Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 1807–1824
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Bootstrap approximation of tail dependence function

Liang Peng[a], Yongcheng Qi[b,*]

[a] School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA
[b] Department of Mathematics and Statistics, University of Minnesota Duluth, SCC 140, 1117 University Drive, Duluth, MN 55812, USA

Received 9 January 2007
Available online 9 February 2008

Abstract

For estimating a rare event via the multivariate extreme value theory, the so-called tail dependence function has to be investigated (see [L. de Haan, J. de Ronde, Sea and wind: Multivariate extremes at work, Extremes 1 (1998) 7–45]). A simple, but effective estimator for the tail dependence function is the tail empirical distribution function, see [X. Huang, Statistics of Bivariate Extreme Values, Ph.D. Thesis, Tinbergen Institute Research Series, 1992] or [R. Schmidt, U. Stadtmüller, Nonparametric estimation of tail dependence, Scand. J. Stat. 33 (2006) 307–335]. In this paper, we first derive a bootstrap approximation for a tail dependence function with an approximation rate via the construction approach developed by [K. Chen, S.H. Lo, On a mapping approach to investigating the bootstrap accuracy, Probab. Theory Relat. Fields 107 (1997) 197–217], and then apply it to construct a confidence band for the tail dependence function. A simulation study is conducted to assess the accuracy of the bootstrap approach.

AMS 2000 subject classifications: 62G32; 62G09

Keywords: Bootstrap; Confidence band; Dependence function; Extremes; Tail empirical process

1. Introduction

Suppose that $\{\mathbf{X}_j = (X_j^{(1)}, X_j^{(2)})^{\mathrm{T}}, 1 \le j \le n\}$ are n independent $\mathbf{R}^2$-valued random vectors having a common distribution F with continuous marginal distributions $F_i,\ i = 1, 2$. In applying bivariate extreme value theory to predict a rare event, one important quantity is the so-called tail

* Corresponding author.
E-mail address: yqi@d.umn.edu (Y. Qi).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2008.01.018

dependence function, which is defined as

$$l(x_1, x_2) = \lim_{n\to\infty} n \left\{1 - F\left(F_1^-\left(1 - \frac{x_1}{n}\right), F_2^-\left(1 - \frac{x_2}{n}\right)\right)\right\} \tag{1.1}$$

for $x_i \geq 0$, $i = 1, 2$, where $(\cdot)^-$ denotes the inverse function of $(\cdot)$ (see [9,10]). Since this tail dependence function is defined as a limit and it is homogeneous, one can extrapolate data into a far tail region via the estimated tail dependence function and then estimate the probability of a rare event.

For estimating the tail dependence function $l(x_1, x_2)$, a natural estimator is used to replace F, F_1, F_2 in the right-hand side of (1.1) by the corresponding empirical distributions. This results in the so-called tail empirical distribution function

$$\hat{l}_n(x_1, x_2) = \frac{n}{m}\left\{1 - F_n\left(F_{n1}^-\left(1 - \frac{m}{n}x_1\right), F_{n2}^-\left(1 - \frac{m}{n}x_2\right)\right)\right\}, \tag{1.2}$$

where

$$F_n(x_1, x_2) = \frac{1}{n}\sum_{j=1}^n I(X_j^{(1)} \leq x_1, X_j^{(2)} \leq x_2),$$

$$F_{ni}(x) = \frac{1}{n}\sum_{j=1}^n I(X_j^{(i)} \leq x), \qquad \text{for } i = 1, 2,$$

and $m = m(n)$ are integers satisfying

$$m \to \infty \quad \text{and} \quad \frac{m}{n} \to 0 \quad \text{as } n \to \infty. \tag{1.3}$$

Huang [12] proved that this estimator is weakly consistent under (1.3) and its limiting distribution is normal under certain additional regularity conditions listed in the next section; see also [17]. Qi [16] obtained the strong consistence of this estimator under condition (1.3) and $m/\log\log n \to \infty$. The optimal rate of convergence for estimating $l(x_1, x_2)$ is given by Drees and Huang [3]. A weighted approximation for $\hat{l}_n(x_1, x_2)$ is derived by Einmahl et al. [4] and is applied to test multivariate extreme value conditions.

It is known that the limiting distribution of $\hat{l}_n(x_1, x_2)$ depends on the partial derivatives of $l(x_1, x_2)$ and a bivariate Gaussian process with a covariance structure depending on $l(x_1, x_2)$ (see (2.4) in the next section), i.e., the asymptotic variance of $\hat{l}_n(x_1, x_2)$ depends on the tail dependence function and its partial derivatives. Therefore, for constructing a confidence interval for the tail dependence function via the normal approximation method, one has to estimate the partial derivatives first. Recently, Peng and Qi [15] proposed smooth estimation of these partial derivatives and obtained confidence intervals for the tail dependence function via the normal approximation method. In order to construct a confidence band via the limit, one has to not only estimate the partial derivatives but also simulate a bivariate Gaussian process with the given covariance structure depending on the tail dependence function and its partial derivatives. Hence this method is quite computationally intensive and also impractical since one cannot tabulate critical values for each tail dependence function. An alternative way to construct a confidence band is by using bootstrap methods. Since full sample bootstrap method fails to catch the bias term of many statistics of extremes, subsample bootstrap method is proposed to approximate the optimal mean squared error of tail index estimation (see [11,2]), to approximate the distribution of tail index estimation (see [7]), and to approximate the distribution of a high quantile (see [8]).

However, when constructing a confidence interval or band for a smooth curve, one commonly used approach is to undersmooth, i.e., to choose a small tuning parameter such that the bias is negligible. By doing this, the full sample bootstrap method is valid in constructing a confidence interval or band in general.

In this paper, we first derive a bootstrap approximation for a tail dependence function with an approximation rate via the construction approach developed by Chen and Lo [1], and then we apply it to construct a confidence band for the tail dependence function; see Section 2 for details. Therefore, we extend the result on copulas in [1] to tail dependence functions. A simulation study and real application are given in Section 3. Proofs are given in Section 4 and in the Appendix.

2. Main results

In order to derive the asymptotic properties of $\hat{l}_n(x_1, x_2)$ we need a stricter condition than (1.1): suppose that there exists a regularly varying function $A(t) \to 0$ such that

$$\lim_{t\downarrow 0} \frac{t^{-1}\{1 - F(F_1^-(1-tx_1), F_2^-(1-tx_2))\} - l(x_1, x_2)}{A(t)} = c(x_1, x_2) \tag{2.1}$$

holds uniformly on $\Theta = \{(x_1, x_2)^{\mathrm{T}} : x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 = 1\}$, where $c(x_1, x_2)$ is non-constant and not a multiple of $l(x_1, x_2)$. Note that the function A is employed to control the bias term introduced by a large value of m. Further we assume that

$$l(x_1, x_2) \text{ has continuous first partial derivatives}$$
$$l_i(x_1, x_2) = \frac{\partial}{\partial x_i} l(x_1, x_2), \quad (i = 1, 2) \tag{2.2}$$

and

$$m \to \infty, \qquad m/n \to 0, \qquad \sqrt{m} A\left(\frac{m}{n}\right) \to 0 \text{ as } n \to \infty. \tag{2.3}$$

Then, under conditions (2.1)–(2.3), we have

$$\sqrt{m}(\hat{l}_n(x_1, x_2) - l(x_1, x_2)) \xrightarrow{D} B(x_1, x_2) \quad \text{in } [0, \infty)^2, \tag{2.4}$$

where $B(x_1, x_2) = W(x_1, x_2) - l_1(x_1, x_2)W(x_1, 0) - l_2(x_1, x_2)W(0, x_2)$ and $W(x_1, x_2)$ is a Gaussian process with zero mean and covariance structure

$$\begin{aligned} E\{W(x_1, x_2)W(y_1, y_2)\} &= l(x_1 \wedge y_1, x_2) + l(x_1 \wedge y_1, y_2) + l(x_1, x_2 \wedge y_2) \\ &\quad + l(y_1, x_2 \wedge y_2) - l(x_1, y_2) - l(y_1, x_2) - l(x_1 \wedge y_1, x_2 \wedge y_2) \end{aligned}$$

(see [12] or [17]). When $\sqrt{m}A(m/n)$ converges to a finite non-zero constant, a bias term will appear on the right-hand side of (2.4). Thus condition (2.3) implies that the asymptotic bias in $\hat{l}_n(x_1, x_2)$ is negligible. This allows us to employ the full sample bootstrap method to approximate the distribution of $\sqrt{m}\{\hat{l}_n(x_1, x_2) - l(x_1, x_2)\}$.

