Nonparametric estimation of tail dependence

RAFAEL SCHMIDT AND ULRICH STADTMÜLLER

London School of Economics and University of Ulm

ABSTRACT. Dependencies of 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 nonparametric estimators for the tail copulae are introduced and weak convergence, asymptotic normality, and strong consistency 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. [1]

Keywords: Tail dependence, Tail dependence coefficient, Tail copula, Nonparametric estimation, Copula, Empirical copula, Asymptotic normality, Strong consistency.

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 and Stulz (1996), Longin and Solnik (2001), and Campbell, Koedijk and 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.

The current standard studying extremal dependencies is to use the concept of *tail dependence* (cf. Joe (1997), Embrechts et al. (2003), Malevergne and Sornette (2003)). 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. Similarly to the well-known copula-concept (cf. Joe (1997), Nelsen (1999)) we may construct multivariate extreme-value distributions with a given tail copula. Recently, the copula-concept has become quite important in theory and applications, see e.g. Sklar (1996), Song (2000), Cuculescu and

[1] Submitted to Blackwell Publishing, December 2003

Theodorescu (2001), 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.

Concerning the estimation of general copula functions, several parametric, semi-parametric, and nonparametric procedures have already been proposed in the literature (cf. Stute (1984) Genest and Rivest (1993), Genest, Ghoudi and Rivest (1995)). Regarding the nonparametric estimation, Deheuvels (1979, 1981) and Fermanian, Radulović and Wegkamp (2002) establish weak convergence of the so-called *empirical copula process* under independent and dependent marginal distributions.

In the present paper we start with the definition of the so-called *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 copulae are introduced and in Section 4 weak convergence, asymptotic normality are shown by means of a functional delta-method as provided in the monograph of Van der Vaart and Wellner (1996). The next section is devoted to strong consistency and further results on functionals of tail ranks. Some simulations 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 $[0,1]$. In particular every n-dimensional distribution function F can be written in the form

$$F(x_1,\dots,x_n) = C(F_1(x_1),\dots,F_n(x_n)), \tag{2.1}$$

where $F_1,\dots,F_n$ are the marginal distribution functions. We assume that these are continuous. Then the copula C is unique and has the representation

$$C(u_1,\dots,u_n) = F(F_1^{-1}(u_1),\dots,F_n^{-1}(u_n)), \quad 0 \le u_1,\dots,u_n \le 1,$$

where $F_1^{-1},\dots,F_n^{-1}$ denote the generalized inverse distribution functions of $F_1,\dots,F_n$, i.e., for all $u_i \in (0,1]: F_i^{-1}(u_i) := \inf\{x \in I\!R \mid F_i(x) \ge u_i\},\ i=1,\dots,n$ with $\inf\{\emptyset\} = \infty$.

Conversely, if C is a copula and $F_1,\dots,F_n$ are distribution functions, then the function F defined by (2.1) is an n-dimensional distribution function with margins $F_1,\dots,F_n$.

The copula function represents the dependence structure of a multivariate random vector. 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 multi-dimensional distributions in the tail and are defined as follows. Throughout this paper we denote by $\bar{I\!R}^n_+ := [0,\infty]^n \backslash \{(\infty,\dots,\infty)\}$.

Definition 1 Tail copulae. *Let F be an n-dimensional distribution function. If for the subsets $I, J \subset \{1,\dots,n\}$, $I \cap J = \emptyset$, the following limit exists everywhere on $\bar{I\!R}^n_+$*

$$\Lambda_U^{I,J}(x) := \lim_{t\to\infty} I\!P(X_i > F_i^{-1}(1-\frac{x_i}{t}),\ \forall i \in I \mid X_j > F_j^{-1}(1-\frac{x_j}{t}),\ \forall j \in J) \tag{2.2}$$

then the function $\Lambda_U^{I,J}: \bar{I\!R}^n_+ \to I\!R$ is called an upper tail-copula associated with F w.r.t. to I, J. The corresponding lower tail-copula is defined by

$$\Lambda_L^{I,J}(x) := \lim_{t\to\infty} I\!P(X_i \le F_i^{-1}(x_i/t),\ \forall i \in I \mid X_j \le F_j^{-1}(x_j/t),\ \forall j \in J) \tag{2.3}$$

provided the limit exists.

For simplicity and notational convenience all further definitions and results are provided for the bivariate case first. The multidimensional extensions are given in Section 8. Within the bivariate framework we consider a random vector $(X,Y)'$ with distribution function F

and continuous marginal distribution functions G and H. The estimation becomes easier if the following slight modification of the tail copula is utilized:

$$\Lambda_U(x,y) := y \cdot \Lambda_U^{\{1\},\{2\}}(x,y) \quad \text{and} \tag{2.4}$$

$$\Lambda_L(x,y) := y \cdot \Lambda_L^{\{1\},\{2\}}(x,y), \quad x \in \bar{\mathbb{R}}_+, y \in \mathbb{R}_+, \tag{2.5}$$

where the indices $\{1\}$ and $\{2\}$ can be dropped. Further, set $\Lambda_U(x,\infty) := x$ and $\Lambda_L(x,\infty) := x$ for all $x \in \mathbb{R}_+$.

The next definition embeds the well-known tail-dependence concept (see Joe (1997), p. 33) within the framework of tail copulae. For an account on tail dependence for elliptically contoured distributions we refer to Schmidt (2002).

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

$$\lambda_U := \Lambda_U(1,1) = \lim_{v\to 1^-} \mathbb{P}(X > G^{-1}(v) \mid Y > H^{-1}(v)) > 0. \tag{2.6}$$

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

$$\lambda_L := \Lambda_L(1,1) \tag{2.7}$$

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.

The tail copula and the tail-dependence coefficient do not depend on the margins but only on the copula. E.g., we have

$$\lambda_U = \lim_{u\to 1-} \frac{1-2u+C(u,u)}{1-u} \text{ and } \Lambda_U(x,y) = x + y - \lim_{t\to\infty} t(1 - C(1-\frac{x}{t}, 1-\frac{y}{t}))$$

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

1. To derive explicit weak convergence results for the following nonparametric estimators of the lower and upper tail-dependence coefficient.
2. As an intuitive and straightforward generalization of the well-known concept of tail dependence.
3. As the immediate counterpart of the so-called stable tail-dependence function (cf. Huang (1992), Chapter 2, and Peng (1998), pp. 96).
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.

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 is not a copula function.

In bivariate EVT the major interest concerns the probability

$$\mathbb{P}(X > G^{-1}(1-x) \text{ or } Y > H^{-1}(1-y)), \tag{2.8}$$

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

$$\mathbb{P}(X > G^{-1}(1-x) \text{ and } Y > H^{-1}(1-y)). \tag{2.9}$$

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

At this point we would like to mention that the nonparametric estimators we propose later base on the empirical counterparts of the probabilities (2.8) and (2.9) and utilize the above homogeneity property. Notice, in case $(X,Y)'$ is tail independent, the latter property does not hold for (2.9). Here, an adjusted homogeneity property can be obtained by assuming that $t \mapsto \mathbb{P}(X > G^{-1}(1-x/t)$ and $Y > H^{-1}(1-y/t))$ is regularly varying at infinity with index $-1/\eta,\ \eta < 1$. The parameter η was introduced by Ledford and 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 and Tawn (1999), Peng (1999), and Draisma, Drees, Ferreira and de Haan (2001). 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 (e.g. 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 $\Lambda_L(tx,ty) = t\Lambda_U(x,y)$ 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.

vi) **(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

$$\begin{aligned} |\Lambda(x,y) - \Lambda(\bar{x},\bar{y})| &= \lim_{t\to\infty} t|C(x/t,y/t) - C(\bar{x}/t,\bar{y}/t)| \\ &\le |x-\bar{x}| + |y-\bar{y}| \le K \left\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \right\|_2 \end{aligned} \tag{3.10}$$

for some constant $K > 0$ because C is a bivariate 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 the latter inequality 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

$$\begin{aligned}\min\{x_n/x_0,\, y_n/y_0\}\tau_n\,C(x_0/\tau_n,\, y_0/\tau_n) &\le t_n\,C(x_n/t_n,\, y_n/t_n)\\ &\le \max\{x_n/x_0,\, y_n/y_0\}\xi_n\,C(x_0/\xi_n,\, y_0/\xi_n).\end{aligned}$$

