Estimating the tail-dependence coefficient: Properties and pitfalls

Gabriel Frahm* Markus Junker† Rafael Schmidt‡

May 26, 2006

Abstract

The concept of tail dependence describes the amount of dependence in the lower-left-quadrant tail or upper-right-quadrant tail of a bivariate distribution. A common measure of tail dependence is given by the so-called tail-dependence coefficient. This paper surveys various estimators for the tail-dependence coefficient within a parametric, semiparametric, and nonparametric framework. Further, a detailed simulation study is provided which compares and illustrates the advantages and disadvantages of the estimators.

Frahm et al., 2005, Estimating the tail dependence coefficient. *Insurance: Mathematics and Economics* **37**, 80-100.

1 Motivation

During the last decade, dependencies between financial asset returns have increased due to globalization effects and relaxed market regulation. However, common dependence measures such as Pearson's correlation coefficient are not always suited for a proper understanding of dependencies in financial markets; see, e.g., Embrechts et al. (2002). In particular, dependencies between extreme events such as extreme negative stock returns or large portfolio losses cause the need for alternative dependence measures to support beneficial asset-allocation strategies.

Several empirical surveys such as Ané and Kharoubi (2003) and Malevergne and Sornette (2004) exhibited that the concept of tail dependence is a useful tool to describe the dependence between extremal data in finance. Moreover, they showed that especially during

*University of Cologne. Email: frahm@statistik.uni-koeln.de.

†Email: ma.junker@gmx.de. The first two authors gratefully acknowledge support by the Center of Advanced European Studies and Research (CAESAR). The main part of this research was carried out at CAESAR.

‡University of Cologne and London School of Economics. Email: rafael.schmidt@uni-koeln.de. The author gratefully acknowledges fellowship support from the Postdoc-Programme of the German Academic Exchange Service (DAAD).

volatile and bear markets, tail dependence plays a significant role. In this context, tail dependence is described via the so-called *tail-dependence coefficient* (TDC) introduced by Sibuya (1960). This concept is reviewed in Section 2.

However, actuaries and statisticians who are not familiar with *extreme value theory* (EVT) often have difficulties in choosing appropriate methods for measuring or estimating tail dependence. One reason for that is the limited amount of (extremal) data which makes the estimation quite sensitive to the choice of method. Another reason is the lack of literature which compares the various estimators developed in (mostly theoretical) articles related to EVT. This paper tries to partially fill this gap by surveying and comparing various methods of tail-dependence estimation. In other words, we will present the most common estimators for the TDC and compare them via a simulation study.

TDC estimators are either based on the entire set of observations or on extremal data. Regarding the latter, EVT is the natural choice for inferences on extreme values. In the one-dimensional setting, the extreme value distributions can be expressed in parametric form, as shown by Fisher and Tippett (1928). Thus it suffices to apply parametric estimation methods only. By contrast, multidimensional extreme value distributions cannot be characterized by a fully parametric model in general. This leads to more complicated estimation techniques.

Parametric estimation methods are efficient if the distribution model under consideration is true, but they suffer from biased estimates in case the underlying model is different. Nonparametric estimation procedures avoid this type of model error but come along with a larger estimation variance. Accordingly, we distinguish in Section 3 between the following types of TDC estimations, namely, TDC estimations which are based on:

a) a specific distribution or a family of distributions;

b) a specific copula or a family of copulas; or

c) a nonparametric model.

We discuss properties of the estimators along with possible applications and give references for further reading. Section 4 presents a detailed simulation study which analyzes and compares selected estimators regarding their finite sample behavior. Statistical methods testing for tail dependence or tail independence are not included in this work. An account on that topic can be found for instance in Draisma et al. (2004) or Falk and Michel (2004).

2 Preliminaries

The following approach, discussed by Sibuya (1960) and Joe (1997, p. 33) among others, represents the most common definition of tail dependence. Let (X, Y) be a random pair with joint cumulative distribution function F and marginals G (for X) and H (for Y). The quantity

$$\lambda_U = \lim_{t \to 1^-} \mathrm{P}\{G(X) > t \mid H(Y) > t\} \tag{1}$$

is called the *upper tail-dependence coefficient* (upper TDC), provided the limit exists. We say that (X, Y) is *upper tail dependent* if $\lambda_U > 0$ and *upper tail independent* if $\lambda_U = 0$. Similarly, we define the lower tail-dependence coefficient by

$$\lambda_L = \lim_{t \to 0^+} \mathrm{P}\{G(X) \le t \mid H(Y) \le t\}. \tag{2}$$

Thus, the TDC roughly corresponds to the probability that one margin exceeds a high/low threshold under the condition that the other margin exceeds a high/low threshold.

The TDC can also be defined via the notion of copula, introduced by Sklar (1959). A copula C is a cumulative distribution function whose margins are uniformly distributed on $[0, 1]$. As shown by Sklar (1959), the joint distribution function F of any random pair (X, Y) with marginals G and H can be represented as

$$F(x, y) = C\{G(x), H(y)\}. \tag{3}$$

in terms of a copula C which is unique when G and H are continuous, as will be assumed in the sequel. Refer to Nelsen (1999) or Joe (1997) for more information on copulas.

If C is the copula of (X, Y), then

$$\lambda_L = \lim_{t \to 0^+} \frac{C(t,t)}{t} \quad \text{and} \quad \lambda_U = \lim_{t \to 1^-} \frac{1 - 2t + C(t,t)}{1-t}.$$

Another representation of the upper TDC is given by $\lambda_U = \lim_{s \to 0^+} \tilde{C}(s, s)/s$, where $\tilde{C}(1-t, 1-t) = 1 - 2t + C(t, t)$ denotes the *survival copula* of C. Thus, the upper TDC of C equals the lower TDC of its survival copula and, vice versa, the lower TDC of C is given by the upper TDC of $\tilde{C}$. Since the TDC is determined by the copula of X and Y, many copula features transfer directly to the TDC. For instance, the TDC is invariant under strictly increasing transformations of the margins.

Consider a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ of observations of (X, Y). Let

$$X_n^* = \max(X_1, \ldots, X_n) \quad \text{and} \quad Y_n^* = \max(Y_1, \ldots, Y_n)$$

be the corresponding componentwise maxima. In order to have a meaningful discussion about tail dependence in the EVT framework, we assume that F belongs to the domain of attraction of an extreme value (EV) distribution. This means that as $n \to \infty$, the joint distribution of the standardized componentwise maxima X_n^* and Y_n^* has the following limiting EV distribution (with non-degenerated margins):

$$F^n(a_n x + b_n, c_n y + d_n) \to F_{EV}(x, y)$$

for some standardizing sequences $(a_n), (c_n) > 0$ and $(b_n), (d_n) \in \mathbb{R}$. Suppose that F_{EV} has unit Fréchet margins G_{EV} and H_{EV}, i.e.,

$$G_{EV}(x) = \exp(-1/x), \ x > 0 \quad \text{and} \quad H_{EV}(y) = \exp(-1/y), \ y > 0.$$

This assumption, which is standard in the EVT framework, is similar to the assumption that the margins can be transformed into uniform distributions in the theory of copulas. Then the EV distribution possesses the following representation, Pickands (1981):

$$F_{EV}(x,y) = \exp\left\{-\left(\frac{1}{x}+\frac{1}{y}\right)A\left(\frac{y}{x+y}\right)\right\}, \qquad x,y>0. \tag{4}$$

Here $A:[0,1]\to[1/2,1]$ is a convex function such that $\max(t,1-t)\le A(t)\le 1$ for every $0\le t\le 1$. The function A is known as Pickands' dependence function. In the sequel, the term *dependence function* always refers to the above representation and should not be confused with the copula of a bivariate random vector.

The copula C^*_ℓ, $\ell\in I\!N$, of the componentwise maxima X^*_l and Y^*_l is related to the copula C as follows:

$$C^*_\ell\,(u,v) = C^\ell\left(u^{1/\ell},v^{1/\ell}\right), \qquad 0\le u,v\le 1.$$

If the diagonal section $C\,(t,t)$ is differentiable for $t\in(1-\varepsilon,1)$ for some $\varepsilon>0$, then it can be shown that

$$2-\lambda_U = \lim_{t\to 1^+}\frac{1-C\,(t,t)}{1-t} = \lim_{t\to 1^+}\frac{1-C^*_\ell\,(t,t)}{1-t} = \lim_{t\to 1^+}\frac{dC\,(t,t)}{dt} = \lim_{t\to 1^+}\frac{dC^*_\ell\,(t,t)}{dt} \tag{5}$$

for all $\ell\in I\!N$. In particular, for $\ell\to\infty$ we obtain

$$C_{EV}\,(t,t) = F_{EV}\left\{-\frac{1}{\log(t)},-\frac{1}{\log(t)}\right\} = t^{2A(1/2)}, \qquad 0<t<1, \tag{6}$$

where C_{EV} denotes the copula of F_{EV}. This implies the following important relationship:

$$\lambda_U = 2-2A\left(\frac{1}{2}\right).$$

Another representation of the EV distribution is frequently encountered in the EVT literature. If F_{EV} has unit Fréchet margins, there exists a finite spectral measure S on $\mathcal{B}=\{(x,y):\,x,y>0,\|(x,y)\|_2=1\}$, where $\|\cdot\|_2$ denotes the Euclidean norm, such that

$$F_{EV}(x,y) = \exp\left\{-\int_{\mathcal{B}}\max\left(\frac{u}{x},\frac{v}{y}\right)\,dS(u,v)\right\}, \qquad x,y>0,$$

with $\int_{\mathcal{B}} u\,dS(u,v)=1$ and $\int_{\mathcal{B}} v\,dS(u,v)=1$. This yields

$$\lambda_U = 2-\int_{\mathcal{B}}\max(u,v)\,\,dS(u,v)$$

and $A(1/2)=\int_{\mathcal{B}}\max(u,v)\,dS(u,v)/2$. The estimation of the spectral measure is discussed by Joe et al. (1992), de Haan and Resnick (1993), Einmahl et al. (1993), Einmahl et al. (1997), and Capéraà and Fougères (2000), among others.