Let $\{\mathbf{X}_j^* = (X_j^{(1)*}, X_j^{(2)*})^{\mathrm{T}}, 1 \leq j \leq n\}$ be a bootstrap sample of $\{\mathbf{X}_j, 1 \leq j \leq n\}$, $\hat{l}_n^*(x_1, x_2)$ be the bootstrap statistic of $\hat{l}_n(x_1, x_2)$, and P^ω and E^ω denote the conditional probability and expectation, respectively, given the data $\mathbf{X}_1, \ldots, \mathbf{X}_n$. For any given compact set $B \subset (0, \infty)^2$

define

$$\Delta_n(t) = P^\omega\left(\sqrt{m}\sup_{(x_1,x_2)^{\mathrm{T}}\in B}|\hat{l}_n^*(x_1,x_2)-\hat{l}_n(x_1,x_2)|\le t\right) - P\left(\sqrt{m}\sup_{(x_1,x_2)^{\mathrm{T}}\in B}|\hat{l}_n(x_1,x_2)-l(x_1,x_2)|\le t\right).$$

A general method to derive a uniform rate for $\Delta_n(t)$ is to identify the asymptotic distributions of $\sqrt{m}\{\hat{l}_n^*(x_1,x_2)-\hat{l}_n(x_1,x_2)\}$ and $\sqrt{m}\{\hat{l}_n(x_1,x_2)-l(x_1,x_2)\}$ to certain order. Since it involves tail empirical processes for bootstrap samples, it is difficult to determine the convergence rate $\Delta_n(t)$ for this naive bootstrap method. Recently, Chen and Lo [1] argued that the essence of the bootstrap accuracy relies on the error between the studied statistic and its bootstrap version rather than their asymptotic distributions, and proposed a simple mapping approach to study such bootstrap accuracy. From this bootstrap accuracy, one can derive the uniform convergence rate for the difference between distributions of a statistic and its bootstrap version. One of the specific examples studied in [1] is on copulas. Here we extend these results to tail dependence functions as follows. The key is to first follow the mapping approach in [1] to construct an independent copy of $\{\mathbf{X}_j, 1\le j\le n\}$, say $\{\mathbf{X}'_j, 1\le j\le n\}$, such that the bootstrap sample $\{\mathbf{X}^*_j, 1\le j\le n\}$ shares the same space with $\{\mathbf{X}'_j, 1\le j\le n\}$. Let $\hat{l}'_n(x_1,x_2)$ denote the tail empirical distribution function based on the sample $\{\mathbf{X}'_j, 1\le j\le n\}$. Since $\sqrt{m}\{\hat{l}_n(x_1,x_2)-l(x_1,x_2)\}$ and $\sqrt{m}\{\hat{l}'_n(x_1,x_2)-l(x_1,x_2)\}$ have the same distribution, assessing the bootstrap accuracy amounts to deriving the rate between $\sqrt{m}\{\hat{l}_n^*(x_1,x_2)-\hat{l}_n(x_1,x_2)\}$ and $\sqrt{m}\{\hat{l}'_n(x_1,x_2)-l(x_1,x_2)\}$. Next we employ similar arguments in the copula example of Chen and Lo [1] through replacing distribution and quantile approximations by tail distribution and high quantile approximations; see the proof of Theorem 2.1 in Section 4 for details.

Our main result is the following theorem. In this theorem, $\{\mathbf{X}'_j,\ j\ge 1\}$ is an independent copy of $\{\mathbf{X}_j,\ j\ge 1\}$, $\hat{l}'_n(x,y)$ is defined as an estimate of $l(x,y)$ like (1.2) based on $\{\mathbf{X}'_j,\ 1\le j\le n\}$, and the bootstrap sample with certain properties will be constructed and explained in the proof as we have mentioned above.

Theorem 2.1. *In addition to conditions (2.1)–(2.3), if $m/(\log n)^6\to\infty$ as $n\to\infty$ and $l(x_1,x_2)$ has bounded second derivatives on a compact set $B\subset(0,\infty)^2$, then, almost surely*

$$\sup_{(x_1,x_2)^{\mathrm{T}}\in B}|\sqrt{m}(\hat{l}_n^*(x_1,x_2)-\hat{l}_n(x_1,x_2))-\sqrt{m}(\hat{l}'_n(x_1,x_2)-l(x_1,x_2))| = O\left(\sqrt{m}A\left(\frac{m}{n}\right)+\frac{(\log n)^{3/5}}{m^{1/10}}\right).$$

Remark. Using the arguments in (2) of [1], it follows from (2.4) and Theorem 2.1 that $\Delta_n(t)=o(1)$ almost surely. Therefore, we could construct a bootstrap confidence band for the tail dependence function as follows. Since (1.1) implies that $l(ax_1,ax_2)=al(x_1,x_2)$ for any $a>0$ (cf. [12, p. 38]) we only need to construct a confidence band for $l(\cos(\theta),\sin(\theta))$ for $\theta\in[0,\frac{\pi}{2}]$. As a matter of fact, we have to exclude the two endpoints of the interval since the second derivatives of the tail dependence functions at these two points are not bounded. Let z_α^* denote the α quantile of $\sup_{\theta_1 \le \theta \le \theta_2} |\hat{l}_n^*(\cos(\theta), \sin(\theta)) - \hat{l}_n(\cos(\theta), \sin(\theta))|$, where $0 < \theta_1 < \theta_2 < \pi/2$. Then a confidence band for $l(\cos(\theta), \sin(\theta))$, $\theta \in [\theta_1, \theta_2]$, with level α is

$$I_\alpha := \left(\hat{l}_n(\cos(\theta), \sin(\theta)) - z_\alpha^*,\ \hat{l}_n(\cos(\theta), \sin(\theta)) + z_\alpha^*\right).$$

Indeed, using the arguments in (2) of [1], we can show that $\Delta_n(t) = O(\delta_n + \sqrt{m} A\left(\frac{m}{n}\right) + \frac{(\log n)^{3/5}}{m^{1/10}})$ almost surely if

$$P\left(\sup_{(x_1,x_2)^{\mathrm{T}} \in B} \sqrt{m}\{\hat{l}_n(x_1, x_2) - l(x_1, x_2)\} \le t\right) - P\left(\sup_{(x_1,x_2)^{\mathrm{T}} \in B} B(x_1, x_2) \le t\right) = O(\delta_n), \tag{2.5}$$

where $B(x_1, x_2)$ and B are given in (2.4) and Theorem 2.1, respectively. Therefore, obtaining a rate for $\Delta_n(t)$ requires finding the rate δ_n in (2.5), which remains unknown in the literature.

3. Simulations and a real data application

We examine the finite sample behavior of the bootstrap confidence bands in terms of coverage accuracy. In general, it is difficult to simulate random vectors with a given tail dependence function. Recently, Klüppelberg, Kuhn and Peng [13] obtained an explicit expression for the tail dependence function of an elliptical random vector. More specifically, the tail dependence function of an elliptical vector RAU is

$$l(x, y) = x + y - \frac{x \int_{g((x/y)^{1/\alpha})}^{\pi/2} \cos^\alpha(\phi)\, \mathrm{d}\phi + y \int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})} \sin^\alpha(\phi + \arcsin\rho)\, \mathrm{d}\phi}{\int_{-\pi/2}^{\pi/2} \cos^\alpha(\phi)\, \mathrm{d}\phi},$$

where $g(t) = \arctan((t - \rho)/\sqrt{1-\rho^2})$, when the elliptical vector RAU satisfies that

$$AA^{\mathrm{T}} = \begin{pmatrix} \sigma^2 & \rho\sigma\nu \\ \rho\sigma\nu & \nu^2 \end{pmatrix},$$

$rank(AA^{\mathrm{T}}) = 2$, $-1 < \rho < 1$, $R > 0$ is a heavy tailed random variable (i.e., $\lim_{t\to\infty} P(R > tx)/P(R > t) = x^{-\alpha}$ for some $\alpha > 0$), $U = (U_1, U_2)^{\mathrm{T}}$ is a random vector uniformly distributed on the unit sphere $\{(u_1, u_2) : u_1^2 + u_2^2 = 1\}$, and U is independent of R. For applications of elliptical distributions and elliptical copulas in risk management, we refer the reader to [14].