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

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 lower and upper bound arise from the Fréchet-Hoeffding bounds for copulae (cf. Nelsen (1999), Theorem 2.2.3), i.e., for every copula function C we have

$$\max(u+v-1, 0) \le C(u,v) \le \min(u,v).$$

Part ii) follows from the proof of part v) in Theorem 1. 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 $\lim_{t\to\infty} t\,C(x/t,y/t) = (\sqrt{x^2+y^2}\lim_{\tau\to\infty}\tau C(x/(\sqrt{x^2+y^2}\tau), y/(\sqrt{x^2+y^2}\tau))\,.$
$\square$

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*

$$0 \le \frac{\partial}{\partial x}\Lambda_L(x,y) \le 1. \tag{3.11}$$

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

$$0 \le \frac{\partial}{\partial y}\Lambda_L(x,y) \le 1. \tag{3.12}$$

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 and Zygmund (1977)). Inequalities (3.11) and (3.12) are implied by the Lipschitz condition (3.10). 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. $\square$

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

In the forthcoming, three nonparametric estimators for the modified upper and lower tail-copula $\Lambda_U(x,y)$ and $\Lambda_L(x,y)$, $(x,y)' \in \bar{R}^2_+$, are proposed. 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

$$C_m(u,v) = F_m(G_m^{-1}(u), H_m^{-1}(v)),\ (u,v)' \in [0,1]^2 \tag{4.13}$$

with F_m, G_m, H_m being the empirical distribution functions corresponding to F,G,H. Then, $C_m((a,b] \times (c,d])$ is called the empirical copula-"measure" of the interval $(a,b] \times (c,d] \subset [-\infty,\infty]^2$ given by

$$C_m((a,b] \times (c,d]) := C_m\big(((a,b] \times (c,d]) \cap [0,1]^2\big).$$

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 and Schmidt (2002).

Let $R^{(j)}_{m1}$ and $R^{(j)}_{m2}$ denote the rank of $X^{(j)}$ and $Y^{(j)}$, $j = 1,\dots,m$, respectively. The first set of estimators are based on formulae (2.2) and (2.3):

$$\begin{aligned}\hat{\Lambda}_{U,m}(x,y) &:= \frac{m}{k}\, C_m\Big(\Big(1 - \frac{kx}{m}, 1\Big] \times \Big(1 - \frac{ky}{m}, 1\Big]\Big) \\ &\approx \frac{1}{k}\sum_{j=1}^{m} \mathbf{1}_{\{R^{(j)}_{m1} > m-kx \text{ and } R^{(j)}_{m2} > m-ky\}}\end{aligned} \tag{4.14}$$

and

$$\hat{\Lambda}_{L,m}(x,y) := \frac{m}{k} C_m\Big(\frac{kx}{m}, \frac{ky}{m}\Big) \approx \frac{1}{k}\sum_{j=1}^{m} \mathbf{1}_{\{R^{(j)}_{m1} \le kx \text{ and } R^{(j)}_{m2} \le ky\}} \tag{4.15}$$

with some parameter $k \in \{1,\dots,m\}$ 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 (4.14) and (4.15) 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 et al. (1995), i.e.,

$$\bar{C}_m(u,v) = \frac{1}{m}\sum_{j=1}^{m} \mathbf{1}_{\{G_m(X^{(j)}) \le u \text{ and } H_m(Y^{(j)}) \le v\}},\quad (u,v)' \in [0,1]^2. \tag{4.16}$$

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 corresponding rank order statistics.

Note that the empirical tail-copulae and the rank order statistics coincide on the grid $\{(i/k, j/k),\ 1 \le i,j \le m\}$. 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

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

A related pair of estimators was introduced and investigated by Huang (1992), Chapter 2, Peng (1998), pp. 96, and Durrleman, Nikeghbali and Roncalli (2000) in the context of stable tail-dependence functions. The relationship between the bivariate 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 function l with respect to questions arising from EVT. They show consistency and asymptotic normality for the estimator below. However, instead of utilizing the Skorokhod representation theorem as in Huang (1992), Chapter 2, we apply a general Delta method to prove asymptotic normality. Observe that $\Lambda_U(x,y)=x+y-\lim_{t\to\infty}t(1-C(1-x/t,1-y/t))$. Thus, an estimator for $\Lambda_U(x,y)$ is given by

$$\begin{aligned}\hat{\Lambda}^{EVT}_{U,m}(x,y)&:=x+y-\frac{m}{k}\Big(1-C_m\Big(1-\frac{kx}{m},1-\frac{ky}{m}\Big)\Big)\\ &\approx x+y-\frac{1}{k}\sum_{j=1}^{m}\mathbf{1}_{\{R^{(j)}_{m1}>m-kx\text{ or }R^{(j)}_{m2}>m-ky\}}\end{aligned}\tag{4.17}$$

with $k=k(m)\to\infty$ and $k/m\to 0$ as $m\to\infty$. The estimator $\hat{\Lambda}^{EVT}_{L,m}$ is defined similarly. One important practical problem 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 present paper concerns the study of the asymptotic behavior of the empirical tail-copulae $\hat{\Lambda}_{U,m}$ and $\hat{\Lambda}_{L,m}$ stated in (4.14) and (4.15). These estimators, although different, are related to the estimator $\hat{\Lambda}^{EVT}_{U,m}$ introduced in Huang (1992). However, as the method of proof used in Huang (1992) cannot be applied in our case, we chose a different approach. Furthermore, the results of Huang (1992) can be shown with the same techniques. Extensions of the latter results to dimensions larger than two are possible. Strong consistency and asymptotic normality of the above estimators are addressed in Sections 5 and 6.

The third set 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:

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

where the finite measure Φ, which lives on $[0,\pi/2]$, denotes the spectral measure of Λ_U. Einmahl, de Haan and Piterbarg (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 $m\to\infty$. The estimator $\hat{\Lambda}^{S}_{L,m}$ is defined similarly.

Based on the above estimators for the lower and upper tail-copula, we propose

$$\hat{\lambda}_{U,m}:=\hat{\Lambda}_{U,m}(1,1),\ \hat{\lambda}^{EVT}_{U,m}:=\hat{\Lambda}^{EVT}_{U,m}(1,1),\text{ and }\hat{\lambda}^{S}_{U,m}:=\hat{\Lambda}^{S}_{U,m}(1,1)\tag{4.18}$$

as nonparametric estimators for the upper tail-dependence coefficient and

$$\hat{\lambda}_{L,m} := \hat{\Lambda}_{L,m}(1,1),\ \hat{\lambda}^{EVT}_{L,m} := \hat{\Lambda}^{EVT}_{L,m}(1,1),\ \text{and}\ \hat{\lambda}^{S}_{L,m} := \hat{\Lambda}^{S}_{L,m}(1,1) \tag{4.19}$$

as nonparametric estimators for the lower tail-dependence coefficient. Note that $\hat{\lambda}^S_{U,m}$ degenerates to $\hat{\Phi}_m(\pi/2)$ and therefore $\hat{\lambda}^S_{U,m} = \hat{\lambda}_{U,m}$ (equivalently $\hat{\lambda}^S_{L,m} = \hat{\lambda}_{L,m}$). Further nonparametric estimators for the lower tail-dependence coefficient are provided in Dobrić and Schmid (2002).

In Section 9, a simulation study compares all introduced nonparametric estimators for the tail-dependence coefficient regarding their finite sample properties.

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 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 an elegant Delta method (see Theorem 13). The techniques to convey this can be found in Van der Vaart and 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 $\Lambda_L(x,y)$:

$$\hat{\Lambda}^*_{U,m}(x,y) := \frac{1}{k}\sum_{j=1}^{m} \mathbf{1}_{\left\{G(X^{(j)}) > 1-\frac{kx}{m} \text{ and } H(Y^{(j)}) > 1-\frac{ky}{m}\right\}} \tag{5.20}$$

and

$$\hat{\Lambda}^*_{L,m}(x,y) := \frac{1}{k}\sum_{j=1}^{m} \mathbf{1}_{\left\{G(X^{(j)}) \le \frac{kx}{m} \text{ and } H(Y^{(j)}) \le \frac{ky}{m}\right\}}. \tag{5.21}$$

Condition 4 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 and Stadtmüller (1996)).

Theorem 5 Asymptotic normality under known margins G and H. *Let F be a bivariate 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 4 with*

$$\sqrt{k}\,A(m/k) \to 0 \quad as \quad m \to \infty \tag{5.22}$$

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

$$\sqrt{k}\left(\hat{\Lambda}^*_{U,m}(x,y) - \Lambda_U(x,y)\right) \xrightarrow{w} \mathbb{G}_{\hat{\Lambda}^*_U}(x,y) \tag{5.23}$$

and

$$\sqrt{k}\left(\hat{\Lambda}^*_{L,m}(x,y) - \Lambda_L(x,y)\right) \xrightarrow{w} \mathbb{G}_{\hat{\Lambda}^*_L}(x,y), \tag{5.24}$$

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 $\mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ and the covariance structure of $\mathbb{G}_{\hat{\Lambda}^*_U}(x,y)$ and $\mathbb{G}_{\hat{\Lambda}^*_L}(x,y)$ is given in Corollary 1 below.