Thus *any* estimator of the upper TDC $\hat{\lambda}_U$ (the index n is dropped for notational convenience) is equivalent to some estimator $\widehat{A}_n(1/2)$ via the relationship $\hat{\lambda}_U = 2 - 2\widehat{A}_n(1/2)$. By considering the dependence function related to the survival copula, this holds also for the lower TDC. An abundant literature exists concerning the estimation of the dependence function A. See for instance Tiago de Oliveira (1984), Tawn (1988), Smith et al. (1990), Hutchinson and Lai (1990) or Coles and Tawn (1991) for fitting parametric (structural) models to A. By contrast, Pickands (1981), Deheuvels (1991), Joe et al. (1992), Abdous et al. (1998), Capéraà and Fougères (2000) or Falk and Reiss (2003) consider nonparametric estimation procedures.

Due to the invariance of the TDC of (X_ℓ^*, Y_ℓ^*) with respect to ℓ, the following estimator arises quite naturally:

$$\hat{\lambda}_U = 2 - 2\widehat{A}_m\left(\frac{1}{2}\right) = 2 - \frac{\widehat{dC_m}}{dt}(t,t)\bigg|_{t\approx 1}, \quad 1 \leq m \leq n.$$

Here $\widehat{dC_m/dt}$ denotes the estimated derivative of the diagonal section of the copula C_ℓ^* from m block maxima, where each block contains $\ell = n/m$ elements of the original data set (we choose m such that $n/m \in I\!N$). The special case $m = n$ (i.e., $\ell = 1$) corresponds to n block maxima which form the original data set. Every TDC estimator has to deal with a bias-variance trade-off arising from the following two sources. The first one is the choice of the threshold t. That is, the larger t the smaller the bias (and the larger the variance) and vice versa. The second source is the number of block maxima. Thus, the larger m the smaller the variance but the larger the bias. An optimal choice of m and t, e.g., with respect to the mean squared error (MSE) of the estimator, is usually difficult to derive. A similar problem exists for univariate tail-index estimations of regular varying distributions.

In Figure 1, we illustrate the latter bias-variance problem via the following estimator which is motivated by (5) and forms the nonparametric counterpart of the parametric estimator $\hat{\chi}$ introduced in Coles et al. (1999):

$$\hat{\lambda}_U^{\mathrm{LOG}} = 2 - \frac{\log \widehat{C}_m\left(\frac{m-k}{m}, \frac{m-k}{m}\right)}{\log\left(\frac{m-k}{m}\right)}, \qquad 0 < k < m, \tag{7}$$

where

$$\widehat{C}_m(u,v) = \frac{1}{m}\sum_{j=1}^{m} 1\!\!1\left(R_{1j}/m \leq u, R_{2j}/m \leq v\right)$$

is called the *empirical copula*. Here $1\!\!1$ denotes the indicator function, while R_{1j} and R_{2j}, respectively, are the ranks of the block maxima $X_{\ell j}^*$ and $Y_{\ell j}^*$, $j = 1, \ldots, m$, $\ell = n/m$. The threshold is denoted by k. As expected, Figure 1 reveals that the estimation via block maxima has a lower bias but a larger variance. The bias-variance tradeoff for various thresholds can be clearly seen, too.

In order to ease the presentation we do not explicitly differentiate between block maxima and the original data set in the forthcoming sections.

3 TDC estimation

The following estimation approaches are classified by the degree of prior information which is available about the distribution of the data. We will either assume a specific distribution or a class of distributions, a specific copula or a class of copulas, or we perform a completely nonparametric estimation. For notational convenience, λ will be written without the subscript L or U whenever we know that $\lambda_L = \lambda_U$. Moreover, the subscript is dropped whenever we neither specifically refer to the upper nor to the lower TDC.

3.1 Estimation using a specific distribution

Suppose that the distribution $F(\cdot\,;\theta)$ is known. Further assume that λ can be represented via a *known* function of θ, i.e., $\lambda = \lambda(\theta)$. Also assume that F allows for tail dependence. Then the parameter θ can be estimated via maximum-likelihood (ML), which suggests the estimator $\hat{\lambda} = \lambda(\hat{\theta})$. Under the usual regularity conditions of ML-theory, as in Casella and Berger (2002, p. 516), the functional estimator $\hat{\lambda} = \lambda(\hat{\theta})$ represents an ML-estimator which possesses the well-known consistency and asymptotic normality properties.

Example 1. Suppose that (X, Y) is bivariate t-distributed, i.e.,

$$(X, Y) \stackrel{\mathrm{d}}{=} \mu + \frac{Z}{\sqrt{\chi^2_\alpha/\alpha}}\,, \qquad \alpha > 0,$$

where $Z \sim \mathcal{N}(0, \Sigma)$, $\mu \in \mathbb{R}^2$, $\Sigma \in \mathbb{R}^{2\times 2}$ positive definite, and Z is stochastically independent of χ^2_α. Then Embrechts et al. (2002) show that

$$\lambda = 2\,\bar{t}_{\alpha+1}\left(\sqrt{\alpha+1}\,\sqrt{\frac{1-\rho}{1+\rho}}\right), \tag{8}$$

where $\bar{t}_{\alpha+1}$ is the survival function of a Student's univariate t-distribution with $\alpha + 1$ degrees of freedom. The parameter $\rho = \sin(\pi\tau/2)$, expressed in terms of Kendall's tau, denotes the correlation parameter of (X, Y). It corresponds to Pearson's correlation coefficient, when it exists. □

Obviously this estimation approach requires prior information about the joint distribution function of the data. Consequently, the TDC estimator generates good estimates (in the sense of MSE) if the proposed distribution is the right one, but it will be biased if the distribution is wrong. In other words, this type of estimation is not expected to reveal surprising results and will be, therefore, excluded from the subsequent discussion.

3.2 Estimation within a class of distributions

Instead of a specific distribution, we now suppose that F belongs to a class of distributions. Because of its popularity in theory and practice, as illustrated, e.g., by Bingham et al. (2003) and Embrechts et al. (2003), we consider the class of elliptically contoured distribution, viz.

$$(X, Y) \stackrel{\mathrm{d}}{=} \mu + \mathcal{R}\Lambda U^{(2)},$$

where $U^{(2)}$ is a random pair uniformly distributed on the unit circle, $\mathcal{R}$ is a nonnegative random variable that is stochastically independent of $U^{(2)}$, $\mu \in I\!R^2$ is a location parameter, and $\Lambda \in I\!R^{2\times 2}$ is nonsingular. Well-known members of the latter distribution family are the multivariate normal, multivariate t and symmetric generalized hyperbolic distributions. Note that $\rho = 0$ does not correspond to independence; see, e.g., Abdous et al. (2005) for additional discussion concerning the dependence properties of this class of copulas.

In case the tail distribution of the Euclidean norm $\|(X, Y)\|_2$ is regularly varying with tail index $\alpha > 0$ [see Bingham et al. (1987) for the definition of regular variation], Schmidt (2002) and Frahm et al. (2003) show that tail dependence is present and that relationship (8) still holds. In particular, we have

$$A\left(\frac{1}{2}\right) = t_{\alpha+1}\left(\sqrt{\alpha+1}\,\sqrt{\frac{1-\rho}{1+\rho}}\right).$$

Various methods for the estimation of the tail index α are discussed, e.g., in Matthys and Beirlant (2002) or Embrechts et al. (1997).

3.3 Estimation using a specific copula

Suppose that the copula $C(\cdot\,;\theta)$ is known. Note that this is a much weaker assumption than assuming a specific distribution. The estimation of the parameter θ can be performed in two steps. First, we transform the observations of X and Y (or the corresponding block maxima) via estimates of the marginal distribution functions G and H and fit the copula from the transformed data in a second step; the transformation is justified by (3). Unless stated otherwise, the marginal distribution functions will be estimated by the empirical distribution functions $\widehat{G}_n$ and $\widehat{H}_n$.

The estimation of G and H via the empirical distribution functions avoids an incorrect specification of the margins. Genest et al. (1995), as well as Shih and Louis (1995), discuss consistency and asymptotic normality of the copula parameter θ if it is estimated in this fashion. Roughly speaking, if the map between θ and λ is smooth enough, then the estimator $\hat{\lambda} = \lambda(\hat{\theta})$ is consistent and asymptotically normal provided $\hat{\theta}$ is consistent and asymptotically normal.

If $G(\cdot\,;\theta_G)$ and $H(\cdot\,;\theta_H)$ are assumed to be specific distributions, then θ_G and θ_H can be estimated, e.g., via ML methods. In particular, the *IFM method* (method of inference functions for margins) consists of estimating θ_G and θ_H via ML and, in a second step, estimating the parameter θ of the copula $C(\cdot\,;\theta)$ via ML also. However, for this approach it is necessary that the parameters θ_G and θ_H do not analytically depend on the copula parameter θ. Results about the asymptotic distribution and the asymptotic covariance matrix of this type of estimation are derived in Joe (1997, Chapter 10); see also the references therein. A simulation study (which is not included in this paper but can be obtained from the authors upon request) shows that there is not much difference between the two step and the one step estimation in terms of the MSE. Also the MSE related to the pseudo-ML and the ML-estimation via empirical margins are roughly the same in this simulation study.