Here we draw 500 random samples of size $n = 1000$ from the above elliptical random vector with

$$A = \begin{pmatrix} \dfrac{\sqrt{1+\rho} + \sqrt{1-\rho}}{2} & \dfrac{\sqrt{1+\rho} - \sqrt{1-\rho}}{2} \\ \dfrac{\sqrt{1+\rho} - \sqrt{1-\rho}}{2} & \dfrac{\sqrt{1+\rho} + \sqrt{1-\rho}}{2} \end{pmatrix},$$

$-1 < \rho < 1$, and $P(R \le x) = \exp\{-x^{-1}\}$. For each random sample, we draw 1000 bootstrap samples of size $n = 1000$ to obtain the bootstrap confidence band I_α, given in Section 2. We consider the confidence band for $l(\cos(\theta), \sin(\theta))$, $0.01 \le \theta \le 1.57$. In computing $\sup_{0.01 \le \theta \le 1.57} |\hat{l}_n^*(\cos(\theta), \sin(\theta)) - \hat{l}_n(\cos(\theta), \sin(\theta))|$ and the empirical coverage probability of I_α, we maximize the values for $\theta = 0.01 + 0.01j$ for $j = 0, 1, \ldots, 156$. Take $\rho = 0.5$. The

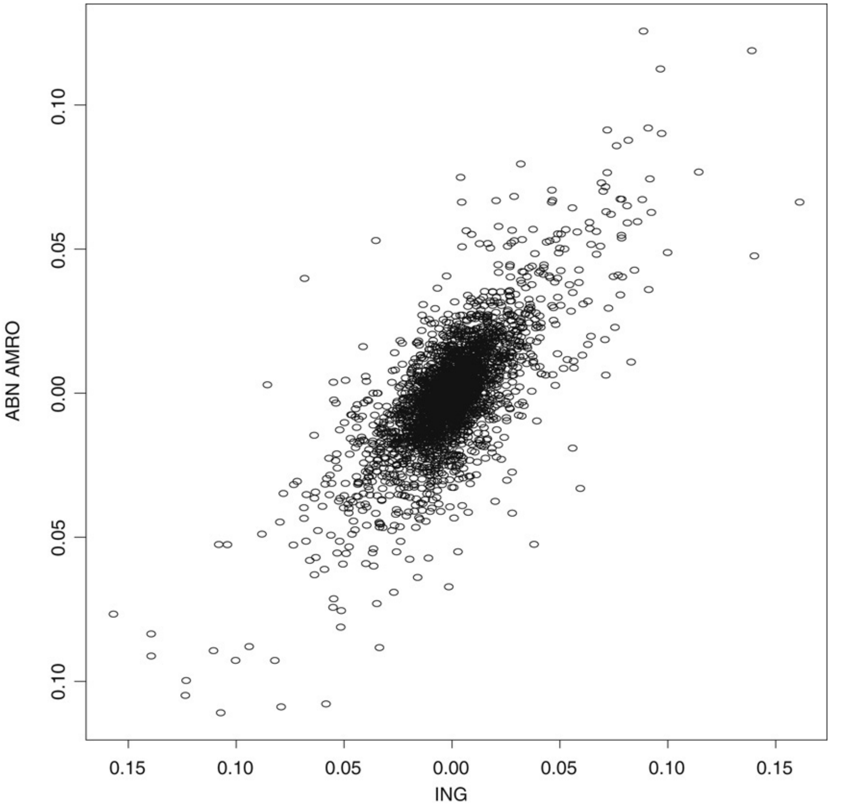

Fig. 1. Log-returns of Equity for two Dutch banks, ING and ABN AMRO, over the period 1991–2003.

empirical coverage probabilities of $I_{0.9}$ are 0.958, 0.928, 0.886 and 0.864 for $m = 50, 100, 150$ and 200, respectively, and those of $I_{0.95}$ are 0.978, 0.958, 0.942 and 0.918 for $m = 50, 100, 150$ and 200, respectively. As we see, the coverage accuracy becomes worse when m is too large. This is due to the fact that we do not take care of the bias term in the bootstrap method. We leave the difficult, but practically important issue of choosing m to future research.

Next, as an illustration example, we apply the bootstrap confidence bands to a real data set on the 3283 daily log-returns of equity for two major Dutch banks (ING and ABN AMRO Bank) over the period 1991–2003; see Fig. 1. This data set has been confirmed by Einmahl, de Haan and Li [4] to be in the domain of attraction of a bivariate extreme value distribution. Here we construct a confidence band for $l(\cos(\theta), \sin(\theta))$, $0.01 \leq \theta \leq 1.57$; see Figs. 2 and 3. In these two figures, we also plot the function $\cos(\theta) + \sin(\theta)$, which corresponds to the case of asymptotic independence. Although our method only works for the case of asymptotic dependence, intuitively the function $\cos(\theta) + \sin(\theta)$ should be outside of the confidence band when the underlying distribution is away from the case of asymptotic independence. Figs. 1 and 2 indicate that observations in the data set have asymptotically dependent tails, and thus bivariate extreme value theory can safely be employed to predict a rare event. If the observations in the data set have asymptotically independent tails, then some additional model assumptions will be needed for predicting a rare event.

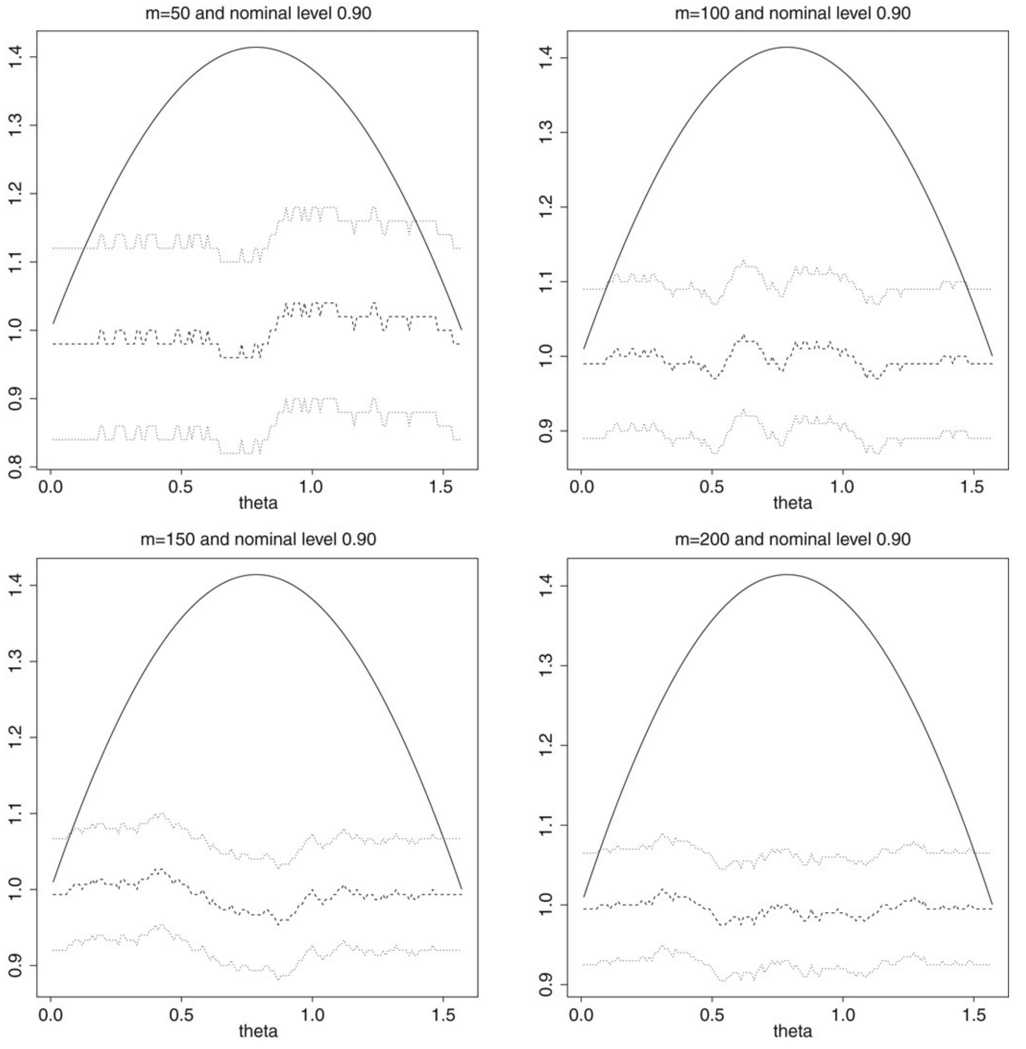

Fig. 2. Bootstrap confidence bands with level 0.90 for the log-returns of equity of two Dutch banks, ING and ABN AMRO. The solid line, dashed line and dotted line represent $\cos(\theta) + \sin(\theta)$, $\hat{l}_n(\cos(\theta), \sin(\theta))$ and endpoints of $I_{0.90}$, respectively.

4. Proofs

We closely follow the construction method and the copula example of [1], but our proof is different from the copula example given in [1] since we have to replace distribution and quantile approximations by tail distribution and high quantile approximations, respectively. In order to present the idea of our proof of Theorem 2.1 clearly, we state these results (see (4.5)–(4.8)) on tail distributions and high quantiles in the proof of Theorem 2.1, but give the proofs in the Appendix.