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

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 (5.22) 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 $k/m \to 0$ 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 $\{(x_1,y_1),\ldots,(x_t,y_t)\}$ of $\bar{I\!R}^2_+$ there exists a centered normal random vector $(\alpha(x_1,y_1),\ldots,\alpha(x_t,y_t))$ with appropriate covariance structure such that

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

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

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

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

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

Notice that $t\,C(x_r/t,y_r/t)\,C(x_s/t,y_s/t) \to 0$ as $t\to\infty$. The matrix $A = (a_{r,s})_{r,s=1,\ldots,t}$ is nonzero if $\Lambda_L \not\equiv 0$ according to Theorem 1. Further, for $Z^{(j)}_m = (Z^{(j)}_{1,m},\ldots,Z^{(j)}_{t,m})$ and the Euclidian norm $||.||_2$ we have

$$||Z^{(j)}_m||_2^2 = \sum_{i=1}^{t}\big(Z^{(j)}_{i,m}\big)^2 \le \frac{t}{k}$$

and thus

$$\sum_{j=1}^{m}\int_{\{||Z^{(j)}_m||_2>\varepsilon\}} ||Z^{(j)}_m||_2^2\, dI\!P \to 0 \quad \text{as } m\to\infty$$

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 $M \in I\!N$ 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 $m_0 \in I\!N$ such that

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

Note that α_m belongs to the space of cádlág functions $D(\bar{I\!R}^2_+)$. 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 $D(\bar{I\!R}_+^2)$ 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 $(M\cdot l_1/L, M\cdot l_2/L)$, $l_i\in\{0,\ldots,L\}, L\in I\!N$, $i=1,2$. Then for arbitrary but fixed $\xi>0$ and $\delta\in(0,1)$ such that $1/L\geq\delta$ we obtain

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

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

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

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 and Wellner (1982), Inequality 1.1. In particular the function $\psi:[-1,\infty)\to I\!R$ 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 $m,L\in I\!N$, $n=1,2$, and $l_i\in\{1,\ldots,L\},\ i=1,2$,

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

In the first case an upper bound is provided by

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

whereas in the second case we utilize the upper bound

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

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 and 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 and Wellner (1996). Finally, weak convergence of α_m in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ is provided by Theorem 12. Continuity of the sample paths in $\bar{I\!R}_+^2$ follows according to the Addendum 1.5.8 in the latter reference. $\square$

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 6 ASYMPTOTIC NORMALITY UNDER UNKNOWN MARGINS G AND H. *Let F be a bivariate 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 4 holds, then for $\sqrt{k}A(m/k) \to 0$ as $m \to \infty$*

$$\sqrt{k}\{\hat{\Lambda}_{U,m}(x,y) - \Lambda_U(x,y)\} \xrightarrow{w} \mathbb{G}_{\Lambda_U}(x,y),$$

and

$$\sqrt{k}\{\hat{\Lambda}_{L,m}(x,y) - \Lambda_L(x,y)\} \xrightarrow{w} \mathbb{G}_{\Lambda_L}(x,y),$$

where $\mathbb{G}_{\Lambda_U}(x,y)$ and $\mathbb{G}_{\Lambda_L}(x,y)$ are centered tight continuous Gaussian random fields. Weak convergence takes place in $\mathcal{B}_\infty(\bar{I\!R}_+^2)$ and the covariance structure of $\mathbb{G}_{\hat{\Lambda}_U}(x,y)$ and $\mathbb{G}_{\hat{\Lambda}_L}(x,y)$ is given in Corollary 1 below.

PROOF. The proof is given for the lower empirical tail-copula as the upper empirical tail-copula can be treated similarly. 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 (A.39) 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 stated in the Appendix A.1 to the following map

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

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

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

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

$$\Phi : \gamma(x,y) \overset{\Phi_1}{\mapsto} (\gamma(x,y), \gamma(x,\infty), \gamma(\infty,y))$$

$$\overset{\Phi_2}{\mapsto} (\gamma(x,y), \gamma^-(x,\infty), \gamma^-(\infty,y)) \overset{\Phi_3}{\mapsto} \gamma \circ (\gamma^-(x,\infty), \gamma^-(\infty,y)),$$

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

$$\Phi : \gamma(x,y) \overset{\Phi_1}{\mapsto} (\gamma(x,y), \gamma(x,\infty), \gamma(x,\infty))$$

$$\overset{\Phi_2}{\mapsto} (\gamma(x,y), \gamma^-(x,\infty), \gamma^-(x,\infty)) \overset{\Phi_3}{\mapsto} \gamma \circ (\gamma^-(x,\infty), \infty),$$

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

$$\Phi : \gamma(x,y) \overset{\Phi_1}{\mapsto} (\gamma(x,y), \gamma(\infty,y), \gamma(\infty,y))$$

$$\overset{\Phi_2}{\mapsto} (\gamma(x,y), \gamma^-(\infty,y), \gamma^-(\infty,y)) \overset{\Phi_3}{\mapsto} \gamma \circ (\infty, \gamma^-(\infty,y)).$$

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 $\mathcal{B}^I_\infty(\bar{I\!R}^2_+)$ at $\gamma_0 = \Lambda_L$ tangentially to $C(\bar{I\!R}^2_+) \subset \mathcal{B}_\infty(\bar{I\!R}^2_+)$.

i) The first map Φ_1 is Hadamard-differentiable on $\mathcal{B}^I_\infty(\bar{I\!R}^2_+)$ 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 $\mathcal{B}^I_\infty(\bar{I\!R}^2_+) \times B^I(I\!R_+) \times B^I(I\!R_+)$ at $(\Lambda_L, \mathrm{id}_{I\!R_+}, \mathrm{id}_{I\!R_+})$ tangentially to $C(\bar{I\!R}^2_+) \times C(I\!R_+) \times C(I\!R_+)$. 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 $\mathcal{B}^I_\infty(\bar{I\!R}^2_+)$. 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 $C(I\!R_+)$. Here

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

denote 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

$$-th(\xi_{pt}) + o(t) \leq \xi_{pt} - p \leq -th(\xi_{pt} - \varepsilon_{pt}) + o(t), \tag{5.25}$$

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 (5.25), 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 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 (A.39) (cf. Heuser (2000) , Satz 164.4, and Van der Vaart and 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 and 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 $\hat{\Lambda}^*_{L,m}(\infty, y)$. $\square$

Corollary 1 Covariance structure. *The covariance structure of $G_{\Lambda_U^*}$ and $G_{\Lambda_L^*}$ in Theorem 5 is given by*

$$\mathbb{E}\big(\mathbb{G}_{\hat{\Lambda}_U^*}(x,y)\cdot \mathbb{G}_{\hat{\Lambda}_U^*}(\bar{x},\bar{y})\big) = \Lambda_U\big(\min\{x,\bar{x}\},\min\{y,\bar{y}\}\big) \quad \text{and} \tag{5.26}$$

$$\mathbb{E}\big(\mathbb{G}_{\hat{\Lambda}_L^*}(x,y)\cdot \mathbb{G}_{\hat{\Lambda}_L^*}(\bar{x},\bar{y})\big) = \Lambda_L\big(\min\{x,\bar{x}\},\min\{y,\bar{y}\}\big) \tag{5.27}$$

for $(x,y)', (\bar{x},\bar{y})' \in \bar{\mathbb{R}}^2_+$.

