NONPARAMETRIC ESTIMATION OF THE SPECTRAL MEASURE OF AN EXTREME VALUE DISTRIBUTION

By John H. J. Einmahl,[1] Laurens de Haan[2] and Vladimir I. Piterbarg[3]

*Tilburg University, Erasmus University and
Moscow Lomonosov State University*

Let $(\mathscr{X}_1, \mathscr{Y}_1), \ldots, (\mathscr{X}_n, \mathscr{Y}_n)$ be a random sample from a bivariate distribution function F in the domain of max-attraction of a distribution function G. This G is characterised by the two extreme value indices and its spectral or angular measure. The extreme value indices determine both the marginals and the spectral measure determines the dependence structure of G. One of the main issues in multivariate extreme value theory is the estimation of this spectral measure. We construct a truly nonparametric estimator of the spectral measure, based on the ranks of the above data. Under natural conditions we prove consistency and asymptotic normality for the estimator. In particular, the result is valid for all values of the extreme value indices. The theory of (local) empirical processes is indispensable here. The results are illustrated by an application to real data and a small simulation study.

1. Introduction. In two-dimensional space as in one-dimensional space, if one has to do inference in the tail of a distribution outside the range of the observations, a rational way to proceed is to use extreme value theory, that is, to model the tail *asymptotically* as an extreme-value distribution. In order to turn this into a useful tool, one has to estimate the parameters of the fitted extreme-value distribution. In fact there is no finite-dimensional parametrisation in the higher-dimensional case: the probability distribution is characterised by the extreme value indices and a finite measure, the *spectral* or *angular* measure. The estimation of this spectral measure is one of the main issues in multivariate extreme value theory. In this paper a natural nonparametric estimator is constructed and its asymptotic properties are derived. In order to describe the setup, we have to start by explaining the probabilistic background.

Let $(\mathscr{X}, \mathscr{Y}), (\mathscr{X}_1, \mathscr{Y}_1), (\mathscr{X}_2, \mathscr{Y}_2), \ldots, (\mathscr{X}_n, \mathscr{Y}_n)$ be i.i.d. with common continuous distribution function F. Suppose that there are norming constants $a_n, c_n >$

0 and b_n, d_n such that the sequence of distribution functions

$$\mathbf{P}\left\{\frac{\max_{1\leq i\leq n}\mathscr{X}_i - b_n}{a_n} \leq x, \frac{\max_{1\leq i\leq n}\mathscr{Y}_i - d_n}{c_n} \leq y\right\}$$

converges to a limit distribution function, say $G(x, y)$, with non-degenerate marginals, that is,

$$(1)\qquad\qquad \lim_{n\to\infty} F^n(a_n x + b_n, c_n y + d_n) = G(x, y)$$

for all but countably many x and y. The two marginal distribution functions are automatically extreme value distribution functions and we choose the constants a_n, c_n, b_n and d_n such that for some $\gamma_1, \gamma_2 \in \mathrm{R}$,

$$G(x, \infty) = \exp\left\{-(1 + \gamma_1 x)^{-1/\gamma_1}\right\},$$

$$G(\infty, y) = \exp\left\{-(1 + \gamma_2 y)^{-1/\gamma_2}\right\}.$$

Then there is a finite measure Φ on $[0, \pi/2]$, the spectral measure, such that

$$(2)\quad G\left(\frac{x^{\gamma_1} - 1}{\gamma_1}, \frac{y^{\gamma_2} - 1}{\gamma_2}\right) = \exp\left\{-\int_0^{\pi/2}\left(\frac{1 \wedge \tan\theta}{x} \vee \frac{1 \wedge \cot\theta}{y}\right)\Phi(d\theta)\right\}$$

and

$$\int_0^{\pi/2}(1 \wedge \tan\theta)\,\Phi(d\theta) = \int_0^{\pi/2}(1 \wedge \cot\theta)\,\Phi(d\theta) = 1.$$

This is a variant, useful for our purposes, of the usual representation; cf. de Haan and Resnick (1977), Deheuvels (1978) and Pickands (1981). For more background material see Einmahl, de Haan and Sinha (1997). Note that G has independent marginals if and only if Φ is concentrated on $\{0, \pi/2\}$.

An alternative useful way to express (1) is

$$\lim_{n\to\infty} n(1 - F(a_n x + b_n, c_n y + d_n)) = -\log G(x, y).$$

A continuous version also holds:

$$(3)\qquad\qquad \lim_{t\to\infty} t(1 - F(a(t)x + b(t), c(t)y + d(t))) = -\log G(x, y)$$

for suitable functions $a, c > 0$, and b and d, or

$$\lim_{t\to\infty} t\mathbf{P}\left\{\frac{\mathscr{X} - b(t)}{a(t)} > \frac{x^{\gamma_1} - 1}{\gamma_1} \text{ or } \frac{\mathscr{Y} - d(t)}{c(t)} > \frac{y^{\gamma_2} - 1}{\gamma_2}\right\}$$

$$= -\log G\left(\frac{x^{\gamma_1} - 1}{\gamma_1}, \frac{y^{\gamma_2} - 1}{\gamma_2}\right)$$

$$= \int_0^{\pi/2}\left(\frac{1 \wedge \tan\theta}{x} \vee \frac{1 \wedge \cot\theta}{y}\right)\Phi(d\theta)$$

for $x, y > 0$, where we can choose $b(t) = F_1^{\leftarrow}(1 - 1/t)$ and $d(t) = F_2^{\leftarrow}(1 - 1/t)$, with F_1 and F_2 the marginals of F. This implies