Proof of Theorem 2.1. Since the estimator $\hat{l}_n(x, y)$ is free of marginals, for the sake of simplicity we assume that the two marginal distributions are uniform over $(0, 1)$, that is, $F_i(x) =$

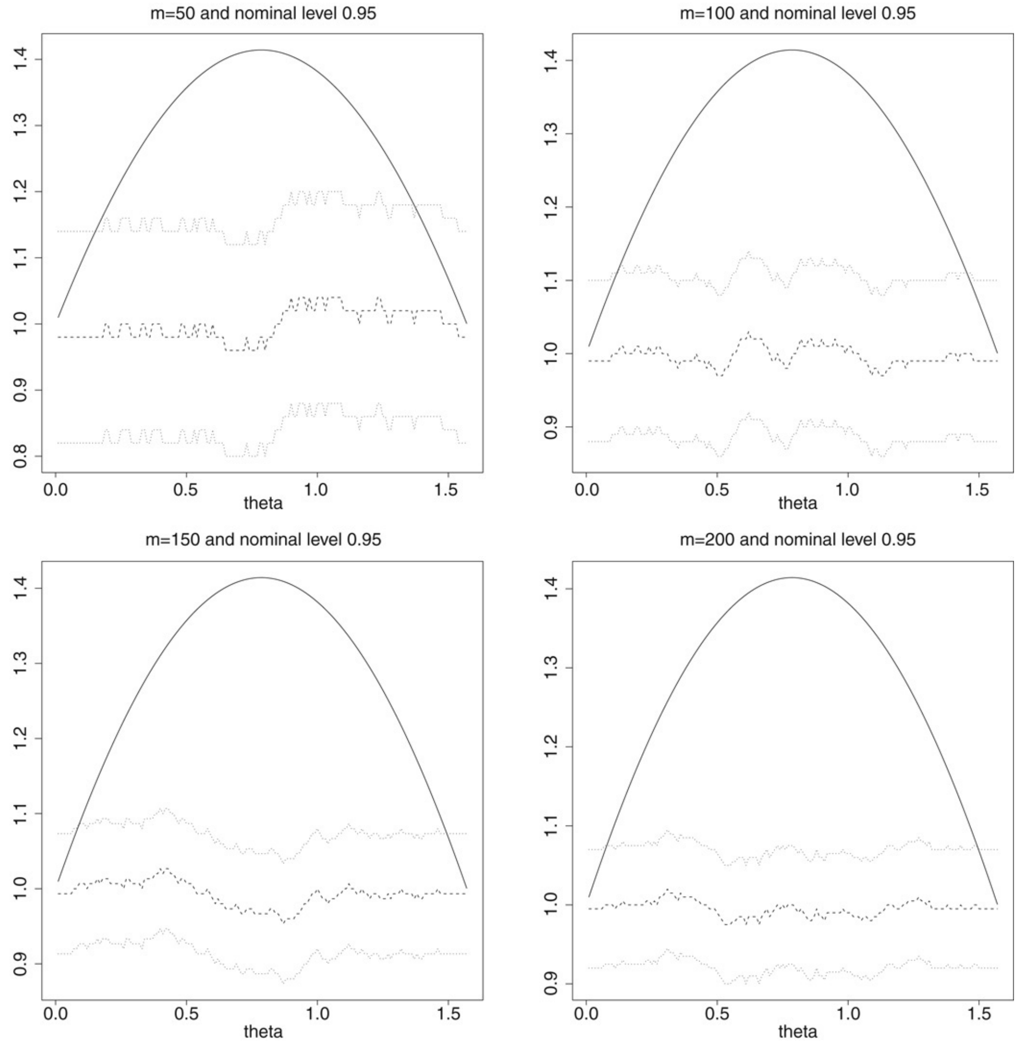

Fig. 3. Bootstrap confidence bands with level 0.95 for the log-returns of equity of two Dutch banks, ING and ABN AMRO. The solid line, dashed line and dotted line represent $\cos(\theta) + \sin(\theta)$, $\hat{l}_n(\cos(\theta), \sin(\theta))$ and endpoints of $I_{0.95}$, respectively.

x, $x \in (0, 1)$ for $i = 1, 2$. Define $\mathbf{Z}_i = (1 - X_i^{(1)}, 1 - X_i^{(2)})^{\mathrm{T}}$ and $\mathbf{Z}_i' = (1 - X_i^{(1)'}, 1 - X_i^{(2)'})^{\mathrm{T}}$ with $\{\mathbf{X}_i' = (X_i^{(1)'}, X_i^{(2)'})^{\mathrm{T}}, 1 \le i \le n\}$ being an independent copy of $\{\mathbf{X}_i, 1 \le i \le n\}$. Below we will construct a bootstrap sample $\{\mathbf{Z}_i^*, 1 \le i \le n\}$ for $\{\mathbf{Z}_i, 1 \le i \le n\}$, and thus obtain the bootstrap sample $\{\mathbf{X}_i^* := \mathbf{1} - \mathbf{Z}_i^*, 1 \le i \le n\}$ for the original sample $\{\mathbf{X}_i, 1 \le i \le n\}$. Since B given in Theorem 2.1 is a compact set, we can assume that $B \subset [0, 1]^2$ for simplicity.

Throughout we define $t_n = m/n$. Let k_n denote an integer such that

$$\left(\frac{m}{\log n}\right)^{3/5} \le 2^{k_n} \le 2\left(\frac{m}{\log n}\right)^{3/5}.$$

For each $1 \le i \le k_n$, set

$$J_{i,r} = \left[\frac{2t_n(r-1)}{2^i}, \frac{2t_n r}{2^i}\right], \quad r = 1, \ldots, 2^i$$

and define

$$\mathcal{F}_n^{(i)} = \{J_{i,r_1} \times J_{i,r_2} : 1 \le r_1, r_2 \le 2^i\}.$$

Set $D_0 = \{(x_1, x_2)^{\mathrm{T}} : x_1 > 2t_n \quad \text{or} \quad x_2 > 2t_n\} \cap [0,1]^2$. It is easy to see that each $\mathcal{F}_n^{(i)} \cup \{D_0\}$ is a partition of $[0,1]^2$, and $\sigma(\mathcal{F}_n^{(i)}, D_0) \subset \sigma(\mathcal{F}_n^{(i+1)}, D_0)$. Then it follows from [1] that there exists a mapping g_n on $\mathcal{F}_n = \cup_{i=1}^{k_n} \mathcal{F}_n^{(i)}$ such that $\{\mathbf{Z}_j^* := g_n(\mathbf{Z}_j'),\ 1 \le j \le n\}$ is a bootstrap sample of $\{\mathbf{Z}_i, 1 \le i \le n\}$, and

$$E^{\omega}(|I(\mathbf{Z}_1^* \in D) - I(\mathbf{Z}_1' \in D)|) = |P_n(D) - P(D)| \tag{4.1}$$

for all $D \in \mathcal{F}_n$, where $P_n(D) = \frac{1}{n}\sum_{i=1}^n I(\mathbf{Z}_i \in D)$; see Equations (7) and (9) and the discussion after Corollary 2.3 in [1].

For each $D \in \mathcal{F}_n^{(i)}$, say $D = J_{i,r_1} \times J_{i,r_2}$, we have

$$P(D) \le \min\{P(Z_1^{(1)} \in J_{i,r_1}), P(Z_1^{(2)} \in J_{i,r_2})\} = 2t_n 2^{-i}. \tag{4.2}$$

Eqs. (4.1) and (4.2) will be used in the proofs in the Appendix.

Set

$$\begin{cases} H_n(x_1, x_2) = \dfrac{1}{m}\displaystyle\sum_{i=1}^n I(Z_i^{(1)} \le t_n x_1 \text{ or } Z_i^{(2)} \le t_n x_2) \\ H_n^*(x_1, x_2) = \dfrac{1}{m}\displaystyle\sum_{i=1}^n I(Z_i^{(1)*} \le t_n x_1 \text{ or } Z_i^{(2)*} \le t_n x_2) \\ H_n'(x_1, x_2) = \dfrac{1}{m}\displaystyle\sum_{i=1}^n I(Z_i^{(1)'} \le t_n x_1 \text{ or } Z_i^{(2)'} \le t_n x_2) \\ \bar{H}_n(x_1, x_2) = \dfrac{n}{m} P(Z_1^{(1)} \le t_n x_1 \text{ or } Z_1^{(2)} \le t_n x_2). \end{cases}$$

Let $H_{n1}(x) = H_n(x, 0)$, $H_{n2}(x) = H_n(0, x)$, $H_{n1}^*(x) = H_n^*(x, 0)$, $H_{n2}^*(x) = H_n^*(0, x)$, $H_{n1}'(x) = H_n'(x, 0)$ and $H_{n2}'(x) = H_n'(0, x)$. Note that $\bar{H}_n(x, 0) = \bar{H}_n(0, x) = x$. From Lemma 4.7 in the Appendix we have

$$\sup_{0 \le x \le 2} |H_{nj}^*(x) - H_{nj}(x) - H_{nj}'(x) + x| = O(m^{-3/5}(\log n)^{3/5}) \quad \text{a.s.} \tag{4.3}$$

for $j = 1, 2$. Further, by combining Lemmas 4.6 and 4.7 we obtain

$$\sup_{0 \le x_1, x_2 \le 2} |H_n^*(x_1, x_2) - H_n(x_1, x_2) - H_n'(x_1, x_2) + \bar{H}_n(x_1, x_2)| \\ = O(m^{-3/5}(\log n)^{3/5}) \quad \text{a.s.} \tag{4.4}$$