Further the limiting process in Theorem 6 can be expressed by

$$\begin{aligned}\mathbb{G}_{\hat{\Lambda}_U}(x,y) &= \mathbb{G}_{\hat{\Lambda}_U^*}(x,y) \\ &\quad - \frac{\partial}{\partial x}\Lambda_U(x,y)\mathbb{G}_{\hat{\Lambda}_U^*}(x,\infty) - \frac{\partial}{\partial y}\Lambda_U(x,y)\mathbb{G}_{\hat{\Lambda}_U^*}(\infty,y)\end{aligned} \tag{5.28}$$

and

$$\begin{aligned}\mathbb{G}_{\hat{\Lambda}_L}(x,y) &= \mathbb{G}_{\hat{\Lambda}_L^*}(x,y) \\ &\quad - \frac{\partial}{\partial x}\Lambda_L(x,y)\mathbb{G}_{\hat{\Lambda}_L^*}(x,\infty) - \frac{\partial}{\partial y}\Lambda_L(x,y)\mathbb{G}_{\hat{\Lambda}_L^*}(\infty,y).\end{aligned} \tag{5.29}$$

Proof. The covariance structure (5.27) and (5.26), respectively, has been derived in the proof of Theorem 5. The second assertion follows from the Delta method (see Theorem 13 in the Appendix). Notice that the derivative Φ'_{Λ_L} of the map Φ (at the point Λ_L) utilized in Theorem 6 is of the form

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

The latter claim follows according to the specific form of the derivative map of the inverse operator provided in the proof of Theorem 6 (part ii)), Lemma 1, and the chain rule (Lemma 3.9.3 in Van der Vaart and Wellner (1996)). $\square$

Corollary 2 Asymptotic normality of $\hat{\lambda}_{U,m}$ and $\hat{\lambda}_{L,m}$. *With the prerequisites of Theorem 6*

$$\sqrt{k}\big\{\hat{\lambda}_{U,m} - \lambda_U\big\} \xrightarrow{d} \mathcal{N}_{0,\sigma_U^2} \quad \text{and} \quad \sqrt{k}\big\{\hat{\lambda}_{L,m} - \lambda_L\big\} \xrightarrow{d} \mathcal{N}_{0,\sigma_L^2},$$

where $\mathcal{N}_{0,\sigma_U^2}$ and $\mathcal{N}_{0,\sigma_L^2}$ are centered normal-distributed random variables with variances

$$\begin{aligned}\sigma_U^2 &= \lambda_U + \Big(\frac{\partial}{\partial x}\Lambda_U(1,1)\Big)^2 + \Big(\frac{\partial}{\partial y}\Lambda_U(1,1)\Big)^2 \\ &\quad + 2\lambda_U\Big(\Big(\frac{\partial}{\partial x}\Lambda_U(1,1)-1\Big)\Big(\frac{\partial}{\partial y}\Lambda_U(1,1)-1\Big)-1\Big)\end{aligned} \tag{5.30}$$

and

$$\begin{aligned}\sigma_L^2 &= \lambda_L + \Big(\frac{\partial}{\partial x}\Lambda_L(1,1)\Big)^2 + \Big(\frac{\partial}{\partial y}\Lambda_L(1,1)\Big)^2 \\ &\quad + 2\lambda_L\Big(\Big(\frac{\partial}{\partial x}\Lambda_L(1,1)-1\Big)\Big(\frac{\partial}{\partial y}\Lambda_L(1,1)-1\Big)-1\Big).\end{aligned} \tag{5.31}$$

Proof. Note that e.g. for the lower tail-copula Λ_L we know from Corollary 1 that

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

$\square$

Example. To illustrate the latter results we calculate the tail copula Λ_L and the asymptotic variance σ_L^2 in (5.31) for the well-known Pareto copula.

The bivariate 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

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

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 (5.31) is given by (see also Figure 1)

$$\sigma_L^2(\theta) = 2^{-1/\theta} - \frac{3}{2}4^{-1/\theta} + \frac{1}{2}8^{-1/\theta}. \tag{5.32}$$

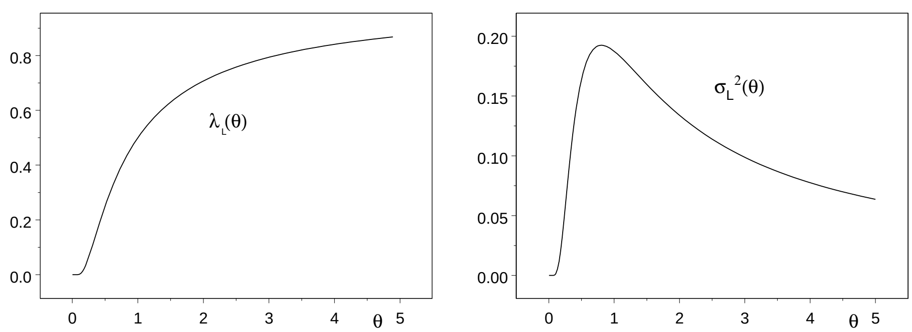

FIG. 1. *Lower tail-dependence coefficient $\lambda_L(\theta)$ (left plot) and corresponding asymptotic variance $\sigma_L^2(\theta)$ as in formula (5.31) (right plot) for the bivariate Pareto copula.*

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 have not 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 9.

6. Strong consistency

THEOREM 7. *Let F be a bivariate 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 $\mathbb{B}_\infty(\bar{I\!R}^2_+)$ (equipped with the metric d as in (A.39)) In particular*

$$\mathbb{P}\big(\lim_{m\to\infty} d(\hat{\Lambda}_{U,m},\Lambda_U)=0\big)=1 \quad \textit{and} \quad \mathbb{P}\big(\lim_{m\to\infty} d(\hat{\Lambda}_{L,m},\Lambda_L)=0\big)=1. \tag{6.33}$$

PROOF. The proof is provided for the upper tail-copula. Recall that a sequence converges in the space $\mathbb{B}_\infty(\bar{I\!R}^2_+)$ with respect to the metric d (cf. (A.39)) 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

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

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

$$+ \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m1}>m-kx\}} - x\Big| + \Big|\frac{1}{k}\sum_{j=1}^m \mathbf{1}_{\{R^{(j)}_{m1}>m-ky\}} - y\Big|$$

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

The proof for the lower tail-copula is similar. □

7. General rank order statistics for extreme events In the present section we restrict ourself to the bivariate lower tail-copula Λ_L. Rank order statistics of the type

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

have been investigated, for example, by Ruymgaart, Shorack and van Zwet (1972), Ruymgaart (1974) and Rüschendorf (1976). Recently, Fermanian et al. (2002) 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:

$$\mathcal{R}_m := \frac{1}{k}\sum_{j=1}^m J\Big(R^{(j)}_{m1}/k, R^{(j)}_{m2}/k\Big). \tag{7.34}$$

The next theorem establishes asymptotic normality of $\mathcal{R}_m$ under certain regularity assumptions on J.

THEOREM 8. *Let F be a bivariate 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*

$$\frac{1}{\sqrt{k}}\sum_{j=1}^m \Big(J\Big(\frac{R^{(j)}_{m1}}{k},\frac{R^{(j)}_{m2}}{k}\Big) - I\!EJ(G(X^{(j)}),H(Y^{(j)}))\Big) \xrightarrow{w} \int_{\bar{I\!R}^2_+} \mathbb{G}_{\Lambda_L}(x,y)dJ(x,y), \tag{7.35}$$

where $\mathbb{G}_{\Lambda_L}$ is a centered continuous Gaussian field as in Theorem 6 and weak convergence takes place in $\mathcal{B}_\infty(\bar{I\!R}^2_+)$. Moreover, the limiting process is also centered Gaussian.

Proof. With the stated prerequisites we have

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

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

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

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

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

Notice that

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

Thus,

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

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

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

which is centered Gaussian. $\square$

8. Multidimensional extensions We consider again the lower (empirical) tail-copula Λ_L ($\hat{\Lambda}_{L,m}$). All results can be similarly stated for the upper (empirical) tail-copula.

Theorem 9. *Let F be a multivariate distribution function with continuous marginal distribution functions. Then the (lower) tail copula $\Lambda_L^{I,J}(x)\not\equiv 0$ exists if for some $j\in J$ the (lower) tail copula $\Lambda_L^{I\cup J\setminus\{j\},\{j\}}(x)\not\equiv 0$ exists. Moreover, (putting $0:=0/0$)*

$$\Lambda_L^{I,J}(x)=\frac{\Lambda_L^{I\cup J\setminus\{j\},\{j\}}(x)}{\Lambda_L^{J\setminus\{j\},\{j\}}(x)}. \tag{8.36}$$

Proof. Suppose $\Lambda_L^{I\cup J\backslash\{j\},\{j\}}(x)\not\equiv 0$ exists for some $j\in J$. Then $\Lambda_L^{J\backslash\{j\},\{j\}}(x)>0$ for all x according to the proof of part v) of Theorem 1. By Definition 1 we conclude that $\Lambda_L^{I,J}(x)\not\equiv 0$ exists and equation (8.36) holds. □

A nonparametric estimator for the lower tail-copula is defined by

$$\hat\Lambda_{L,m}^{I,J}(x) := \sum_{j=1}^{m}\mathbf{1}_{\{R_{ml}^{(j)}\le kx_l,\forall l\in I\cup J\}}\Big/\sum_{j=1}^{m}\mathbf{1}_{\{R_{ml}^{(j)}\le kx_l,\forall l\in J\}} \tag{8.37}$$

with $k=k(m)\to\infty$ and $k/m\to 0$ as $m\to\infty$.