Example 2. Suppose that the data stem from a bivariate t-copula

$$C\,(u, v\,; \alpha, \rho) = t_\alpha \left\{t_\alpha^{-1}(u), t_\alpha^{-1}(v)\,; \rho\right\},$$

where $t_\alpha(\cdot\,; \rho)$ represents the bivariate t-distribution function with α degrees of freedom and correlation parameter ρ. □

Note that elliptical copulas (i.e., copulas of elliptical random vectors) are restricted to transpositional symmetry, i.e., $C = \tilde{C}$ and thus $\lambda_L = \lambda_U$. Hence, if the TDC is estimated from the entire sample via a single copula, the elliptical copulas are not appropriate if $\lambda_L \neq \lambda_U$. For example it is well known that investors react differently to negative and positive news. In particular for asset return modeling, the symmetry assumption has to be considered with care; see, e.g., Junker (2004) for an empirical study of commodity returns and U.S. dollar yield-returns using likelihood ratio tests. In such a case, λ_L and λ_U are better estimated by utilizing two different elliptical copulas and taking only the lower left or the upper right observations of the copula into consideration (see the example below).

Example 3. Suppose that C is a specific Archimedean copula such as the Gumbel copula

$$C_{GU}\,(u, v) = \exp\left[-\left\{(-\log u)^{\frac{1}{\theta}} + (-\log v)^{\frac{1}{\theta}}\right\}^{\theta}\right],$$

where $0 < \theta \leq 1$. It is easy to show that $\lambda_U(\theta) = 2 - 2^\theta$ and therefore $A(1/2) = 2^\theta/2$. Thus, λ_U may be estimated via $\lambda_U(\hat{\theta})$ where $\hat{\theta}$ is obtained from a fitted Gumbel copula. □

In general, Archimedean copulas are described by a continuous, strictly decreasing and convex generator function $\phi : [0,\, 1] \to [0,\, \infty]$ with $\phi(1) = 0$. The copula C is then given by

$$C(u, v) = \phi^{[-1]}\left\{\phi\,(u) + \phi\,(v)\right\}. \tag{9}$$

Here $\phi^{[-1]} : [0,\, \infty] \to [0,\, 1]$ denotes the pseudo-inverse of ϕ. The generator ϕ is called strict if $\phi(0) = \infty$ and in this case $\phi^{[-1]} = \phi^{-1}$; see Genest and MacKay (1986) or Nelsen (1999, Chapter 4).

Suppose (U, V) is distributed with Archimedean copula C with generator $\phi(\cdot\,; \theta)$ involving an unknown parameter θ. Recall that the TDC is defined along the copula's diagonal. In this context, we mention the following useful relationship

$$\mathrm{P}\{\max(U, V) \leq t\} = C\,(t, t) = \phi^{-1}\left\{2\,\phi\,(t\,; \theta)\,; \theta\right\}.$$

Example 4. Consider the following conditional distribution function:

$$\mathrm{P}\,(U \leq u, V \leq v \mid U, V \leq t) = \frac{C(u, v)}{C(t, t)}, \qquad 0 < t < 1,\ 0 \leq u, v \leq t.$$

Observe that we may only consider data which fall below the threshold t in order to estimate the lower TDC. The conditional distribution function of the upper right quadrant of C is similarly defined. The point is that it is useful to allow completely different conditional distributions for lower left and upper right observations of the copula. Note that the typical bias-variance trade-off appears again for the choice of the threshold t (as discussed in Section 1). □

3.4 Estimation within a class of copulas

Let us consider the important class of Archimedean copulas. Juri and Wüthrich (2002) have derived the following limiting result for the bivariate excess distribution of Archimedean copulas C. Define

$$F_t(x) = \frac{C\{\min(x,t), t\}}{C(t,t)}, \qquad 0 \le x \le 1,$$

where $0 < t < 1$ is a *low* threshold. Note that F_t can be also defined via the second argument of C since $C(u,v) = C(v,u)$. Now consider the "copula of small values" defined by

$$C_t(u,v) = \frac{C\left\{F_t^{-1}(u), F_t^{-1}(v)\right\}}{C(t,t)}, \tag{10}$$

where F_t^{-1} is the generalized inverse of F_t. It can be shown that if C has a differentiable and regularly varying generator ϕ with tail index $\alpha > 0$ then

$$\lim_{t \to 0^+} C_t(u,v) = C_{\mathrm{Cl}}(u,v\,;\alpha),$$

for every $0 \le u, v \le 1$, where

$$C_{\mathrm{Cl}}(u,v\,;\alpha) = \left(u^{-\alpha} + v^{-\alpha} - 1\right)^{-1/\alpha}$$

is the Clayton copula with parameter α. One may verify that $\lambda_L = \lambda_L(\alpha) = 2^{-1/\alpha}$. Thus, the lower TDC can be estimated by fitting the Clayton copula to small values of the approximate copula realizations and set $\hat{\lambda}_L = 2^{-1/\hat{\alpha}}$.

Remarks.

i) Archimedean copulas that belong to a domain of attraction are necessarily in the domain of attraction of the Gumbel copula, which is an EV copula; see, e.g., Genest and Rivest (1989) and Capéraà et al. (2000). Hence, the Gumbel copula seems to be a natural choice regarding the TDC estimation if we work in an Archimedean framework.

ii) The marginal distributions of financial asset returns are commonly easier to model than the corresponding dependence structure; this is often due to the limited availability of data. Consider for instance the pricing of so-called *basket credit derivatives*. Here the marginal survival functions of the underlying credits are usually estimated via parametric hazard-rate models by utilizing observable default spreads.

The choice of an appropriate dependence structure, however, is still a debate and several approaches are currently discussed; see, e.g., Li (1999) or Laurent and Gregory (2003).

3.5 Nonparametric estimation

In the present section, no parametric assumptions are made for the copula and the marginal distribution functions. TDC estimates are obtained from the empirical copula $\widehat{C}_n$. Note that the empirical copula implies the following relationship

$$\widehat{F}_n(x, y) = \widehat{C}_n\{\widehat{G}_n(x), \widehat{H}_n(y)\},$$

where $\widehat{F}_n$, $\widehat{G}_n$, and $\widehat{H}_n$ denote the empirical distributions.

In (7), we presented the nonparametric upper TDC estimator $\hat{\lambda}_U^{\mathrm{LOG}}$ which is based on the empirical copula. This estimator was motivated by equation (5). Note that if the bivariate data are stochastically independent (or comonotonic), $\hat{\lambda}_U^{\mathrm{LOG}}$ is well behaved for all thresholds k in terms of the bias, as in that case $C(t, t) = t^2$ (or $C(t, t) = t$) and thus $\hat{\lambda}_U^{\mathrm{LOG}} \approx 2 - 2 = 0$ (or $\hat{\lambda}_U^{\mathrm{LOG}} \approx 2 - 1 = 1$) holds independently of k.

Another estimator appears as a special case in Joe et al. (1992):

$$\hat{\lambda}_U^{\mathrm{SEC}} = 2 - \frac{1 - \widehat{C}_n\left(\frac{n-k}{n}, \frac{n-k}{n}\right)}{1 - \frac{n-k}{n}}, \qquad 0 < k \leq n. \tag{11}$$

This estimator can also be motivated by equation (5), which explains the superscript SEC illustrating the relationship to the secant of the copula's diagonal. Asymptotic normality and strong consistency of $\hat{\lambda}_U^{\mathrm{SEC}}$ are, e.g., addressed in Schmidt and Stadtmüller (2004).

A third nonparametric estimator is proposed below which is motivated in Capéraà et al. (1997). Let $\{(U_1, V_1), \ldots, (U_n, V_n)\}$ be a random sample obtained from the copula C. Assume that the empirical copula function approximates an EV copula C_{EV} (take block maxima if necessary) and define

$$\hat{\lambda}^{\mathrm{CFG}} = 2 - 2 \exp\left[\frac{1}{n}\sum_{i=1}^{n} \log\left\{\sqrt{\log\frac{1}{U_i}\log\frac{1}{V_i}} \Big/ \log\frac{1}{\max(U_i, V_i)^2}\right\}\right].$$

3.6 Pitfalls

From finitely many observations $(x_1, y_1), \ldots, (x_n, y_n)$ of (X, Y), it is difficult to conclude whether (X, Y) is tail dependent or not. As for tail-index estimation, one can always specify thin-tailed distributions which produce sample observations suggesting heavy tails even for large sample sizes. For example the upper left plot of Figure 2 shows the scatter plot of 2000 realizations from a distribution with standard normal univariate margins and copula C_{NM} corresponding to a mixture distribution of different bivariate Gaussian distributions, namely:

$$NM = \frac{7}{10}\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.49 & 0.245 \\ 0.245 & 0.49 \end{bmatrix}\right) + \frac{3}{10}\mathcal{N}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0.49 & 0.441 \\ 0.441 & 0.49 \end{bmatrix}\right).$$

At first glance, the scatter plot reveals upper tail dependence although any finite mixture of normal distributions is tail independent. The upper right plot of Figure 2 shows the scatter plot of 2000 realizations from a distribution with standard normal univariate margins and a Gumbel copula with $\theta = 2.25$. As expected, the sample reveals a large upper TDC of $\lambda_U = 2 - 2^{\frac{1}{\theta}} \approx 0.64$. Nevertheless, the upper left plot with $\lambda_U = 0$ looks more or less like the upper right plot. The lower two plots of Figure 2 give the corresponding TDC estimates of $\hat{\lambda}_U^{\text{LOG}}$ for different choices of k. It can be seen that for any choice of k the TDC estimate for copula C_{NM} has nearly the same value as the TDC estimate for the Gumbel copula. Conversely one may create samples which seem to be tail independent but they are realizations of a tail dependent distribution. Thus, the message is that one must be careful by inferring tail dependence from a finite random sample. The best way to protect against misidentifications is the application of several estimators, test or plots to the same data set.