In the Appendix we shall prove the following equations

$$\begin{cases} \sup_{0\le x\le 2} |H_{nj}^{-}(x)-x| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.} \\ \sup_{0\le x\le 2} |H_{nj}^{'-}(x)-x| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.} \\ \sup_{0\le x\le 2} |H_{nj}^{*-}(x)-x| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.,} \end{cases} \tag{4.5}$$

$$\begin{cases} \sup_{0\le x\le 2} |H_{nj}^{'-}(x)-x+H'_{nj}(x)-x| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \\ \sup_{0\le x\le 2} |H_{nj}^{*-}(x)-H_{nj}^{-}(x)+H_{nj}^{*}(x)-H_{nj}(x)| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \end{cases} \tag{4.6}$$

for $j=1,2$,

$$\sup_{(x_1,x_2,y_1,y_2)\in S} |H_n(x_1,x_2)-\bar{H}_n(x_1,x_2)-H_n(y_1,y_2) \\ +\bar{H}_n(y_1,y_2)| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \tag{4.7}$$

and

$$\sup_{(x_1,x_2,y_1,y_2)\in S} |H'_n(x_1,x_2)-\bar{H}_n(x_1,x_2)-H'_n(y_1,y_2) \\ +\bar{H}_n(y_1,y_2)| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \tag{4.8}$$

where $S=\{(x_1,x_2,y_1,y_2): 0\le x_1,x_2,y_1,y_2\le 2, |x_1-y_1|+|x_2-y_2| = O(m^{-1/2}(\log n)^{1/2})\}$. We remark that similar results can be found in the literature, cf. Theorem 5 of [5] and Theorem 3.1 of [6], but more restrictive conditions on the sequence $\{m(n)\}$ are imposed and those results cannot be applied to the bootstrap sample directly.

By (4.3) and (4.6),

$$\sup_{0\le x\le 2} |H_{nj}^{*-}(x)-H_{nj}^{-}(x)-H_{nj}^{'-}(x)+x| = O(m^{-3/5}(\log n)^{3/5}) \quad \text{a.s.} \tag{4.9}$$

By (4.4), (4.7) and (4.8),

$$\sup_{(x_1,x_2,y_1,y_2)\in S} |H^*_n(x_1,x_2)-\bar{H}_n(x_1,x_2)-H^*_n(y_1,y_2) \\ +\bar{H}_n(y_1,y_2)| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \tag{4.10}$$

Hence, by (2.1), (4.4), (4.5) and (4.7)–(4.10),

$$\begin{aligned} &\hat{l}^*_n(x_1,x_2)-\hat{l}_n(x_1,x_2)-\hat{l}'_n(x_1,x_2)+l(x_1,x_2) \\ &= H^*_n(H_{n1}^{*-}(x_1),H_{n2}^{*-}(x_2))-H_n(H_{n1}^{-}(x_1),H_{n2}^{-}(x_2)) \\ &\quad -H'_n(H_{n1}^{'-}(x_1),H_{n2}^{'-}(x_2))+l(x_1,x_2) \\ &= H^*_n(H_{n1}^{*-}(x_1),H_{n2}^{*-}(x_2))-\bar{H}_n(H_{n1}^{*-}(x_1),H_{n2}^{*-}(x_2)) \\ &\quad -H^*_n(x_1,x_2)+\bar{H}_n(x_1,x_2)-H_n(H_{n1}^{-}(x_1),H_{n2}^{-}(x_2)) \\ &\quad +\bar{H}_n(H_{n1}^{-}(x_1),H_{n2}^{-}(x_2))+H_n(x_1,x_2)-\bar{H}_n(x_1,x_2) \\ &\quad -H'_n(H_{n1}^{'-}(x_1),H_{n2}^{'-}(x_2))+\bar{H}_n(H_{n1}^{'-}(x_1),H_{n2}^{'-}(x_2))+H'_n(x_1,x_2)-\bar{H}_n(x_1,x_2) \end{aligned}$$

$$
\begin{aligned}
&+ H_n^*(x_1, x_2) - H_n(x_1, x_2) - H_n'(x_1, x_2) + \bar{H}_n(x_1, x_2)\\
&+ \bar{H}_n(H_{n1}^{*-}(x_1), H_{n2}^{*-}(x_2)) - l(H_{n1}^{*-}(x_1), H_{n2}^{*-}(x_2))\\
&- \bar{H}_n(H_{n1}^{-}(x_1), H_{n2}^{-}(x_2)) + l(H_{n1}^{-}(x_1), H_{n2}^{-}(x_2))\\
&- \bar{H}_n(H_{n1}^{'-}(x_1), H_{n2}^{'-}(x_2)) + l(H_{n1}^{'-}(x_1), H_{n2}^{'-}(x_2))\\
&+ l(H_{n1}^{*-}(x_1), H_{n2}^{*-}(x_2)) - l(x_1, x_2) - l(H_{n1}^{-}(x_1), H_{n2}^{-}(x_2)) + l(x_1, x_2)\\
&- l(H_{n1}^{'-}(x_1), H_{n2}^{'-}(x_2)) + l(x_1, x_2)\\
={}& O(m^{-3/5}(\log n)^{3/5}) + O(|A(m/n)|)\\
&+ l_1(x_1, x_2)\{H_{n1}^{*-}(x_1) - x_1\} + l_2(x_1, x_2)\{H_{n2}^{*-}(x_2) - x_2\}\\
&- l_1(x_1, x_2)\{H_{n1}^{-}(x_1) - x_1\} - l_2(x_1, x_2)\{H_{n2}^{-}(x_2) - x_2\}\\
&- l_1(x_1, x_2)\{H_{n1}^{'-}(x_1) - x_1\} - l_2(x_1, x_2)\{H_{n2}^{'-}(x_2) - x_2\} \quad \text{a.s.}\\
={}& O(m^{-3/5}(\log n)^{3/5}) + O(|A(m/n)|)\\
&+ l_1(x_1, x_2)\{H_{n1}^{*-}(x_1) - H_{n1}^{-}(x_1) - H_{n1}^{'-}(x_1) + x_1\}\\
&+ l_2(x_1, x_2)\{H_{n2}^{*-}(x_2) - H_{n2}^{-}(x_2) - H_{n2}^{'-}(x_2) + x_2\}\\
={}& O(m^{-3/5}(\log n)^{3/5}) + O(|A(m/n)|) \quad \text{a.s.}
\end{aligned}
$$

uniformly for $(x_1, x_2) \in B$, i.e., the theorem holds.

Acknowledgments

We thank two reviewers for their helpful comments. Peng's research was supported by NSF grant SES 0631608 and Qi's research was supported by NSF grant DMS 0604176.

Appendix

Before proving (4.5)–(4.8) used in the proof of Theorem 2.1, we need some lemmas.

Let $\mathbf{Y}_{n1}, \mathbf{Y}_{n2}, \ldots, \mathbf{Y}_{nn}$ be independent bivariate random vectors taking values in $(0, 1]^2$ with a common distribution function. For any Borel subset B in $[0, 1]^2$ set

$$C_n(B) = P(\mathbf{Y}_{n1} \in B), \qquad \hat{C}_n(B) = \frac{1}{n}\sum_{i=1}^{n} I(\mathbf{Y}_{ni} \in B).$$

In particular, for $\mathbf{x} \in (0, 1]^2$ define

$$\lambda_n(\mathbf{x}) = \hat{C}_n((\mathbf{0}, \mathbf{x}]) - C_n((\mathbf{0}, \mathbf{x}]),$$

and for $x \in (0, 1]$ set

$$\beta_n(x) = C_n((0, x] \times (0, 1]), \qquad \rho_n(x) = C_n((0, 1] \times (0, x]).$$

Lemma 4.1. *Let $\upsilon_n = C_n(A_n) \le 1/2$, where $A_n \subset [0, 1]^2$ are half-open rectangles. Then*

$$\sup_{\widetilde{A} \subset A_n} |\hat{C}_n(\widetilde{A}) - C_n(\widetilde{A})| = O\left(\sqrt{\frac{\upsilon_n \log n}{n}}\right) \quad \text{a.s.}$$

if $\log n = o(n\upsilon_n)$*, where $\widetilde{A}$ denotes any half-open rectangle.*