Theorem 10 Asymptotic normality. *Assume that the prerequisites of Theorem 9 are fulfilled, Λ_L possesses continuous partial derivatives, and the multivariate version of the Second order condition 4 holds. Then, for $\sqrt{k}A(m/k)\to 0$ as $m\to\infty$*

$$\sqrt{k}\{\hat\Lambda_{L,m}^{I,J}(x)-\Lambda_L^{I,J}(x)\}\xrightarrow{w}\mathbb{G}_{\Lambda_L}^{I,J}(x),$$

where $\mathbb{G}_{\Lambda_L}^{I,J}(x)$ is a centered tight continuous Gaussian random field, and weak convergence takes place in $\mathbb{B}_\infty(\bar{I\!R}_+^n\backslash[0,\varepsilon)^n)$ for every $\varepsilon>0$.

The *proof* follows along the same lines as the proof of Theorem 6.

Theorem 11 Strong consistency. *Let F be a multivariate distribution function with continuous marginal distribution functions. With the prerequisites of Theorem 9 and $k/\log\log m\to\infty$ as $m\to\infty$, $\hat\Lambda_{L,m}^{I,J}(x)$ converges almost surely to $\Lambda_L^{I,J}(x)$ in the space $\mathbb{B}_\infty(\bar{I\!R}_+^n\backslash[0,\varepsilon)^n)$ for every $\varepsilon>0$.*

Proof. For some $j\in J$, put $I_1:=I\cup J\backslash\{j\}$ and $I_2:=J\backslash\{j\}$. Let $k/\log\log m\to\infty$. Then

$$\begin{aligned}
&|\hat\Lambda_{L,m}^{I,J}(x)-\Lambda_L^{I,J}(x)|\\
&=|\hat\Lambda_{L,m}^{I_1,\{j\}}(x)-\Lambda_L^{I_1,\{j\}}(x)|/\Lambda_L^{I_2,\{j\}}(x)\\
&\quad+\frac{\hat\Lambda_{L,m}^{I_1,\{j\}}(x)}{\hat\Lambda_{L,m}^{I_2,\{j\}}(x)}\frac{1}{\Lambda_L^{I_2,\{j\}}(x)}\left|\hat\Lambda_{L,m}^{I_2,\{j\}}(x)-\Lambda_L^{I_2,\{j\}}(x)\right|\\
&=: C_{1,m}(x)+C_{2,m}(x).
\end{aligned}$$

Regarding the first term: $C_{1,m}(x)\to 0$ as $m\to\infty$ almost surely. This follows essentially from Theorem 1.2 in Qi (1997) and the fact that for every $\varepsilon>0$, the estimator $\hat\Lambda_{L,m}$ is bounded from above and below by strictly positive constants on $T_i\backslash[0,\varepsilon)^n$ (cf. Definition 4). The second term: $C_{2,m}(x)\to 0$ as $m\to\infty$ almost surely with the same reasoning. □

9. Simulation and empirical study The simulation study considers four 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 introduced in Section 4. Second, we construct asymptotic confidence intervals related to one particular TDC estimator and discuss the applicability of the proposed approximation method. Following, a robustness study regarding the choice of the threshold k is provided. Fourth, we investigate the finite-sample behavior of the nonparametric estimator for the tail copula, i.e., the empirical tail copula. Finally, the TDCs and the tail copulae are estimated for two well-known financial time series.

9.1. *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 and Schmidt (2003)). We restrict ourself 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.

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 1
Sample-bias and MSE for the nonparametric upper TDC estimator $\hat{\lambda}_U = \hat{\lambda}_U^S$ (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)

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).

Note that for spherically contoured distributions the lower and upper TDCs coincide, in which case one can double the data for the upper TDC-estimation by mirroring them at the origin.

In Figures 2 and 3 we provide plots illustrating the results of the TDC estimation. Presented are 3×1000 TDC estimations for $m = 500, 1000, 2000$ in a consecutive ordering. These plots visualize the decreasing empirical bias and variance for increasing samplesize m.

As the main result we conclude that the TDC estimator $\hat{\lambda}_U = \hat{\lambda}_U^S$ 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 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.

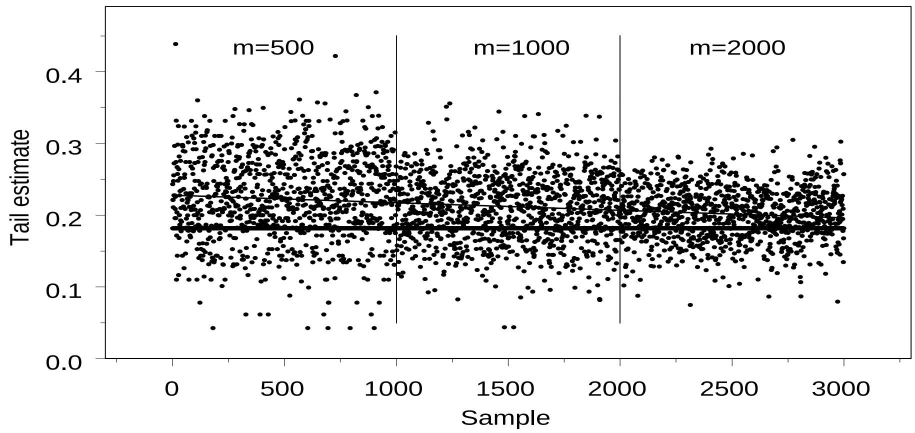

FIG. 2. *Nonparametric upper TDC estimates* $\hat{\lambda}_U = \hat{\lambda}_U^S$ *for* 3×1000 *iid samples of size* $m = 500, 1000, 2000$ *from a bivariate t-distribution with parameters* $\nu = 2$, $\rho = 0$, *and* $\lambda_U = 0.1817$. *The thick line corresponds to* λ_U *and the thin line denotes the regression line.*

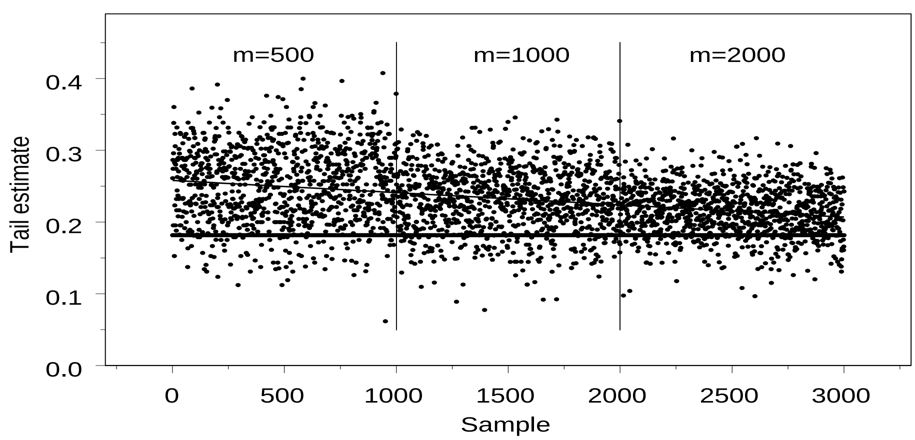

FIG. 3. *Nonparametric upper TDC estimates* $\hat{\lambda}_U^{EVT}$ *for* 3×1000 *iid samples of size* $m = 500, 1000, 2000$ *from a bivariate t-distribution with parameters* $\nu = 2$, $\rho = 0$, *and* $\lambda_U = 0.1817$. *The thick line corresponds to* λ_U *and the thin line denotes the regression line.*

9.2. *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 ourself 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 (5.32). 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 Pareto copula itself and for the t-copula.

Consider 200 independent copies of $m = 500, 1000$ iid pseudo-random vectors which are generated from a bivariate Pareto copula with various parameters θ. The threshold k is chosen according to the plateau algorithm described above. The estimation results are illustrated in Figures 4 and 5. Figure 4 shows that the sample-means of the nonparametric estimator $\hat{\lambda}_L$ and the approximative standard deviation $\sigma_L(\hat{\theta})$ nearly coincide with the true values. However, there is an increasing bias observable for smaller parameters θ. The estimation improves with increasing samplesize. The results are characteristic for all other estimations which we have not listed in this work. Figure 5 represents the means of the approximated confidence intervals. Again, the larger sample-bias for smaller values of θ result in an asymmetric confidence band for smaller θ.