We address another pitfall regarding the estimation of the marginal distribution functions. The use of parametric margins instead of empirical margins bears a model risk and may lead to wrong interpretations of the dependence structure. For instance, consider 3000 realizations of a random pair with distribution function

$$H(x, y) = C_{GU}\{t_\nu(x), t_\nu(y); \theta\},$$

where t_ν denotes the univariate standard t-distribution with ν degrees of freedom and C_{GU} is the Gumbel copula with parameter θ. Set $\theta = 2$ and $\nu = 3$. In Figure 3, we compare the empirical copula densities which are either obtained via empirical marginal distributions or via fitted normal marginal distributions. Precisely, in the second case we plot the pairs $(G(x_i), H(y_i))$, where G and H are normal distribution functions with parameters estimated from the data. The left panel of Figure 3 clearly illustrates the dependence structure of a Gumbel copula. By contrast, the data in the right panel have nearly lost all the appearance for upper tail dependence. Thus we have shown that not testing or ignoring the quality of the marginal fit can cause dramatic misinterpretation of the underlying dependence structure.

4 Simulation Study

In order to compare the finite sample properties of the discussed TDC estimators, we run an extensive simulation study. Each simulated data set consists of 1000 independent copies of n realizations from a random sample $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ having one particular distribution out of four. Three different sample sizes $n = 250, 1000, 5000$ are considered for each data set. The four different distributions are denoted by H, T, G and AG. For example the data set S_H^{250} contains realizations of 1000 samples with sample size 250 which are generated from a bivariate symmetric generalized hyperbolic distribution H. This is an elliptical distribution (see (3.2)), where $\mathcal{R}U^{(2)}$ has density

$$f(s) = \frac{K_0\left(\sqrt{1+s's}\right)}{2\pi \cdot K_0(1)}$$

and K_0 is the Bessel function of the third kind with index 0. The correlation parameter is set to $\rho = 0.5$.

Further, T refers to the bivariate t-distribution with $\nu = 1.5$ degrees of freedom and $\rho = 0.5$. Distribution G is determined by the distribution function

$$F_G(x, y) = C_G \left\{\Phi\left(x\right), \Phi\left(y\right); \vartheta, \delta\right\},$$

where Φ denotes the univariate standard normal distribution and C_G is an Archimedean copula with generator function

$$\phi_G(t\,; \vartheta, \delta) = \{\phi_{\text{Frank}}(t\,; \vartheta)\}^\delta = \left(-\log\frac{e^{-\vartheta t}-1}{e^{-\vartheta}-1}\right)^\delta, \qquad \vartheta \neq 0, \delta \geq 1,$$

considered by Junker and May (2002). Here, ϕ_{Frank} is the generator of the Frank copula, and values $\vartheta = -0.76$ and $\delta = 1.56$ are chosen, for reasons given below.

Finally, distribution AG is an asymmetric Gumbel copula, as defined by Tawn (1988), combined with standard normal margins, viz.

$$F_{AG}(x, y) = C_{AG} \left\{\Phi\left(x\right), \Phi\left(y\right); \theta, \phi, \delta\right\},$$

where

$$C_{AG}\left(u, v; \theta, \phi, \delta\right) = u^{1-\theta} v^{1-\phi} \exp\left(-\left[\{-\theta \ln\left(u\right)\}^\delta + \{-\phi \ln\left(v\right)\}^\delta\right]^{\frac{1}{\delta}}\right).$$

We set $\theta = 0.5, \phi = 0.9, \delta = 2.78$. For additional ways of generating asymmetric models and multi-parameter Archimedean copulas, see Genest et al. (1998).

Note that distribution H has no tail dependence; e.g., see Schmidt (2002). Thus the set of generalized hyperbolic data is used to control the performance of the TDC estimation methods under the absence of tail dependence. By contrast, distribution T possesses tail dependence; see also (8). Further, copula C_G is lower tail independent but upper tail dependent, i.e., $\lambda_L^G = 0$ and $\lambda_U^G = 2 - 2^{1/\delta}$. The parametrization of the distributions T, G, and AG is chosen such that $\lambda_L^T = \lambda_U^T = \lambda_U^G = \lambda_U^{AG} = 0.4406$ and Kendall's tau $\tau^H = \tau^T = \tau^G = \tau^{AG} = 1/3$ in order to provide comparability of the estimation results. Figure 4 illustrates the different tail behavior of distributions H, T, G, and AG by presenting the scatter plots of the respective simulated data-sample with sample size $n = 5000$, together with the corresponding empirical copula realizations. Regarding the copula mapping, we use empirical marginal distribution functions.

The different estimation methods are compared via the sample means $\widehat{\mu}\,(\hat{\lambda}_n)$ and the sample standard deviations $\widehat{\sigma}\,(\hat{\lambda}_n)$ of the estimates $\hat{\lambda}_{n,i}$, $i = 1, \ldots, 1000$, depending on the sample size n. Furthermore, to analyze the bias-variance trade-off for different sample sizes and estimation methods we compare the corresponding root mean squared errors:

$$\mathrm{RMSE}\left(\hat{\lambda}_n\right) = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} \left(\hat{\lambda}_{n,i} - \lambda\right)^2}. \tag{12}$$

Moreover, we introduce another statistical quantity called MESE (mean error to standard error):

$$\mathrm{MESE}\left(\hat{\lambda}_n\right) = \frac{\left|\widehat{\mu}\,(\hat{\lambda}_n) - \lambda\right|}{\widehat{\sigma}\,(\hat{\lambda}_n)}. \tag{13}$$

MESE quantifies the sample bias normalized by the sample standard error. Thus, it measures the degree of possible misinterpretation caused by considering the standard error as a criterion for the quality of the estimator. For instance, assume a situation where the standard error of the estimate is small but the bias is large. In that case the true parameter is far away from the estimate, though the approximated confidence bands suggest the opposite. This situation is represented by a large MESE. In particular, if the bias of the estimator decreases with a slower rate as the standard error (for $n \to \infty$) then MESE tends to infinity. One aim of this quantity is to investigate the danger of this sort of misinterpretation.

In the following, the TDC is estimated via the various methods discussed in Section 3. It is reasonable to discard those models which are obviously not compatible with the observed data. Further, we do not consider TDC estimations using a specific distribution since the results are not surprising (due to the strong distributional assumptions).

4.1 Estimation within a class of distributions

The following estimation approach is based on the expositions in Section 3.2. We have to estimate the tail index α and the correlation parameter ρ. For any elliptical distribution, the correlation parameter is determined by Kendall's τ via the relationship of Lindskog et al. (2003), viz.

$$\rho = \sin\left(\frac{\pi}{2}\,\tau\right).$$

Hence, using $\hat{\tau} = (c-d)/(c+d)$, where c is the number of concordant pairs and d is the number of discordant pairs of the sample, the correlation parameter may be estimated by $\hat{\rho} = \sin(\pi\hat{\tau}/2)$. Alternatively, one can apply Tyler's M-estimator for the covariance matrix, which is completely robust within the class of elliptical distributions; see Tyler (1987) or Frahm (2004). Given the covariance matrix, the random variable $\mathcal{R}$ can be extracted by the Mahalanobis norm of (X, Y). Our pre-simulations showed that there is no essential difference regarding the finite sample properties between these two estimation procedures. Hence we use the approach via Kendall's τ for the sake of simplicity.

The tail index α could be estimated via traditional methods of EVT, i.e., by taking only extreme values or excesses of the Euclidean norm $\|(X,Y)\|_2$ into consideration. Different methods for estimating the tail index are discussed, e.g., in Matthys and Beirlant (2002) or Embrechts et al. (1997). For our purposes, we used a Hill-type estimator with optimal sample fraction proposed by Drees and Kaufmann (1998).

For the data sets S_G and S_{AG}, we do not assume an elliptical distribution due to the obvious asymmetry of the data; see the scatter plots in Figure 4. Consequently we will not apply the latter estimation procedure to these data sets. The estimation results for S_H and S_T are summarized in Table 1.

4.2 Estimation using a specific copula

As mentioned in Section 3.3, the marginal distribution functions are now estimated by their empirical counterparts, whereas the copula is chosen according to our decision. For the elliptical data sets S_H and S_T, we opted for a t-copula, which seems to be a realistic choice by glancing at the scatter plots in Figure 4. However, we know that the t-copula is not suitable for S_H. The TDC is estimated via relation (8). Regarding the data set S_G, we fit a Gumbel copula since the empirical copula, which is illustrated in Figure 4, shows transpositional asymmetry, i.e., the underlying copula does not seem to coincide with its survival copula. Moreover, the symmetry of the Gumbel copula with respect to the copula's diagonal appears to be satisfied by S_G, too. Here the upper TDC is estimated via $\hat{\lambda}_U^G = 2 - 2^{1/\widehat{\theta}}$. However, the original copula of S_G is not the Gumbel copula and thus the assumed model is wrong. We disregard the data set S_{AG} since it is not obvious which specific copula might be appropriate in this framework. Note that the empirical copula is even asymmetric with respect to the copula's diagonal (see Figure 4). The estimation results are summarized in Table 1.

4.3 Estimation within a class of copulas

We follow the approach given in Section 3.4, which is based on a result by Juri and Wüthrich (2002). The upper TDC is estimated, but in the following we refer to the lower left corner of the underlying survival copula. We choose a small threshold t for the latter copula in order to obtain the conditional copula (10). In order to increase the robustness of the copula estimates with respect to the threshold choice, we take the mean of estimates which correspond to 10 equidistant thresholds between $n^{-1/2}$ and $n^{-1/4}$. Note that if the margins of the underlying distribution are completely dependent, then $n^{1/2}$ data points are expected in the copula's lower left quadrant which is determined by the threshold $t = n^{-1/2}$. For the smaller threshold $t = n^{-1/4}$, the same amount of data ($n^{1/2}$) is expected for an independence copula.

We assume that the data S_H, S_T and S_G have an Archimedean dependence structure. Since S_{AG} is not permutational symmetric, we reject the Archimedean hypothesis for this data set. We point out that the Frank copula, Frank (1979), is the only transpositional symmetric Archimedean copula and thus suitable for S_H and S_T but it does not comprise tail dependence. The statistical results for the data sets S_H, S_T, and S_G are provided in Table 1.

