirical tail copula) in Section 5 requires the following lemma. The lemma is stated in the original version as provided in Van der Vaart & Wellner (1996), p.388. However, in our context, the space of uniformly bounded real function $l^\infty(T)$ has to be substituted by the appropriate space of locally uniformly bounded real functions on compact sets; likewise the corresponding metrics.

Let $\mathcal{Y}$ and $\mathcal{Z}$ be subsets of normed spaces. Consider the maps $A : \mathcal{X} \mapsto \mathcal{Y}$ and $B : \mathcal{Y} \mapsto \mathcal{Z}$ which define the composition map $\Phi(A, B) : \mathcal{X} \mapsto \mathcal{Z}$ via

$$\Phi(A, B)(x) = B \circ A(x) = B(A(x)).$$

If B is a uniformly norm-bounded map from $\mathcal{Y} \mapsto \mathcal{Z}$, then $\Phi(A, B)$ is a uniformly norm-bounded map from $\mathcal{X} \mapsto \mathcal{Z}$. Consider now Φ as a map with domain $l^\infty(\mathcal{X}) \times l^\infty(\mathcal{Y})$ equipped with the norm $||(A, B)||_\infty = \sup_x ||A(x)||_{\mathcal{Y}} \vee \sup_y ||B(y)||_{\mathcal{Z}}$.

Lemma 1 *Suppose $B : \mathcal{Y} \mapsto \mathcal{Z}$ is Fréchet-differentiable uniformly in y in the range of A with derivatives B'_y such that $y \mapsto B'_y$ is uniformly norm-bounded. Then the composition map $\Phi : l^\infty(\mathcal{X}) \times l^\infty(\mathcal{Y}) \mapsto l^\infty(\mathcal{X})$ is Hadamard-differentiable at (A, B) tangentially to the set $l^\infty(\mathcal{X}) \times UC(\mathcal{Y})$ ($UC(T)$ denotes the space of uniformly continuous function from T to $\mathbb{R}$). The derivative is given by*

$$\Phi'_{A,B}(\alpha, \beta)(x) = \beta \circ A(x) + B'_{A(x)}\big(\alpha(x)\big), \quad x \in \mathcal{X}.$$