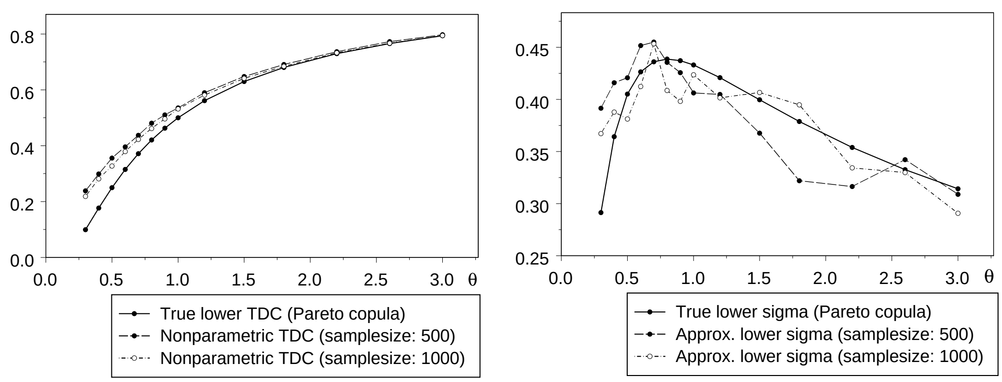

FIG. 4. **Left:** *True lower TDC for the Pareto copula for various θ and the corresponding sample-means of the nonparametric estimator $\hat{\lambda}_L$ for samplesizes $m = 500, 1000$.* **Right:** *True asymptotic standard deviation $\sigma_L(\theta)$ and the corresponding sample-means of the approximated standard deviations $\sigma_L(\hat{\theta})$ (see formula (5.32)) for samplesizes $m = 500, 1000$, where θ results from an ML-fitted Pareto copula.*

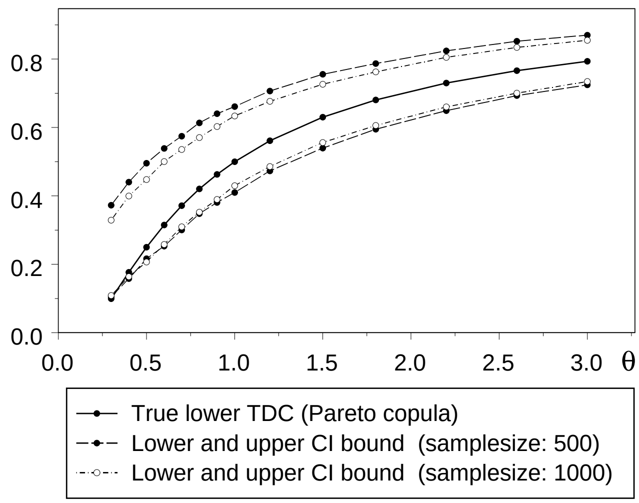

FIG. 5. *True lower TDC for the Pareto copula for various θ and the sample-means of the approximated confidence intervals (to the confidence level $\alpha = 0.05$) for samplesizes $m = 500, 1000$.*

Now 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 sample-means of the nonparametric estimator $\hat{\lambda}_L$ and the approximated confidence intervals for the t-copula are presented in Figure 6. Observe that the sample-bias is much smaller in comparison to the Pareto copula.

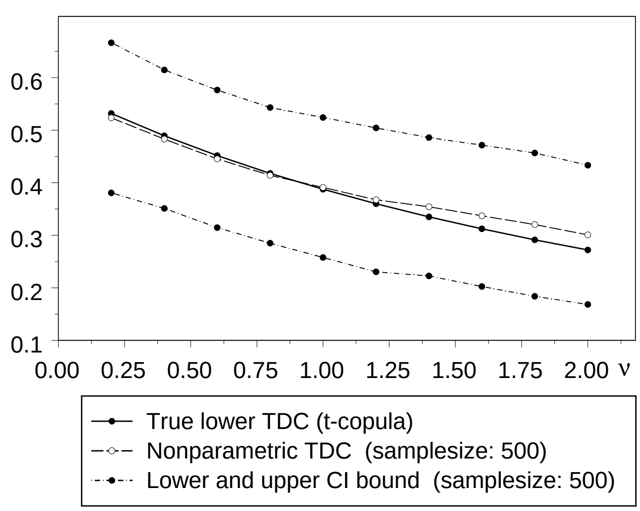

FIG. 6. *True lower TDC for the t-copula for various parameters ν and correlation coefficient $\rho = 0.25$, the corresponding sample-means of the nonparametric estimator $\hat{\lambda}_L$, and the sample-means of the approximated confidence intervals (confidence level $\alpha = 0.05$) for samplesize $m = 500$.*

The quality of the approximation of the true confidence interval is illustrated in Table 3. In contrast to the Pareto copula (see Table 4), 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 representing 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.

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

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 $\alpha = 0.05$).

9.3. *Robustness of the estimation with respect to the threshold choice* In the following we investigate the sensitivity of the nonparametric estimator $\hat{\lambda}_L$ regarding different choices of the threshold k. Further, we analyze the sensitivity of the sample standard deviation $\hat{\sigma}_L$. We consider 200 independent copies of 500 iid pseudo-random vectors which are generated from a bivariate Pareto copula with various parameters θ. The simulation results are presented in Table 4. The main result shows that the estimator $\hat{\sigma}_L$ is not excessively sensitive towards different threshold choices of the following type: We chose k equal to 0.05%, 0.1%, and 0.15% of the total samplesize of 500. This justifies the usage of the plateau algorithm for the threshold choice as described above. However, for very small TDCs we observe a quite sever sample bias.

θ	λ_L	$mean(\hat{\lambda}_L)$ $k = 0.05\%$	$mean(\hat{\lambda}_L)$ $k = 0.1\%$	$mean(\hat{\lambda}_L)$ $k = 0.15\%$	σ_L	$mean(\sigma_L(\hat{\theta}))$	$\hat{\sigma}_L$ $k = 0.05\%$	$\hat{\sigma}_L$ $k = 0.1\%$	$\hat{\sigma}_L$ $k = 0.15\%$
0.3	0.099	0.204	0.254	0.294	0.291	0.280	0.38	0.384	0.405
0.4	0.177	0.258	0.304	0.342	0.364	0.355	0.453	0.413	0.424
0.5	0.25	0.306	0.35	0.378	0.405	0.400	0.45	0.415	0.424
0.6	0.315	0.361	0.386	0.415	0.426	0.423	0.476	0.44	0.465
0.7	0.371	0.399	0.426	0.447	0.436	0.434	0.463	0.451	0.45
0.8	0.42	0.44	0.469	0.487	0.439	0.437	0.449	0.446	0.429
0.9	0.463	0.472	0.498	0.513	0.437	0.436	0.449	0.471	0.416
1	0.5	0.502	0.525	0.536	0.433	0.432	0.473	0.436	0.412
1.2	0.561	0.559	0.571	0.589	0.421	0.420	0.435	0.443	0.42
1.5	0.63	0.62	0.635	0.643	0.4	0.399	0.411	0.372	0.382
1.8	0.68	0.674	0.68	0.687	0.379	0.378	0.441	0.36	0.362
2.2	0.73	0.715	0.728	0.728	0.354	0.354	0.393	0.33	0.382
2.6	0.766	0.764	0.759	0.763	0.333	0.332	0.355	0.363	0.318
3	0.794	0.786	0.792	0.79	0.314	0.314	0.328	0.342	0.308

TABLE 4

Sensitivity analysis of the nonparametric estimator $\hat{\lambda}_L$ *and the sample-standard deviation* $\hat{\sigma}_L$ *regarding three different choices of the threshold* k, *i.e.,* k *equals* 0.05%, 0.1%, *and* 0.15% *of the total samplesize of* 500.

9.4. *Estimation of the tail copula* So far we have concentrated on the special case of estimating the TDC via the nonparametric estimators $\hat{\lambda}_L$ and $\hat{\lambda}_U$ given in (4.18) and (4.19). Now we turn to the estimation of the entire tail copula utilizing the proposed nonparametric estimators. 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 (4.15). For the simulation experiment we utilize 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 7.