4.4 Nonparametric estimation

No specific distributional assumptions for the upper TDC estimation are made in the present section. Recall that for $\hat{\lambda}_U^{\mathrm{LOG}}$ and $\hat{\lambda}_U^{\mathrm{SEC}}$, we have to choose the threshold k as indicated in (7) and (11). By contrast, $\hat{\lambda}_U^{\mathrm{CFG}}$ needs no additional decision regarding the threshold. This, however, goes along with the assumption that the underlying copula can be approximated by an EV copula.

The diagonal section of the copula is supposed to be smooth in the neighborhood of 1, and the second derivative of the diagonal section is expected to be small (i.e., the first derivative is approximately constant). Then $\hat{\lambda}_U^{\mathrm{SEC}}(k)$ is homogeneous for small (thresholds) k. However, k should be sufficiently large in order to decrease variance. We consider the graph $k \mapsto \hat{\lambda}_U^{\mathrm{SEC}}(k)$ in order to identify the plateau which is induced by the homogeneity. Note that $\hat{\lambda}_U^{\mathrm{LOG}}$ possesses this homogeneity property even for larger thresholds k if the diagonal section of the copula follows a power law.

The plateau is chosen according to the following heuristic plateau-finding algorithm. First, the map $k \mapsto \hat{\lambda}_k$ is smoothed by a simple box kernel with bandwidth $b \in \mathbb{N}$. That is, the means of $2b+1$ successive points of $\hat{\lambda}_1, \ldots, \hat{\lambda}_n$ lead to the new smoothed map $\overline{\lambda}_1, \ldots, \overline{\lambda}_{n-2b}$. Here we have taken $b = \lfloor 0.005n \rfloor$ such that each moving average consists of 1% of the data, approximately. In a second step, a plateau of length $m = \lfloor \sqrt{n-2b} \rfloor$ is defined as a vector $p_k = \left(\overline{\lambda}_k, \ldots, \overline{\lambda}_{k+m-1}\right)$, $k = 1, \ldots, n-2b-m+1$. The algorithm stops at the first plateau p_k which elements fulfill the condition

$$\sum_{i=k+1}^{k+m-1} \left|\overline{\lambda}_i - \overline{\lambda}_k\right| \leq 2\sigma,$$

where σ represents the standard deviation of $\overline{\lambda}_1, \ldots, \overline{\lambda}_{n-2b}$. Then the TDC estimate is set to

$$\hat{\lambda}_U(k) = \frac{1}{m} \sum_{i=1}^{m} \overline{\lambda}_{k+i-1}.$$

If there is no plateau fulfilling the stopping condition, the TDC estimate is set to zero.

As outlined above, we may choose a greater bandwidth b for the $\hat{\lambda}^{\mathrm{LOG}}$ in order to reduce the variance of the estimation. However, for a better comparison we do not change b. The statistical results related to these nonparametric estimators for the data sets S_H, S_T,S_G and S_{AG} are provided in Table 2.

Method	Data set	$\hat{\mu}(\hat{\lambda}_U)$	BIAS$(\hat{\lambda}_U)$	$\hat{\sigma}(\hat{\lambda}_U)$	RMSE$(\hat{\lambda}_U)$	MESE$(\hat{\lambda}_U)$
Estimation	S_H^{250}	0.1618	0.1618^+	0.0817	0.1812^+	1.9802
for a	S_H^{1000}	0.1698	0.1698	0.0413	0.1747^+	4.1116
specific copula	S_H^{5000}	0.1739	0.1739	0.0187	0.1749	9.3141
(t- and	S_T^{250}	0.4281	-0.0125	0.0403^+	0.0422^+	0.3078
Gumbel	S_T^{1000}	0.4374	-0.0032	0.0204^+	0.0206^+	0.1403
copula)	S_T^{5000}	0.4400	-0.0006^+	0.0092^+	0.0092^+	0.0652^+
	S_G^{250}	0.3905	-0.0501^-	0.0437^+	0.0664	1.1466^-
	S_G^{1000}	0.3922	-0.0484^-	0.0212^+	0.0529	2.2819^-
	S_G^{5000}	0.3919	-0.0487^-	0.0097^+	0.0497	5.0252^-
Estimation	S_H^{250}	0.2031	0.2031	0.0588	0.2114	3.4541
within a class	S_H^{1000}	0.1815	0.1815	0.0377	0.1854	4.8143
of distributions	S_H^{5000}	0.1575	0.1575	0.0220	0.1590	7.1591
(elliptical	S_T^{250}	0.4379	-0.0027^+	0.0465	0.0466	0.0490^+
distributions)	S_T^{1000}	0.4432	0.0026^+	0.0242	0.0243	0.1041^+
	S_T^{5000}	0.4437	0.0031	0.0109	0.0113	0.2849
Estimation	S_H^{250}	0.2278	0.2278	0.1910^-	0.2972	1.1921^+
within a class	S_H^{1000}	0.1671	0.1671^+	0.1357^-	0.2152	1.2309^+
of copulas	S_H^{5000}	0.1237	0.1237^+	0.0977^-	0.1576^+	1.2657^+
(Archimedean	S_T^{250}	0.5317	0.0911	0.1864^-	0.2074^-	0.4880
copulas)	S_T^{1000}	0.5575	0.1169^-	0.1175^-	0.1657^-	0.9943^-
	S_T^{5000}	0.5701	0.1295^-	0.0647^-	0.1448^-	2.0022^-
	S_G^{250}	0.4352	-0.0054^+	0.1948^-	0.1948^-	0.0277^+
	S_G^{1000}	0.4495	0.0089^+	0.1312^-	0.1314^-	0.0548^+
	S_G^{5000}	0.4554	0.0148	0.0792^-	0.0805^-	0.1818^+

Table 1: Various statistical results for the TDC estimation under a specific copula assumption or within a class of distributions or copulas. Best values of the different methods (including the nonparametric methods in Table 2) are ticked with a plus, worst values are ticked with a minus.

Method	Data set	$\hat{\mu}(\hat{\lambda}_U)$	BIAS($\hat{\lambda}_U$)	$\hat{\sigma}(\hat{\lambda}_U)$	RMSE($\hat{\lambda}_U$)	MESE($\hat{\lambda}_U$)
Nonparametric	S_H^{250}	0.3553	0.3553	0.0444^+	0.3580	8.0008^-
estimator	S_H^{1000}	0.3558	0.3558^-	0.0229^+	0.3566^-	15.5400^-
$\hat{\lambda}_U^{\text{CFG}}$	S_H^{5000}	0.3568	0.3568^-	0.0104^+	0.3570^-	34.3123^-
	S_T^{250}	0.4462	0.0056	0.0471	0.0474	0.1133
	S_T^{1000}	0.4509	0.0103	0.0234	0.0256	0.4437
	S_T^{5000}	0.4511	0.0105	0.0107	0.0150	0.9825
	S_G^{250}	0.3922	-0.0484	0.0450	0.0661^+	1.0759
	S_G^{1000}	0.3939	-0.0467	0.0216	0.0514^+	2.1593
	S_G^{5000}	0.3936	-0.0470	0.0099	0.0480	4.7443
	S_{AG}^{250}	0.4377	-0.0029	0.0480^+	0.0481^+	0.0648^+
	S_{AG}^{1000}	0.4400	-0.0006	0.0243^+	0.0243^+	0.0247^+
	S_{AG}^{5000}	0.4406	0.0000^+	0.0107^+	0.0107^+	0.0000^+
Nonparametric	S_H^{250}	0.3636	0.3636^-	0.1016	0.3775^-	3.5787^-
estimator	S_H^{1000}	0.3056	0.3056	0.0717	0.3139	4.2622
$\hat{\lambda}_U^{\text{SEC}}$	S_H^{5000}	0.2390	0.2390	0.0932	0.2565^-	2.5644
	S_T^{250}	0.4681	0.0275	0.0800	0.0845	0.3436
	S_T^{1000}	0.4587	0.0181	0.0513	0.0545	0.3534
	S_T^{5000}	0.4463	0.0057	0.0431	0.0435	0.1322
	S_G^{250}	0.4841	0.0435	0.0796	0.0907	0.5467
	S_G^{1000}	0.4650	0.0244	0.0482	0.0541	0.5062
	S_G^{5000}	0.4453	0.0047^+	0.0603	0.0605	0.0775^+
	S_{AG}^{250}	0.5042	0.0636^-	0.0810^-	0.1029^-	0.7835^-
	S_{AG}^{1000}	0.4763	0.0357^-	0.0523^-	0.0633^-	0.6818^-
	S_{AG}^{5000}	0.4567	0.0161^-	0.0340^-	0.0376^-	0.4722^-
Nonparametric	S_H^{250}	0.3144	0.3144	0.0828	0.3251	3.7968
estimator	S_H^{1000}	0.2893	0.2893	0.0539	0.2943	5.3677
$\hat{\lambda}_U^{\text{LOG}}$	S_H^{5000}	0.2567	0.2567	0.0377	0.2595	6.8103
	S_T^{250}	0.3951	-0.0455^-	0.0727	0.0857	0.6242^-
	S_T^{1000}	0.4132	-0.0274	0.0491	0.0562	0.5569
	S_T^{5000}	0.4240	-0.0166	0.0352	0.0389	0.4704
	S_G^{250}	0.4016	-0.0390^+	0.0719	0.0818	0.5426
	S_G^{1000}	0.4098	-0.0308	0.0448	0.0544	0.6850
	S_G^{5000}	0.4229	-0.0177	0.0233	0.0293	0.7624
	S_{AG}^{250}	0.4424	0.0018^+	0.0696	0.0696	0.0259
	S_{AG}^{1000}	0.4403	-0.0003^+	0.0386	0.0386	0.0078
	S_{AG}^{5000}	0.4412	0.0006	0.0200	0.0200	0.0030

Table 2: Statistical results for the methods of nonparametric TDC estimation. Best values of the different methods (including the parametric methods in Table 1) are ticked with a plus, worst are ticked with a minus.