Proof. It follows from inequality 2.5 in [6] that

$$P\left(\sup_{\widetilde{A}\subset A}|n^{\frac{1}{2}}(\hat{C}_n(\widetilde{A})-C_n(\widetilde{A}))|\geq t\right)\leq K\exp\left\{\frac{-(1-\delta)t^2}{2C_n(A)}\psi\left(\frac{t}{n^{\frac{1}{2}}C_n(A)}\right)\right\},\quad t>0,\tag{4.11}$$

where

$$\psi(t)=2t^{-2}((1+t)\log(1+t)-t).\tag{4.12}$$

Hence, the lemma follows from (4.11) and the Borel–Cantelli lemma. □

Lemma 4.2. *Assume* $\varepsilon_n>0$ *are constants such that* $\log n=o(n\varepsilon_n)$, $\varepsilon_n=o(1)$ *and*

$$\sup_{0\leq x\leq 1}\max\{\beta_n(x)-\beta_n(x-),\rho_n(x)-\rho_n(x-)\}=o(\varepsilon_n).$$

Then

$$\sup_{\{(a,b,x,y):|\beta_n(y)-\beta_n(x)|\leq\varepsilon_n\}}|\hat{C}_n((x,y]\times(a,b])-C_n((x,y]\times(a,b])|=O\left(\sqrt{\frac{\varepsilon_n\log n}{n}}\right)$$

a.s.

and

$$\sup_{\{(a,b,x,y):|\rho_n(y)-\rho_n(x)|\leq\varepsilon_n\}}|\hat{C}_n((a,b]\times(x,y])-C_n((a,b]\times(x,y])|=O\left(\sqrt{\frac{\varepsilon_n\log n}{n}}\right)$$

a.s.

Proof. By the given conditions, for any large n there exist points $0=x_0<x_1<x_2<\cdots<x_{q(n)}<x_{q(n)+1}=1$ such that $2\varepsilon_n<\beta_n(x_{i+2})-\beta_n(x_i)<4\varepsilon_n$ for $0\leq i\leq q(n)-1$, where $q(n)\leq n/\varepsilon_n\leq n^2/\log n$, and for any rectangle $(x,y]\times(a,b]$, if $|\beta_n(y)-\beta_n(x)|\leq\varepsilon_n$, it must be in some rectangle $(x_i,x_{i+1}]\times(0,1]$. For these rectangles we can apply inequality (4.11) and use the Borel–Cantelli lemma as in the proof of Lemma 4.1 to get the first equation. Since $\psi(t)$ in (4.12) is decreasing, by using (4.11) with $\delta=1/2$ we have

$$\begin{aligned}
&P(\sup_{\{(a,b,x,y):|\beta_n(y)-\beta_n(x)|\leq\varepsilon_n\}}|\hat{C}_n((x,y]\times(a,b])-C_n((x,y]\times(a,b])|>s)\\
&\leq\sum_{i=0}^{q(n)-1}P(\sup_{0\leq a<b\leq 1,x_i\leq x<y\leq x_{i+1}}|\hat{C}_n((x,y]\times(a,b])-C_n((x,y]\times(a,b])|>s)\\
&\leq\sum_{i=0}^{q(n)-1}K\exp\left\{\frac{-ns^2}{4(\beta_n(x_{i+2})-\beta_n(x_i))}\psi\left(\frac{s}{\beta_n(x_{i+2})-\beta_n(x_i)}\right)\right\}\\
&\leq\frac{Kn^2}{\log n}\exp\left\{\frac{-ns^2}{16\varepsilon_n}\psi\left(\frac{s}{\varepsilon_n}\right)\right\},
\end{aligned}$$

which is dominated by $Kn^{-2}(\log n)^{-1}$, where K is a positive constant, if $s=s_n=(128\varepsilon_n\log n)^{1/2}n^{-1}$. Hence, the first equation follows from the Borel–Cantelli lemma. Similarly, we can show the second equation. □

Lemma 4.3. *If $m/\log n \to \infty$ as $n \to \infty$, then*

$$P\left(\max_{1\le i\le k_n}\max_{D\in\mathcal{F}_n^{(i)}} 2^{i/2}|P_n(D)-P(D)| > s_n \text{ infinitely often}\right)=0, \tag{4.13}$$

where $s_n = 6(m\log n)^{1/2}n^{-1}$.

Proof. It follows from Bernstein's inequality and (4.2) that

$$\begin{aligned}
&P\left(\max_{1\le i\le k_n}\max_{D\in\mathcal{F}_n^{(i)}} 2^{i/2}|P_n(D)-P(D)| > s_n\right)\\
&\le \sum_{i=1}^{k_n}\sum_{r_1=1}^{2^i}\sum_{r_2=1}^{2^i} P(|P_n(J_{i,r_1}\times J_{i,r_2})-P(J_{i,r_1}\times J_{i,r_2})|>s_n 2^{-i/2})\\
&\le \sum_{i=1}^{k_n}\sum_{r_1=1}^{2^i}\sum_{r_2=1}^{2^i} 2\exp\left\{-\frac{1}{2}ns_n^2 2^{-i}(s_n 2^{-i/2}+P(J_{i,r_1}\times J_{i,r_2})(1-P(J_{i,r_1}\times J_{i,r_2})))^{-1}\right\}\\
&\le \sum_{i=1}^{k_n} 2^{2i+1}\exp\left\{-\frac{1}{2}\frac{ns_n^2}{s_n 2^{i/2}+2t_n}\right\}\\
&\le \sum_{i=1}^{k_n} 2^{2i+1}\exp\left\{-\frac{1}{2}\frac{ns_n^2}{s_n 2^{k_n/2}+2t_n}\right\}\\
&= O\left(2^{2k_n}\exp\left\{-\frac{1}{2}\frac{nt^2}{s_n 2^{k_n/2}+2t_n}\right\}\right)\\
&= O\left(\left(\frac{m}{\log n}\right)^{6/5}\exp\left\{-\frac{1}{2}\frac{ns_n^2}{3t_n}\right\}\right)\\
&= O(n^{-2})
\end{aligned}$$

as n is large enough. Thus, (4.13) follows by the Borel–Cantelli lemma. □

Let A_n be the collection of all the vertices of the rectangles in $\mathcal{F}_n$ and define $C_{\mathbf{x}} = [0, x_1]\times[0, x_2]$ for $\mathbf{x} = (x_1, x_2)^{\mathrm{T}}$. For each $\mathbf{x}\in[0, 2t_n]^2$, define

$$\underline{\mathbf{x}} = \max\{\mathbf{y}\in A_n : \mathbf{y}\le\mathbf{x}\},$$

where supremum is taken coordinatewise. Then we have $C_{\underline{\mathbf{x}}}\subset C_{\mathbf{x}}$.

Define for any set $D\subset[0,1]^2$

$$P_n'(D)=\frac{1}{n}\sum_{i=1}^n I(\mathbf{Z}_i'\in D)\quad\text{and}\quad P_n^*(D)=\frac{1}{n}\sum_{i=1}^n I(\mathbf{Z}_i^*\in D).$$

Lemma 4.4. *If $m/\log n\to\infty$ as $n\to\infty$, then*

$$\max_{\mathbf{x}\in A_n}|P_n^*(C_{\mathbf{x}})-P_n(C_{\mathbf{x}})-P_n'(C_{\mathbf{x}})+P(C_{\mathbf{x}})| = O\left(\frac{m^{2/5}(\log n)^{3/5}}{n}\right) \tag{4.14}$$

a.s. under P^ω, which implies

$$\max_{\mathbf{x}\in A_n} |P_n^*(C_\mathbf{x}) - P_n(C_\mathbf{x}) - P_n'(C_\mathbf{x}) + P(C_\mathbf{x})| = O(m^{2/5}(\log n)^{3/5}n^{-1}) \tag{4.15}$$

a.s. under P.