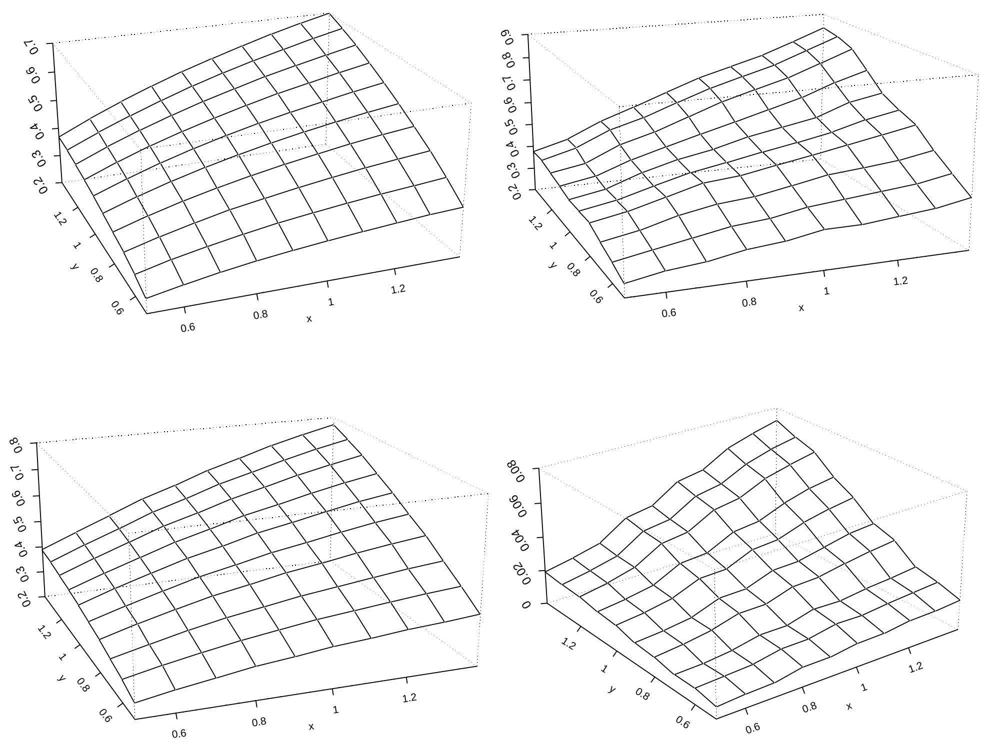

FIG. 7. **Upper left:** *True lower tail-copula of the Pareto copula with parameter* $\theta = 1$. **Upper right:** *Nonparametric estimate* $\hat{\Lambda}_L(x, y)$ *stated in (4.15) for one particular sample-set with samplesize* 500. **Lower left:** *Mean of nonparametric estimates* $\hat{\Lambda}_L(x, y)$ *for* 200 *samples of a Pareto copula with parameter* $\theta = 1$ *and samplesize* 500. **Lower right:** *Mean-bias of nonparametric estimates* $\hat{\Lambda}_L(x, y)$ *for* 200 *samples of a Pareto copula with parameter* $\theta = 1$ *and samplesize* 500.

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 above 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.

A particular point of interest concerns the estimation of the tail copula in case the data result form 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 (see Bingham et al. (2003)). The left picture in Figure 8 reveals that in case of the standard normal distribution (copula) the tail-copula estimates are performing good. 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 8.

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.

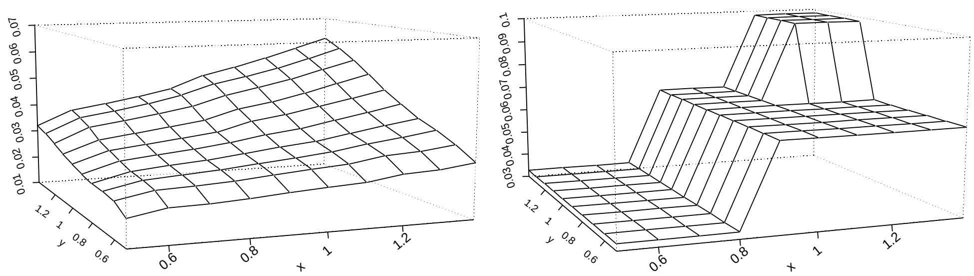

FIG. 8. **Left:** *Mean of nonparametric estimates* $\hat{\Lambda}_L(x, y)$ *for* 200 *samples of a data set of samplesize* 500 *generated from a bivariate standard normal distribution.* **Right:** *Particular nonparametric estimate* $\hat{\Lambda}_L(x, y)$ *for a data set of samplesize* 500 *generated from a normal distribution with parameter* $\rho = 0.25$.

9.5. *Application to financial data* The present section reveals that tail dependence is indeed often found in financial data. Provided are two scatter plots of daily negative log-returns of a tuple 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 samplesize). 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. Both plots show 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 9.1. 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 $\hat{\lambda}_{U,D_2} = 0.14$ with k between 40 and 100.

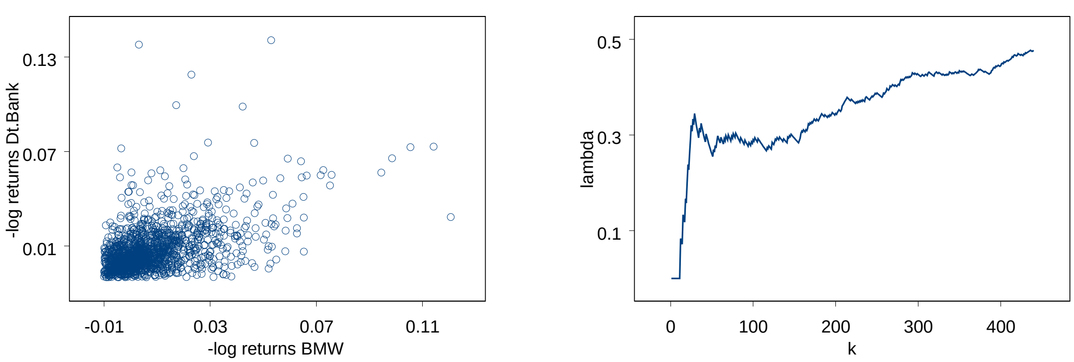

FIG. 9. *Scatter plot of BMW versus Dt. Bank negative daily stock log-returns (2325 data points) and the corresponding TDC estimates* $\hat{\lambda}_U$ *for various* k.

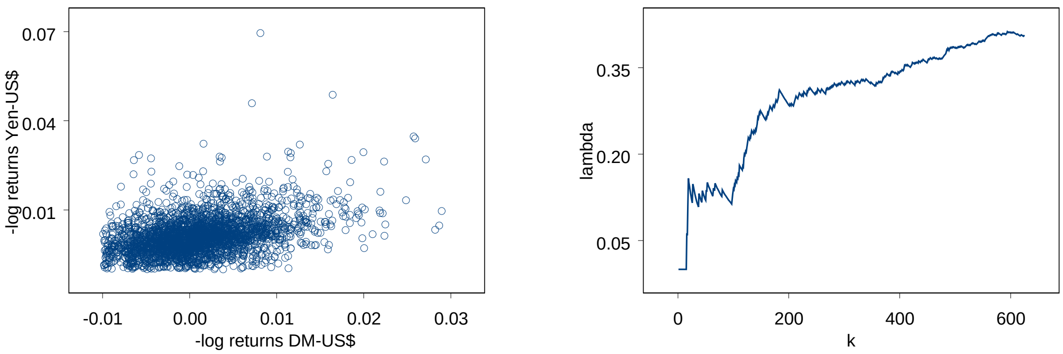

Fig. 10. *Scatter plot of DM-USD versus Yen-USD negative daily exchange-rate log-returns (3126 data points) and the corresponding TDC estimates $\hat{\lambda}_U$ for various k.*

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. See Section 8 for an extension of the bivariate tail-dependence concept to the multidimensional framework. Observe that upper tail-dependence for a bivariate random vector $(X_1, X_2)'$ is equivalent to

$$\lambda_U = \lim_{\alpha \to 0} \mathbb{P}(X_2 > VaR_{1-\alpha}(X_2) | X_1 > VaR_{1-\alpha}(X_1)) > 0. \tag{9.38}$$

Finally, in Figure 11 we provide the estimation of the tail copula related to both financial data sets.

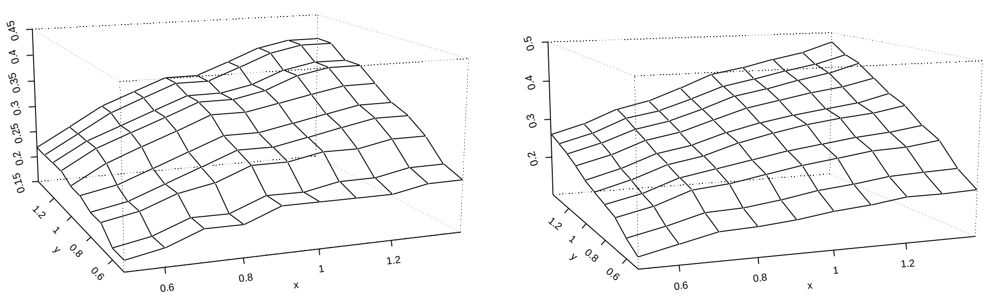

Fig. 11. **Left:** *Nonparametric tail-copula estimate $\hat{\Lambda}_L(x, y)$ related to negative daily stock log-returns of BMW versus Dt. Bank.* **Right:** *Nonparametric tail-copula estimate $\hat{\Lambda}_L(x, y)$ related to negative daily exchange-rate log-returns of DM-USD versus Yen-USD.*

10. Conclusion Summarizing the results we have introduced an appealing concept to model extremal dependencies of random vectors, namely the concept of tail copulae. We have shown that tail copulae have several analogies to the well-known theory of copulae. Further we provided various nonparametric estimators for the tail copula and the tail-dependence coefficient, and we proved asymptotic normality and strong consistency. Within a simulation study we showed that the finite sample behavior of the estimators under consideration is very satisfying. Beside other results, our simulations pointed out that, in addition to solely glancing on

the tail dependence coefficient, the estimation of the tail copula helps a lot to decide whether data are tail dependent or not.