4.5 Discussion of the simulation results

We discuss the simulation results with regard to the following statistical measures: sample variance, sample bias, RMSE, and MESE.

Sample variance. The TDC estimations within the class of elliptically contoured distributions and for specific copulas show the lowest sample variances among all considered methods of TDC estimation. Of course, the small variances go along with restrictive model assumptions. Nevertheless the estimation within the class of elliptically contoured distributions has a surprisingly low variance, even though the tail index α is estimated from few extremal data. Further, a comparably small variance is obtained for the estimator $\hat{\lambda}^{\text{CFG}}$ which is based on the weaker assumption of an EV copula. By contrast, the sample variances of $\hat{\lambda}^{\text{SEC}}$ and $\hat{\lambda}^{\text{LOG}}$ are much larger. In particular the TDC estimation within the class of Archimedean copulas, as described in Section 3.4, shows an exceptionally large variance. However, note that the latter three estimation methods utilize only sub-samples of extremal (excess) data. Besides, there is another explanation for the large variance of the last estimation method: Here, the TDC is estimated (in a second step) from a copula which is fitted from extremal (excess) data.

We conclude that an effective variance reduction of the TDC estimation is possible for those estimation methods which use the entire data sample.

Sample bias. It is not surprising that the estimation methods with distributional assumption have a quite low sample bias if the underlying distribution is true. See, for example, the bias related to S_T for TDC estimations within the class of elliptical distributions or for specific copulas; see also S_G for the estimation within the class of Archimedean copulas. By contrast, the estimation with regard to the sample bias performs badly if we assume an inappropriate distribution, as can be seen for the data set S_G under the estimation using a specific copula; see Section 3.3 and the data set S_T under the estimation within the class of Archimedean copulas. It is, however, surprising that the TDC estimation from S_H (recall that H is an elliptical distribution) shows a larger sample bias for the estimation within the class of elliptical distributions than for the estimation with a specific copula or within the class of Archimedean copulas. We point out that the largest sample bias is observed for the nonparametric estimation methods. Further, all estimation methods, in particular $\hat{\lambda}^{\text{CFG}}$, yield a large MESE value (which indicates a large sample bias relative to the sample variance) for the data set S_H which exhibits tail independence. In most cases, the MESE is greater than 2, which means that the true TDC value is not included in the 2σ confidence interval. Moreover, this illustrates that the sole consideration of the sample variance may lead to the fallacy of an exceptionally large TDC, even in the case of tail independence. Thus it is absolutely necessary to test for tail dependence in the first instance; see Ledford and Tawn (1996), Draisma et al. (2004) or Falk and Michel (2004).

RMSE. The TDC estimation using a specific copula represents the smallest RMSE if the underling copula is true, as applies, e.g., to the data set S_T. The second best RMSE for the latter data set comes from the estimation within the class of elliptically contoured distributions. This estimation goes along with a larger variance, due to the estimation of the tail index α. It is remarkable that the nonparametric estimator $\hat{\lambda}^{\text{CFG}}$ (which assumes an EV copula) possesses an RMSE in the same range as the two aforementioned estimation methods for the data sets S_T, S_G, and S_{AG}. By contrast, the estimators $\hat{\lambda}^{\text{SEC}}$ and $\hat{\lambda}^{\text{LOG}}$ have a much larger RMSE. The TDC estimation based on the class of Archimedean copulas as described in Section 3.4 yields by far the largest RMSE even for the (Archimedean) data set S_G. Further, the estimation using a specific copula has a similar RMSE under the wrong model assumption (see S_G) due to its low variance. An RMSE in the same range is found for the nonparametric estimators $\hat{\lambda}^{\text{SEC}}$ and $\hat{\lambda}^{\text{LOG}}$. This also suggest a consideration of the following statistical measure.

MESE. The MESE detects all misspecified models such as S_G under the estimation using a specific (Gumbel) copula, or S_T under the estimation within the class of Archimedean copulas. However, if the underlying model is true, then the MESE is quite low (e.g., estimation within a class of copulas or distributions for the data set S_T). In this case and for all nonparametric estimations, the MESE is usually smaller than 1, which indicates that the true TDC lies within the 1σ confidence band. Only the data set S_H represents an exception. Especially the estimator $\hat{\lambda}^{\text{CFG}}$ shows an exceptionally bad performance, which is due to its small variance. Thus, the sample variance has to be considered with caution for the latter estimator.

There exists an interesting aspect regarding the estimator $\hat{\lambda}^{\text{SEC}}$. Due to its geometric interpretation as the slope of the secant along the copula's diagonal (at the point $(1, 1)$), the latter estimator reacts sensitively if the extremal data are not accumulated along the diagonal. Such is the case, e.g., for the data set S_{AG} and might also explain the bad performance of $\hat{\lambda}^{\text{SEC}}$ regarding the latter data set.

5 Conclusions

On the basis of the results of our simulation study, we have ranked the various TDC estimators according to their finite-sample performance. Table 3 illustrates the corresponding rankings in terms of numbers between 1 (very good performance) and 6 (very poor performance). Thereby we distinguish between a true and a wrong assumption of the underlying distribution. Moreover, we rank the estimators according to their performance under the assumption of tail independence. The second column of Table 3 indicates the heaviness of the model assumptions required for each TDC estimator.

Clearly the (semi-)parametric TDC estimators perform well if the underlying distribution/copula is the right one (except the TDC estimator within a class of copulas as described in Section 3.4). However, their performance is very poor if the assumed model is wrong. Thus, we definitely recommend to test any distributional assumptions. For instance, in the case of empirical marginal distributions and specific copula, we suggest to test the goodness-of-fit of the copula via (non-)parametric procedures such as those developed in Fermanian (2003), Chen et al. (2003), Dobrić and Schmid (2004) or Genest et al. (2006). Further, if one makes use of an elliptically contoured distribution, then we suggest to test for ellipticity; see, e.g., Manzotti et al. (2002). However, we do not recommend a TDC estimation as presented in Section 3.4, since we do not know a suitable test

TDC estimation methods	degree of assumptions	perform. under true assumpt.	perform. under wrong assumpt.	perform. under TDC = 0
specific copula	strong	1	6	1-2
distr. class	strong	2–3	–$^{\star)}$	3
copula class	medium	4–5	5	1–2
$\hat{\lambda}^{\mathrm{CFG}}$	weak	2–3	2–3	6
$\hat{\lambda}^{\mathrm{SEC}}$	weak	4	4	5
$\hat{\lambda}^{\mathrm{LOG}}$	weak	3	3	5

Table 3: Overview of the performance of the TDC estimation methods. Grades rank from 1 to 6 with 1 excellent and 6 poor. $^{\star)}$ This TDC estimation method (via elliptical distributions) is disregarded due to the obvious asymmetry of the data arising from the wrong distributional assumption.

for Archimedean copulas and the estimator does not perform well if the assumption of an Archimedean copula is wrong. Goodness-of-fit tests within the family of Archimedean copulas are developed, e.g., in Wang and Wells (2000) or Genest et al. (2006).

Among the nonparametric estimators, the TDC estimator $\hat{\lambda}^{\mathrm{CFG}}$ does well, although we advise caution regarding the sometimes low variance relative to bias. Further, $\hat{\lambda}^{\mathrm{CFG}}$ shows a weak performance in the case of tail independence. This estimator is followed by $\hat{\lambda}^{\mathrm{LOG}}$ and $\hat{\lambda}^{\mathrm{SEC}}$ whereas the last estimator is not robust for non-transpositional symmetric data. Further, the variance of $\hat{\lambda}^{\mathrm{LOG}}$ could be possibly reduced by enlarging the estimation kernel (see Section 4.4).

We conclude that, among the nonparametric TDC estimators, $\hat{\lambda}^{\mathrm{CFG}}$ shows the best performance whereas for (semi-)parametric estimations we recommend a specific copula (such as the *t*-copula). For the latter, we suggest to work with empirical marginal distributions. Further, we point out that the decision for a specific distribution or class of distributions should be influenced by the visual appearance of the data, e.g., via the related scatter plots. Unfortunately, if the number of data is small (such as 250 points), it is difficult to draw sensible conclusions from the scatter plot. Moreover, the nonparametric estimators are too sensitive in case of small sample sizes. Thus, under these circumstances, a parametric TDC estimation might be favorable in order to increase the stability of the estimation although the model error could be large.

The previous simulation is based on a limited number of distributions, although we tried to incorporate a large spectrum of possible distributions. Nevertheless, according to the pitfalls in Section 3.6 and the statistical results for the tail independent data set H, we see that tests for tail dependence are absolutely mandatory for every TDC estimation. Unfortunately, the current literature on this kind of test is only limited; see Coles, Heffernan, and Tawn (1999), Draisma et al. (2004), or Falk and Michel (2004).

Acknowledgement

The authors are grateful to the Editor and two referees for their constructive comments which substantially improved an earlier version of the paper.

References

Abdous, B., Ghoudi, K., Khoudraji, A., 1998. Non-parametric estimation of the limit dependence function of multivariate extremes. Extremes 2, 245–268.

Abdous, B., Genest, C., Rémillard, B., 2005. Dependence properties of meta-elliptical distributions. In: Statistical Modelling and Analysis for Complex Data Problems, P. Duchesne and B. Rémillard, Eds. Springer, New York, pp. 1–15.

Ané, T., Kharoubi, C., 2003. Dependence structure and risk measure. Journal of Business 76, 411–438.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular variation. In: Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge University Press, Cambridge, England.

Bingham, N.H., Kiesel, R., Schmidt, R., 2003. Semi-parametric models in finance: econometric applications. Quantitative Finance 3, 426–441.

Capéraà, P., Fougères, A.-L., 2000. Estimation of a bivariate extreme value distribution. Extremes 3, 311–329.

Capéraà, P., Fougères, A.-L., Genest, C., 1997. A nonparametric estimation procedure for bivariate extreme value copulas. Biometrika 84, 567–577.