Proof. Note that for each $\mathbf{x} \in A_n$, $C_\mathbf{x}$ can be written as a union of the unions of at most 2×2^i sets in $\mathcal{F}_n^{(i)}$ for $1 \le i \le k_n$. Hence, by (4.13) and (4.2)

$$\begin{aligned}
&\sup_{\mathbf{x}\in A_n} E^\omega(|I(\mathbf{Z}_1^* \in C_\mathbf{x}) - I(\mathbf{Z}_1' \in C_\mathbf{x})|) \\
&\quad \le \sum_{i=1}^{k_n} 2^{i+1}2^{-i/2} \max_{D\in\mathcal{F}_n^{(i)}} 2^{i/2} E^\omega(|I(\mathbf{Z}_1^* \in D) - I(\mathbf{Z}_1' \in D)|) \\
&\quad = \sum_{i=1}^{k_n} 2^{i/2+1} \max_{D\in\mathcal{F}_n^{(i)}} 2^{i/2}|P_n(D) - P(D)| \\
&\quad = O(m^{4/5}(\log n)^{1/5}n^{-1}) \quad \text{a.s.}
\end{aligned} \tag{4.16}$$

It follows from Bernstein's inequality that

$$\begin{aligned}
&P^\omega\left(\max_{\mathbf{x}\in A_n} |P_n^*(C_\mathbf{x}) - P_n(C_\mathbf{x}) - P_n'(C_\mathbf{x}) + P(C_\mathbf{x})| > t\right) \\
&\quad \le \sum_{i=1}^{k_n} 2^{2i} \max_{\mathbf{x}\in A_n} P^\omega(|P_n^*(C_\mathbf{x}) - P_n(C_\mathbf{x}) - P_n'(C_\mathbf{x}) + P(C_\mathbf{x})| > t) \\
&\quad \le \sum_{i=1}^{k_n} 2^{2i} \max_{\mathbf{x}\in A_n} 2\exp\left\{-\frac{1}{2}\frac{nt^2}{t + E^\omega(|I(\mathbf{Z}_1^* \in C_\mathbf{x}) - I(\mathbf{Z}_1' \in C_\mathbf{x})|)}\right\}.
\end{aligned} \tag{4.17}$$

Let $t = Cm^{2/5}(\log n)^{3/5}n^{-1}$. From (4.16), C can be chosen to be a large number depending on the sample path ω such that

$$nt^2/(t + E^\omega(|I(\mathbf{Z}_1^* \in C_\mathbf{x}))) \ge 6\log n$$

for all large n. Then we have from (4.17) that

$$\begin{aligned}
&P^\omega\left(\max_{\mathbf{x}\in A_n} |P_n^*(C_\mathbf{x}) - P_n(C_\mathbf{x}) - P_n'(C_\mathbf{x}) + P(C_\mathbf{x})| > Cm^{2/5}(\log n)^{3/5}n^{-1}\right) \\
&\quad = O(2^{2k_n+1}n^{-6}) = O(n^{-2})
\end{aligned}$$

which, together with the Borel–Cantelli lemma, yields (4.14). □

Lemma 4.5. *Let $\delta_\mathbf{x} = C_\mathbf{x} \setminus C_{\underline{\mathbf{x}}}$. If $m/\log n \to \infty$ as $n \to \infty$, then*

$$\max_{\mathbf{x}\in[0,2t_n]^2} |P_n(\delta_\mathbf{x}) - P(\delta_\mathbf{x})| = O\left(\frac{m^{1/5}(\log n)^{4/5}}{n}\right) \quad \textit{a.s.}, \tag{4.18}$$

$$\max_{\mathbf{x}\in[0,2t_n]^2} |P_n'(\delta_\mathbf{x}) - P(\delta_\mathbf{x})| = O\left(\frac{m^{1/5}(\log n)^{4/5}}{n}\right) \quad \textit{a.s.} \tag{4.19}$$

and

$$\max_{\mathbf{x}\in[0,2t_n]^2} |P_n^*(\delta_{\mathbf{x}}) - P_n(\delta_{\mathbf{x}})| = O\left(\frac{m^{1/5}(\log n)^{4/5}}{n}\right) \tag{4.20}$$

a.s. under P^ω (also under P). Therefore,

$$\sup_{\mathbf{x}\in[0,2t_n]^2} |P_n^*(\delta_{\mathbf{x}}) - P_n'(\delta_{\mathbf{x}}) - P_n(\delta_{\mathbf{x}}) + P(\delta_{\mathbf{x}})| = O\left(\frac{m^{2/5}(\log n)^{3/5}}{n}\right) \quad \textit{a.s.} \tag{4.21}$$

Proof. For each $\mathbf{x} = (x, y) \in [0, 2t_n]^2$, let $\underline{\mathbf{x}} = (\underline{x}, \underline{y})$. Then

$$\delta_{\mathbf{x}} = ((\underline{x}, x] \times (0, y]) \cup ((0, \underline{x}] \times (\underline{y}, y]).$$

By definition, $x - \underline{x} \le t_n 2^{-k_n} = 2m^{2/5}(\log n)^{3/5}n^{-1}$ and $y - \underline{y} \le 2m^{2/5}(\log n)^{3/5}n^{-1}$. By applying Lemma 4.2 with $\varepsilon_n = 2m^{2/5}(\log n)^{3/5}n^{-1}$ we have

$$\begin{aligned}
&\max_{\mathbf{x}\in[0,2t_n]^2} |P_n(\delta_{\mathbf{x}}) - P(\delta_{\mathbf{x}})| \\
&\quad\le \sup_{\{(a,b,x,y):|x-y|\le\varepsilon_n\}} |P_n((x, y] \times (a, b]) - P((x, y] \times (a, b])| \\
&\qquad + \sup_{\{(a,b,x,y):|x-y|\le\varepsilon_n\}} |P_n((a, b] \times (x, y]) - P((a, b] \times (x, y])| \\
&\quad= O\left(\frac{m^{1/5}(\log n)^{4/5}}{n}\right) \quad \text{a.s.,}
\end{aligned}$$

proving (4.18). (4.19) holds trivially since $\{P_n'\}$ is an independent copy of $\{P_n\}$.

From the above inequality we conclude that

$$\begin{aligned}
&\sup_{\{(x<y):y-x\le\varepsilon_n\}} \{|P_n((x, y] \times (0, 1]) - (y - x)| \\
&\quad + |P_n((0, 1] \times (x, y]) - (y - x)|\} = O\left(\frac{m^{1/5}(\log n)^{4/5}}{n}\right) \quad \text{a.s.,}
\end{aligned}$$

which implies

$$\begin{aligned}
&\sup_{\{(x<y):y-x\le\varepsilon_n\}} \{|P_n((x, y] \times (0, 1]) + P_n((0, 1] \times (x, y])\} \\
&\quad\le \frac{m^{2/5}(\log n)^{3/5}}{n}(1 + o(1)) \text{ a.s.}
\end{aligned}$$

By using Lemma 4.2 again we get (4.20) under P^ω.

The last equation follows from Eqs. (4.18)–(4.20) and the estimate

$$\max_{\mathbf{x}\in[0,2t_n]^2} P(\delta_{\mathbf{x}}) \le \frac{4m^{2/5}(\log n)^{3/5}}{n}.$$

This completes the proof. □

Lemma 4.6. *If $m/\log n \to \infty$ as $n \to \infty$, then*

$$\sup_{\mathbf{x}\in[0,2t_n]^2} |P_n^*(C_{\mathbf{x}}) - P_n'(C_{\mathbf{x}}) - P_n(C_{\mathbf{x}}) + P(C_{\mathbf{x}})| = O\left(\frac{m^{2/5}(\log n)^{3/5}}{n}\right) \quad \textit{a.s.} \tag{4.22}$$

Proof. Note that for $\mathbf{x} \in [0, 2t_n]^2$

$$C_{\mathbf{x}} = C_{\underline{\mathbf{x}}} \cup \delta_{\mathbf{x}}.$$

The lemma follows from the inequality

$$\begin{aligned}
&\sup_{\mathbf{x}\in[0,2t_n]^2} |P_n^*(C_{\mathbf{x}}) - P_n'(C_{\mathbf{x}}) - P_n(C_{\mathbf{x}}) + P(C_{\mathbf{x}})| \\
&\quad \le \sup_{\mathbf{x}\in[0,2t_n]^2} |P_n^*(C_{\underline{\mathbf{x}}}) - P_n'(C_{\underline{\mathbf{x}}}) - P_n(C_{\underline{\mathbf{x}}}) + P(C_{\underline{\mathbf{x}}})| \\
&\qquad + \sup_{\mathbf{x}\in[0,2t_n]^2} |P_n^*(\delta_{\mathbf{x}}) - P'(\delta_{\mathbf{x}}) - P_n(\delta_{\mathbf{x}}) + P(\delta_{\mathbf{x}})|
\end{aligned}$$

and Lemmas 4.4 and 4.5. □

Lemma 4.7. *For $x \in (0, 1]$ define $M_x = (0, x] \times (0, 1]$ and $N_x = (0, 1] \times (0, x]$. If $m/\log n \to \infty$ as $n \to \infty$, then*

$$\sup_{x\in(0,2t_n]} |P_n^*(M_x) - P_n'(M_x) - P_n(M_x) + P(M_x)| = O\left(\frac{m^{2/5}(\log n)^{3/5}}{n}\right) \quad \textit{a.s.} \tag{4.23}$$

and

$$\sup_{x\in(0,2t_n]} |P_n^*(N_x) - P_n'(N_x) - P_n(N_x) + P(N_x)| = O\left(\frac{m^{2/5}(\log n)^{3/5}}{n}\right) \quad \textit{a.s.} \tag{4.24}$$