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 uniform metric on compacta and establish necessary and sufficient conditions for weak convergence. The latter space turns out to be useful in the context of empirical tail-copulae. Further, a general *Delta method* is formulated which we utilized to prove asymptotic normality for empirical tail-copulae.

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 living in some space $(\mathbb{D}, \mathcal{D})$, these concepts have to be modified as no probability measures can be defined on the corresponding Borel σ-field $\mathcal{D}$. Loosely speaking, the Borel σ-field turns out to be too large. To overcome this obstacle, several approaches can be distinguished. First, we 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 a way 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 and 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 ($\xrightarrow{w}$) 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 $\mathbb{D}$.

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, precisely all functions $f : T \to \mathbb{R}$ 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 $\bar{\mathbb{R}}_+^2$. If not stated otherwise, T will always denote a compact subset of $\bar{\mathbb{R}}_+^2$ in the following.

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

Skorokhod construction (see Neuhaus (1971)). The space $l^\infty(T)$ will be equipped with the Borel σ-field. The only measurability property we require for X_m relates to the measurability of the maps $X_m(t) : \Omega_m \to \mathbb{R},\ t \in T$ (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 bridge 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 projection σ-field.

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

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

$$d(f_1, f_2) = \sum_{i=1}^{\infty} 2^{-i}(||f_1 - f_2||_{T_i} \wedge 1) \tag{A.39}$$

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

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

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

Theorem 12. *For each $m \geq 1$, let $X_m : \Omega_m \to \mathcal{B}_\infty(\bar{\mathbb{R}}^2_+)$ 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 (see Casella and Berger (2002), Section 5.5.4) is a well-known technique in statistics to prove results concerning asymptotic normality of certain 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 $(\mathbb{D}, d)$ and $(\mathbb{E}, e)$ be metrizable or topological vector spaces, in particular vector addition and scalar multiplication are continuous operations.

Definition 5 Hadamard differentiability. *A map $\phi : \mathbb{D}_\phi \subset \mathbb{D} \to \mathbb{E}$ is called Hadamard-differentiable at $\theta \in \mathbb{D}_\phi$ if there exists a continuous linear map $\phi'_\theta : \mathbb{D} \to \mathbb{E}$ such that*

$$\frac{\phi(\theta + t_m h_m) - \phi(\theta)}{t_m} \to \phi'_\theta(h), \quad \text{as } m \to \infty, \tag{A.40}$$

for all converging sequences $t_m \to 0$ and $h_m \to h$ such that $\theta + t_m h_m \in \mathbb{D}_\phi$ for all m. Further, $\phi : \mathbb{D}_\phi \subset \mathbb{D} \to \mathbb{E}$ is called Hadamard-differentiable tangentially to a set $\mathbb{D}_0 \subset \mathbb{D}$ by requiring that $h_m \to h$ with $h \in \mathbb{D}_0$. In that case the derivative ϕ'_θ needs only to be defined on $\mathbb{D}_0$.

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 13 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 $r_m(X_m - \theta) \xrightarrow{w} X$ for some sequence of constants $r_m \to \infty$, where $X : \Omega \to \mathbb{D}_\phi$ is separable. Then*

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

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

A.2. *Hadamard differentiability* The proof of Theorem 6 (asymptotic normality of the empirical tail-copula) in Section 5 needs the following lemma. The lemma is stated in the original version as provided in Van der Vaart and 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*

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

Acknowledgments. Both authors gratefully acknowledge the financial support of Daimler-Chrysler AG Research and Technologies. This work was partly supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD).

REFERENCES

Araujo, A. and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*, John Wiley & Sons, New York.

Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae, *Scandinavian Journal of Statistics* **5**: 151–157.

Baron, S., Liflyand, E. and Stadtmüller, U. (2000). Complementary spaces and multipliers of double fourier series for functions of bounded variation, *Journal of Mathematical Analysis and Applications* **250**: 706–721.

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

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

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

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

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

Cuculescu, I. and Theodorescu, R. (2001). Copulas: diagonals, tracks, *Rev. Roumaine Math. Pures Appl.* **46**: 731-742.

de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order, *Journal of the Australian Mathematical Society* **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. and Schmid, F. (2002). Nonparametric estimation of the Lower Tail Dependence λ_L in Bivariate Couplas, *University of Cologne preprint*.

Draisma, G., Drees, H., Ferreira, A. and de Haan, L. (2001). Tail dependence in independence, *Eurandom preprint*.

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

Dudley, R. M. (1967). Measures on nonseparable metric spaces, *Illinois Journal of Mathematics*.

Durrett, R. (1996). *Probability Theory and Examples*, 2. edn, Duxbury Press.

Durrleman, V., Nikeghbali, A. and Roncalli, T. (2000). Which copula is the right one?, *Groupe de Recherche Opérationnelle, Crédit Lyonnais, working paper*.

Einmahl, J., de Haan, L. and Piterbarg, V. (2001). Multivariate extremes estimation, *Annals of Statistics* **19**: 1401–1423.

Embrechts, P., Lindskog, F. and 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.

Fermanian, J.-D., Radulović, D. and Wegkamp, M. (2002). Weak convergence of empirical copula processes, *preprint*.

Frahm, G., Junker, M. and Schmidt, R. (2002). Estimating the tail-dependence coefficient, *http://stats.lse.ac.uk/schmidt*.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate archimedean copulas, *Journal of the American Statistical Association* **88**: 1034–1043.

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

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

Huang, X. (1992). *Statistics of Bivariate Extreme Values*, PhD thesis, Tinbergen Institute Research Series, Thesis Publishers and Tinbergen Institute.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Monographs on Statistics and Applied Probabilty 73, Chapman and Hall, London.

Karolyi, G. and Stulz, R. (1996). Why do markets move together? an investigation of u.s.-japan stock return comovements, *Journal of Finance* **51**(3): 951–989.

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

Ledford, A. and Tawn, J. (1997a). Concomitant tail behavior for extremes, *Advances in Applied Probability* **30**: 197–215.

Ledford, A. and Tawn, J. (1997b). Modelling dependence within joint tail regions, *Journal of the Royal Statistical Society* **59**: 475–499.

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

Nelsen, R. (1999). *An Introduction to Copulas*, Lecture Notes in Statistics 139, Springer Verlag.

Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time paramter, *Annals of Mathematical Statistics* **42**: 1285–1295.

Malevergne, Y. and Sornette, D. (2003). Investigating Extreme Dependencies, *forthcoming in: The Review of Financial Studies*.

Ong, M. (1999). *Internal Credit Risk Models*, Risk Books, Haymarket.

Peng, L. (1998). *Second Order Condition and Extreme Value Theory*, Vol. 178 of *PhD thesis, Tinbergen Institute Research Series*, Thesis Publishers and Tinbergen Institute.

Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes, *Statistics & Probability Letters* **43**: 399–409.

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

Ruymgaart, F. and 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, *Annals of Statistics* **2**: 892–910.

Ruymgaart, F. H., Shorack, G. R. and 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, *Mathematical Methods of Operations Research* **55**(2): 301–327.

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, and M.D. Taylor, Institute of Mathematical Statistics, Hayward: 1–14.

Skorokhod, A. V. (1956). Limit theorems for stochastic processes, *Theory of Probability and Its Applications* **1**: 261–290.

Song, P. X.-K. (2000). Multivariate dispertion models generated from gaussian copula, *Scandianvian Journal of Statistics* **27**(2): 305–320.

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

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

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

London School of Economics
Department of Statistics
Houghton Street
London, WC2A 2AE, UK

Universität Ulm
Abteilung Zahlentheorie und Wahrscheinlichkeitstheorie
89069 Ulm, Germany

Correspond to: r.schmidt@lse.ac.uk