Capéraà, P., Fougères, A.-L., Genest, C., 2000. Bivariate distributions with given extreme value attractor. Journal of Multivariate Analysis 72, 30–49.

Casella, G., Berger, R.L., 2002. Statistical Inference, Second Edition. Duxbury Press, Pacific Grove, CA.

Chen, X., Fan, Y., Patton, A., 2003. Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates. FMG/IAM, Discussion Paper 48, `http://www.greta.it/credit/credit2003/Poster/Chen Fan Patton.pdf`.

Coles, S.G., Heffernan, J.E., Tawn, J.A., 1999. Dependence measures for extreme value analyses. Extremes 2, 339–365.

Coles, S.G., Tawn, J.A., 1991. Modelling multivariate extreme events. Journal of the Royal Statistical Society, Series B 53, 377–392.

Deheuvels, P., 1991. On the limiting behaviour of the Pickands estimator for bivariate extreme value distributions. Statistics and Probability Letters 12, 429–439.

Dobrić, J., Schmid, F., 2004. Test of fit for parametric families of copulas: applications to financial data. Discussion Paper 03–04, Department of Economic and Social Statistics, University of Cologne, Germany.

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

Drees, H., Kaufmann, E., 1998. Selecting the optimal sample fraction in univariate extreme value estimation. Stochastic Processes and Their Applications 75, 149–172.

Einmahl, J., de Haan, L., Huang, X., 1993. Estimating a multidimensional extreme value distribution. Journal of Multivariate Analysis 47, 35–47.

Einmahl, J., de Haan, L., Sinha, A.K., 1997. Estimating the spectral measure of an extreme value distribution. Stochastic Processes and Their Applications 70, 143–171.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

Embrechts, P., Lindskog, F. McNeil, A., 2003. Modelling dependence with copulas and applications to risk management. In: Handbook of Heavy Tailed Distributions in Finance, S. Rachev, Ed. Elsevier, pp. 329–384.

Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependency in risk management: properties and pitfalls. In: Risk Management: Value at Risk and Beyond, M.A.H. Dempster, Ed. Cambridge University Press, pp. 176–223.

Falk, M., Michel, R., 2004. Testing for tail independence in extreme value models. Talk at DeMoSTAFI Québec 2004, `http://statistik.mathematik.uni-wuerzburg.de/~falk/archiv/vextra 03.pdf`.

Falk, M., Reiss, R.D., 2003. Efficient estimation of the canonical dependence function. Extremes 6, 61–82.

Fermanian, J.-D., 2003. Goodness-of-fit tests for copulas. Journal of Multivariate Analysis, in press.

Fisher, R.A., Tippett, L.H.C., 1928. Limiting forms of the frequency distributions of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society 24, 180–190.

Frahm, G., 2004. Generalized elliptical distributions: Theory and applications. Dissertation, Department of Statistics, Faculty of Management, Economics, and Social Sciences, University of Cologne, Germany. `http://kups.ub.uni-koeln.de/volltexte/2004/1319/`.

Frahm, G., Junker, M., Szimayer, A., 2003. Elliptical copulas: applicability and limitations. Statistics and Probability Letters 63, 275–286.

Frank, M.J., 1979. On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. Aequationes Mathematicae 19, 194–226.

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

Genest, C., Ghoudi, K., Rivest, L.-P., 1998. Comment on the article of E.W. Frees and E.A. Valdez intitled "Understanding relationships using copulas." North American Actuarial Journal 2, 143–149.

Genest, C., MacKay, R.J., 1986. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. The Canadian Journal of Statistics 14, 145–159.

Genest, C., Quessy, J.-F., Rémillard, B., 2006. Goodness-of-fit procedures for copula models based on the probability integral transformation. Scandinavian Journal of Statistics, in press.

Genest, C., Rivest, L.-P., 1989. A characterization of Gumbel's family of extreme value distributions. Statistics and Probability Letters 8, 207–211.

de Haan, L., Resnick, S.I., 1993. Estimating the limit distribution of multivariate extremes. Communications in Statistics, Stochastic Models 9, 275–309.

Hutchinson, T.P., Lai, C.D., 1990. Continuous Bivariate Distributions, Emphasising Applications. Rumsby Scientific Publishing, Adelaide, Australia.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Joe, H., Smith, R.L., Weissman, I., 1992. Bivariate threshold models for extremes. Journal of the Royal Statistical Society, Series B 54, 171–183.

Junker, M., 2004. Modelling, estimating and validating multidimensional distribution functions: with applications to risk management. Dissertation, Technical University of Kaiserslautern, Germany.

Junker, M., May, A., 2002. Measurement of aggregate risk with copulas. Report, Research Center caesar, Germany.

Juri, A., Wüthrich, M.V., 2002. Copula convergence theorems for tail events. Insurance: Mathematics and Economics 24, 139–148.

Laurent, J.-P., Gregory, J., 2003. Basket default swaps, CDOs and factor copulas. Journal of Risk, in press.

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

Li, D.X., 1999. The valuation of basket credit derivatives. CreditMetrics Monitor, April, 34–50.

Lindskog, F., McNeil, A., Schmock, U., 2003. Kendall's tau for elliptical distributions. In: Credit Risk: Measurement, Evaluation and Management, G. Bohl, G. Nakhaeizadeh, S.T. Rachev, T. Ridder, K.H. Vollmer, Eds. Physica-Verlag, Heidelberg, pp. 649–676.

Malevergne, Y., Sornette, D., 2004. Investigating extreme dependencies. The Review of Financial Studies, in press.

Manzotti, A., Perez, F.J., Quiroz, A.J., 2002. A statistic for testing the null hypothesis of elliptical symmetry. Journal of Multivariate Analysis 81, 274–285.

Matthys, G., Beirlant, J., 2002. Adaptive threshold selection in tail index estimation. In: Extremes and Integrated Risk Management, P. Embrechts, Ed. Risk Books London, pp. 37–49.

Nelsen, R.B., 1999. An Introduction to Copulas. Springer, New York.

Pickands, J., 1981. Multivariate extreme value distributions. Bulletin of the International Statistical Institute, pp. 859–878.

Schmidt, R., 2002. Tail dependence for elliptically contoured distributions. Mathematical Methods of Operations Research 55, 301–327.

Schmidt, R., Stadtmüller, U., 2004. Nonparametric estimation of tail dependence. University of Cologne, Department of Economic and Social Statistics, `http://www.uni-koeln.de/wiso-fak/wisostatsem/autoren/schmidt`.

Shih, J.H., Louis, T.A., 1995. Inferences on the association parameter in copula models for bivariate survival data. Biometrics 51, 1384–1399.

Sibuya, M., 1960. Bivariate extreme statistics. Annals of the Institute of Statistical Mathematics 11, 195–210.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de statistique de l'Université de Paris 8, 229–231.

Smith, R.L., Tawn, J.A., Yuen, H.K., 1990. Statistics of multivariate extremes. International Statistical Review 58, 47–58.

Tawn, J.A., 1988. Bivariate extreme value theory: models and estimation. Biometrika 75, 397–415.

Tiago de Oliveira, J., 1984. Bivariate models for extremes. In: Statistical Extremes and Applications, J. Tiago de Oliveira, Ed. Dordrecht, Reidel, pp. 131–153.

Tyler, D.E., 1987. A distribution-free M-estimator of multivariate scatter. The Annals of Statistics 15, 234–251.

Wang, W., Wells, M.T., 2000. Model selection and semiparametric inference for bivariate failure-time data. Journal of the American Statistical Association 95, 62–76.