Proof. Set $A_n' = \left\{\frac{2t_n i}{2^{k_n}} : 0 \le i \le 2^{k_n}\right\}$. For each $x \in (0, 2t_n]$ define $\underline{x} = \{y \in A_n' : y \le x\}$. Set $\delta_x' = M_x \setminus M_{\underline{x}}$. In the same manner as we obtained (4.21) in Lemma 4.5 we can show that

$$\sup_{x\in(0,2t_n]} |P_n^*(\delta_x') - P_n'(\delta_x') - P_n(\delta_x') + P(\delta_x')| = O\left(\frac{m^{2/5}(\log n)^{3/5}}{n}\right) \quad \text{a.s.}$$

By using the same reasoning as in the proof of Lemma 4.6 we obtain (4.23). Likewise, we can prove (4.24). □

Proof of (4.5). We only need to consider the case $j = 1$. By applying Lemma 4.1 to $\{\mathbf{Z}_i, 1 \le i \le n\}$ with $A_n = (0, 3t_n] \times (0, 1]$ and $\tilde{A} = (0, xt_n] \times (0, 1]$ we have

$$\sup_{0<x\le 3} |H_{n1}(x) - x| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.}, \tag{4.25}$$

which implies $\sup_{0<x\le 2} H_{n1}^-(x) \le H_{n1}^-(2) \le 3$ for all large n. Therefore, it follows from (4.25) that

$$\sup_{0<x\le 2} |H_{n1}(H_{n1}^-(x)) - H_{n1}^-(x)| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.},$$

which, together with the fact that $|H_{n1}(H_{n1}^-(x)) - x| \le m^{-1}$, yields

$$\sup_{0<x\le 2} |H_{n1}^-(x) - x| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.} \tag{4.26}$$

This is the first statement in (4.5). The second statement can be proved in the same way. For the third statement, we can apply Lemma 4.1 to the bootstrap sample $\{\mathbf{Z}_i^*, 1 \le i \le n\}$ and get

$$\sup_{0<x\le 3} |H_{n1}^*(x) - H_{n1}(x)| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.}, \tag{4.27}$$

which coupled with (4.25) yields

$$\sup_{0<x\le 3} |H_{n1}^*(x) - x| = O(m^{-1/2}(\log n)^{1/2}) \quad \text{a.s.} \tag{4.28}$$

Then, the third statement of (4.5) is obtained by using the same argument in the proof of (4.26). Hence (4.5) is proved.

Proof of (4.6). We only need to consider the case $j = 1$. It follows from Lemma 4.2 that

$$\sup_{\{(x_1,x_2):|\beta_n(x_1)-\beta_n(x_2)|\le\varepsilon_n\}} |\lambda_n(x_1, 1) - \lambda_n(x_2, 1)| = O\left(\sqrt{\frac{\varepsilon_n \log n}{n}}\right) \quad \text{a.s.} \tag{4.29}$$

Hence, we will apply (4.29) to $\{\mathbf{Z}'_j, 1 \le i \le n\}$, $\{\mathbf{Z}^*_j, 1 \le i \le n\}$, and $\{\mathbf{Z}_j, 1 \le i \le n\}$, respectively. For all the three cases below we need to verify that $\beta_n(x_1) - \beta_n(x_2) = O((m \log n)^{1/2} n^{-1})$ a.s., which can be concluded from (4.5) and (4.25)–(4.27).

Note that for the sample $\{\mathbf{Z}'_j, 1 \le i \le n\}$, we have

$$\beta_n(t_n x) = t_n H'_{n1}(x), \qquad \lambda_n(t_n x, 1) = t_n(H'_{n1}(x) - x)$$

and set $x_1 = t_n x$ and $x_2 = t_n H'^{-}_{n1}(x)$. From (4.29) with $\varepsilon_n = O((m \log n)^{1/2} n^{-1})$ we have

$$\sup_{0<x\le 2} |H'_{n1}(x) - x - (H'_{n1}(H'^{-}_{n1}(x)) - H'^{-}_{n1}(x))| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.},$$

which implies the first part in (4.6) by noting that $|H'_{n1}(H'^{-}_{n1}(x)) - x| \le m^{-1}$. Similarly, for the bootstrap sample $\{\mathbf{Z}^*_j, 1 \le i \le n\}$ we have

$$\beta_n(t_n x) = t_n H^*_{n1}(x), \quad \lambda_n(t_n x, 1) = t_n(H^*_{n1}(x) - H_{n1}(x)),$$

and set $x_1 = t_n H_n^-(x)$ and $x_2 = t_n H^{*-}_{n1}(x)$. Hence

$$\begin{aligned} &\sup_{0<x\le 2} |H^*_{n1}(x) - H_{n1}(x) - (H^*_{n1}(H^{*-}_{n1}(x)) - H_{n1}(H^{*-}_{n1}(x)))| \\ &\quad = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.}, \end{aligned}$$

i.e.,

$$\sup_{0<x\le 2} |H^*_{n1}(x) - H_{n1}(x) - x + H_{n1}(H^{*-}_{n1}(x))| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \tag{4.30}$$

Applying (4.29) to $\{\mathbf{Z}_j, 1 \le i \le n\}$ with choices $x_1 = t_n H_n^-(x)$ and $x_2 = t_n H^{*-}_{n1}(x)$, we have

$$\sup_{0<x\le 2} |x - H^{-}_{n1}(x) - H_{n1}(H^{*-}_{n1}(x)) + H^{*-}_{n1}(x)| = O(m^{-3/4}(\log n)^{3/4}) \quad \text{a.s.},$$

which coupled with (4.30) proves the second part of (4.6).

Proof of (4.7). It follows from Lemma 4.2 that

$$\sup_{\{|\beta_n(x_1)-\beta_n(x_2)|\le\varepsilon_n,|\rho_n(y_1)-\rho_n(y_2)|\le\varepsilon_n\}} |\lambda_n(x_1,y_1)-\lambda_n(x_2,y_2)| = O\left(\sqrt{\frac{\varepsilon_n\log n}{n}}\right) \quad \text{a.s.} \tag{4.31}$$

Hence, (4.7) can be shown straightforwardly by (4.31).

Proof of (4.8). Similar to the proof of (4.7).

References

[1] K. Chen, S.H. Lo, On a mapping approach to investigating the bootstrap accuracy, Probab. Theory Related Fields 107 (1997) 197–217.
[2] J. Danielsson, L. de Haan, L. Peng, C.G. de Vries, Using a bootstrap method to choose the sample fraction in tail index estimation, J. Multivariate Anal. 76 (2001) 226–248.
[3] H. Drees, X. Huang, Best attainable rates of convergence for estimates of the stable tail dependence functions, J. Multivariate Anal. 64 (1998) 25–47.
[4] J.H.J. Einmahl, L. de Haan, D. Li, Weighted approximations of tail copula processes with application to testing the multivariate extreme value distribution, Ann. Statist. 34 (4) (2006) 1987–2014.
[5] J.H.J. Einmahl, D.M. Mason, Strong limit theorems for weighted quantile process, Ann. Probab. 16 (1988) 1623–1643.
[6] J.H.J. Einmahl, F.H. Ruymgaart, The almost sure behavior of the oscillation modulus of the multivariate empirical process, Statist. Probab. Lett. 6 (1987) 87–96.
[7] C. El-Nouty, A. Guillou, On the bootstrap accuracy of the Pareto index, Statist. Decisions 18 (3) (2000) 275–289.
[8] J. Geluk, L. de Haan, On bootstrap sample size in extreme value theory, Publ. de l'Inst. Math. Nouvelle serie 71 (85) (2002) 21–25.
[9] L. de Haan, J. de Ronde, Sea and wind: Multivariate extremes at work, Extremes 1 (1998) 7–45.
[10] L. de Haan, A.K. Sinha, Estimating the probability of a rare event, Ann. Statist. 27 (1999) 732–759.
[11] P. Hall, Using the bootstrap to estimate mean squared error and selecting parameter in nonparametric problem, J. Multivariate Anal. 32 (1990) 177–203.
[12] X. Huang, Statistics of Bivariate Extreme Values, Ph.D. Thesis, Tinbergen Institute Research Series 1992.
[13] C. Klüppelberg, G. Kuhn, L. Peng, Estimating the tail dependence of an elliptical distribution, Bernoulli 13 (1) (2007) 229–251.
[14] A.J. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management, Princeton University Press, 2005.
[15] L. Peng, Y. Qi, Partial derivatives and confidence intervals for bivariate tail dependence functions, J. Statist. Plann. Inference 137 (2007) 2089–2101.
[16] Y. Qi, Almost sure convergence of the stable tail empirical dependence function in multivariate extreme statistics, Acta Math. Appl. Sin. Engl. Ser. 13 (2) (1997) 167–175.
[17] R. Schmidt, U. Stadtmüller, Nonparametric estimation of tail dependence, Scand. J. Statist. 33 (2006) 307–335.