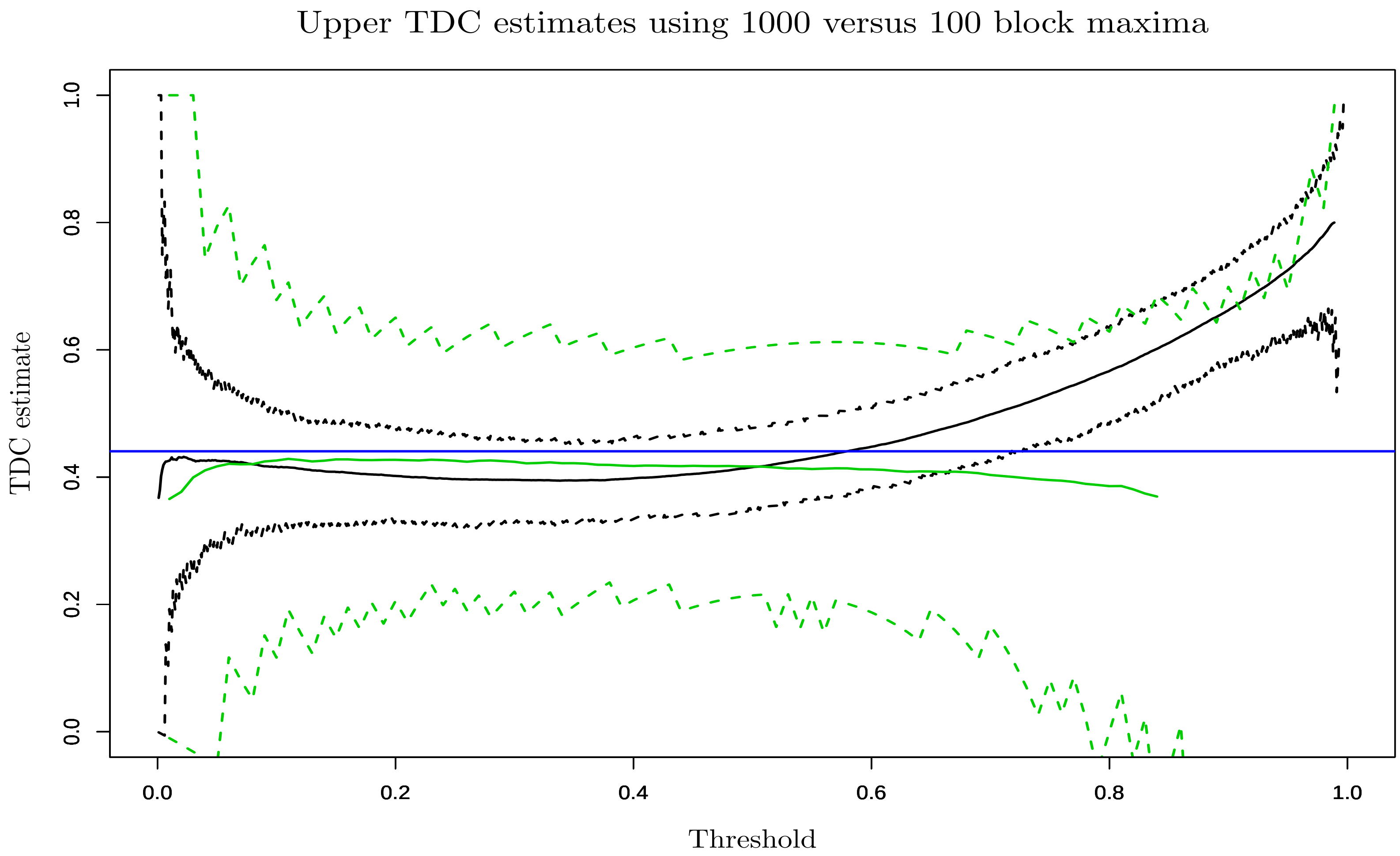

Figure 1: Sample means of upper TDC estimates for various thresholds k using estimator (7) for 1000 samples of a bivariate t-distribution with 1.5 degrees of freedom and correlation parameter $\rho = 0.5$. The black solid line represents the case of 1000 block maxima ($\triangleq$ original data set) and the gray solid line corresponds to 100 block maxima. The related empirical 95% confidence intervals are indicated by the dashed lines. The true value of the TDC is marked by the solid straight line.

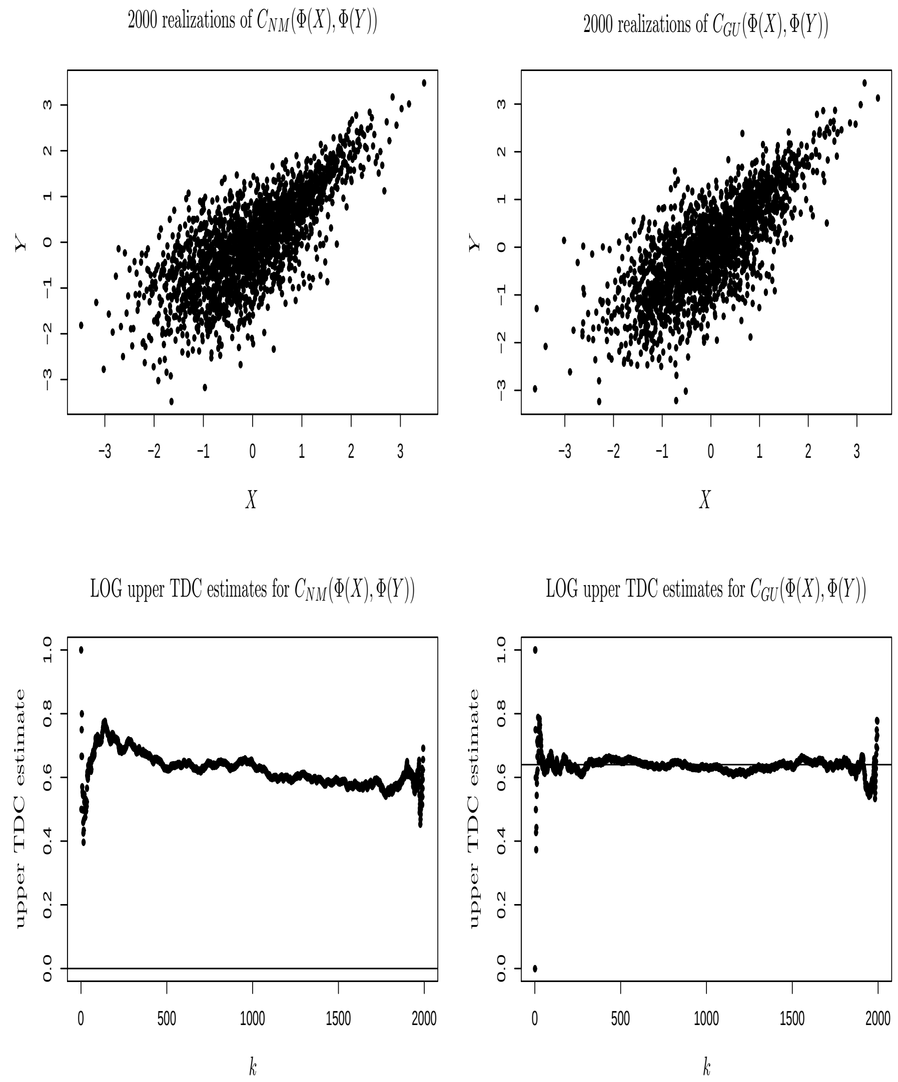

Figure 2: Scatterplots of 2000 simulated data with standard normal margins and copula C_{NM} (upper left) and Gumbel copula C_{GU} (upper right), respectively. The lower plots show the corresponding TDC estimates $\hat{\lambda}_U^{\mathrm{LOG}}$ for different choices of k. The horizontal lines indicate the true TDCs.

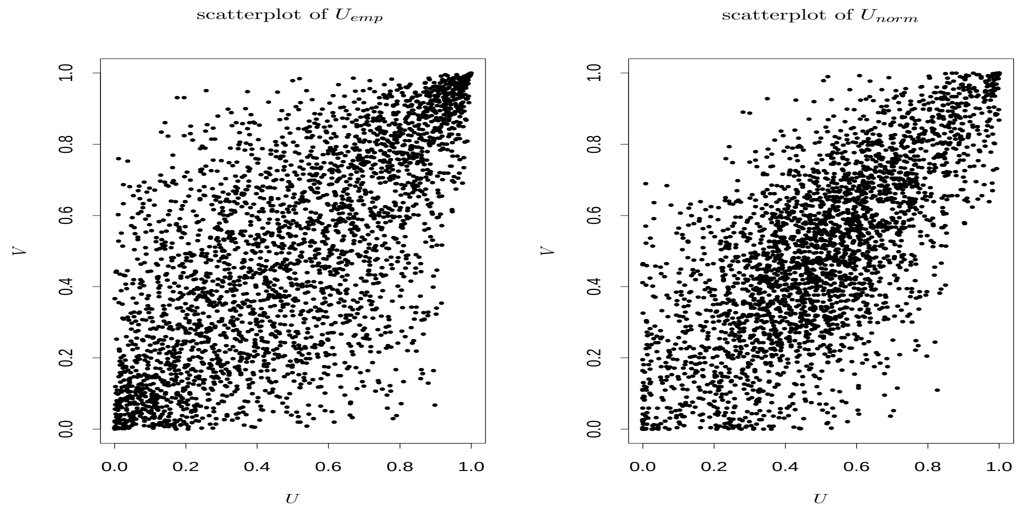

Figure 3: Comparison of empirical copula densities obtained via empirical marginal distributions (left panel) and via fitted normal marginal distributions (right panel). The marginal transformations in the right panel are misspecified.

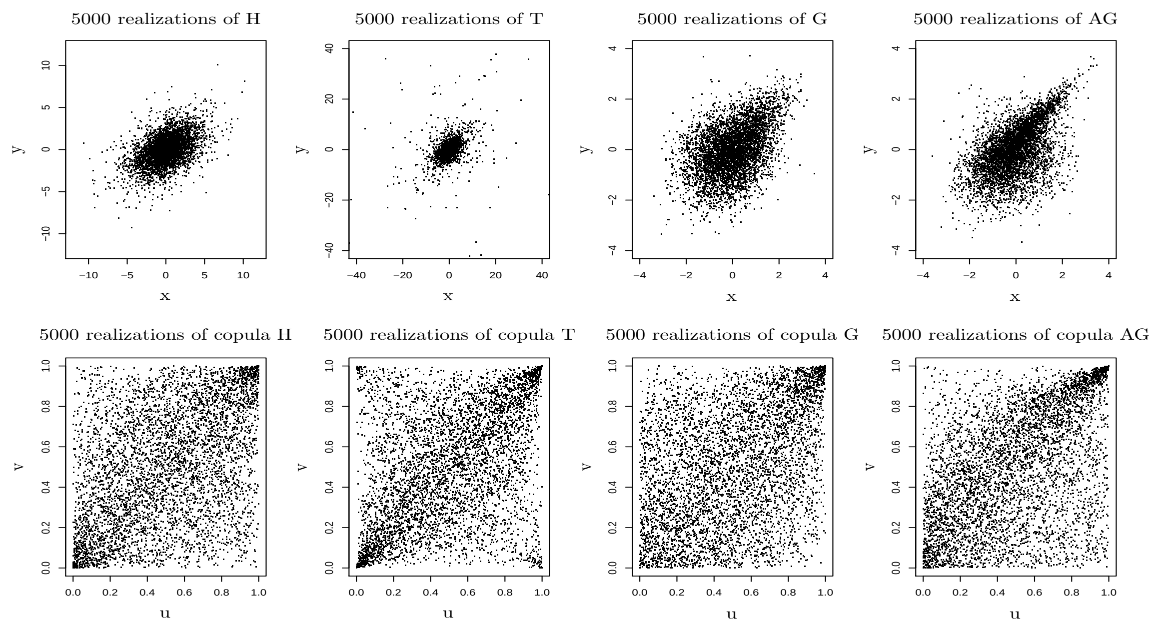

Figure 4: Scatter plots of simulated distributions and corresponding empirical copula realizations.