(4)
$$\lim_{t \to \infty} \mathbf{P}\left\{\frac{\mathscr{X} - b(t)}{a(t)} > \frac{x^{\gamma_1} - 1}{\gamma_1} \text{ or } \frac{\mathscr{Y} - d(t)}{c(t)} > \frac{y^{\gamma_2} - 1}{\gamma_2} \,\middle|\, \mathscr{X} > b(t) \text{ or } \mathscr{Y} > d(t)\right\}$$
$$= \int_0^{\pi/2} \left(\frac{1 \wedge \tan\theta}{x} \vee \frac{1 \wedge \cot\theta}{y}\right) \Phi(d\theta)/\Phi\left(\left[0, \frac{\pi}{2}\right]\right).$$

Relation (4) has an interpretation analogous to the Generalised Pareto setup in one-dimensional extreme value theory: observations outside a large rectangle $(-\infty, b(t)] \times (-\infty, d(t)]$ can be considered as i.i.d. random variables with approximate distribution function

$$1 - \int_0^{\pi/2} \left(\frac{1 \wedge \tan\theta}{(1 + \gamma_1 x/a(t))^{1/\gamma_1}} \vee \frac{1 \wedge \cot\theta}{(1 + \gamma_2 y/c(t))^{1/\gamma_2}}\right) \Phi(d\theta)/\Phi\left(\left[0, \frac{\pi}{2}\right]\right).$$

This interpretation is the basis for estimating Φ.

Relation (3) becomes simpler if we apply a preliminary transformation to the marginals:

(5)
$$\lim_{t \downarrow 0} t^{-1} \mathbf{P}\left\{1 - F_1(\mathscr{X}) \leq tx \text{ or } 1 - F_2(\mathscr{Y}) \leq ty\right\}$$
$$= \int_0^{\pi/2} \left(\frac{x}{1 \vee \cot\theta} \vee \frac{y}{1 \vee \tan\theta}\right) \Phi(d\theta),$$

[cf. de Haan and Resnick (1977)] or, more generally, for any Borel set A in $[0, \infty]^2 \setminus \{(\infty, \infty)\}$,

(6)
$$\lim_{t \downarrow 0} t^{-1} \mathbf{P}\left\{(1 - F_1(\mathscr{X}), 1 - F_2(\mathscr{Y})) \in tA\right\} = \Lambda(A),$$

(with $tA = \{(tx, ty) : (x, y) \in A\}$) provided $\Lambda(\partial A) = 0$ with the measure Λ on $[0, \infty]^2 \setminus \{(\infty, \infty)\}$ defined by

(7)
$$\Lambda\left(([x, \infty] \times [y, \infty])^c\right) = \int_0^{\pi/2} \left(\frac{x}{1 \vee \cot\theta} \vee \frac{y}{1 \vee \tan\theta}\right) \Phi(d\theta).$$

Or, with P the measure on $[0, 1]^2$ induced by $(X, Y) := (1 - F_1(\mathscr{X}), 1 - F_2(\mathscr{Y}))$,

(8)
$$\lim_{t \downarrow 0} t^{-1} P(tA) = \Lambda(A).$$

These relations show how one can get Λ from F and hence show a way to estimate Λ. A slightly more complicated relation shows how to get Φ from F: apply (6) to the set

$$C_\theta := \left\{(x, y) \in [0, \infty]^2 \,:\, x \wedge y \leq 1, \ y \leq x \tan\theta\right\},$$

with the convention that $\infty \cdot 0 = 1$. The result is

(9)
$$\lim_{t \downarrow 0} t^{-1} \mathbf{P}\{(1 - F_1(\mathscr{X})) \wedge (1 - F_2(\mathscr{Y})) \leq t, 1 - F_2(\mathscr{Y}) \leq (1 - F_1(\mathscr{X})) \tan\theta\}$$
$$= \lim_{t \downarrow 0} t^{-1} \mathbf{P}\{X \wedge Y \leq t, Y \leq X \tan\theta\} = \Lambda(C_\theta) = \Phi(\theta)$$

for all but countably many θ. Note that we simplified the notation: $\Phi(\theta) :=$ $\Phi([0, \theta])$. It is also useful to note that we have, for example, on $[0, \pi/4]$, for continuity points of Φ:

$$(10) \qquad \Phi(\theta) = \frac{\partial}{\partial y} \Lambda \left(([x, \infty] \times [y, \infty])^c \right) |_{y=x \tan \theta}.$$

In order to turn the left-hand side of (9) into an estimator for Φ, we have to replace F_1 and F_2 and the unknown probability measure P with empirical counterparts. In Einmahl, de Haan and Sinha (1997) this has been done by replacing P with the empirical measure and the tails $1 - F_1(x)$ and $1 - F_2(y)$ with the fitted Pareto tails,

$$t \left(1 + \gamma_1 \frac{x - b(t)}{a(t)} \right)^{-1/\gamma_1} \quad \text{and} \quad t \left(1 + \gamma_2 \frac{y - d(t)}{c(t)} \right)^{-1/\gamma_2}$$

[based on one-dimensional versions of (3)]. The use of Pareto tails (which depend on γ_1 and γ_2) for the marginals and the ensuing necessity to estimate six parameters cause mathematical problems: asymptotic normality was only proved for $\gamma_1, \gamma_2 > 0$; the limiting process does not even exist when $\gamma_1 \leq -1$ or $\gamma_2 \leq -1$.

In this paper we replace P, F_1 and F_2 by the corresponding empirical measures and consider the following purely non-parametric estimator based on the relations (8) and (9):

$$\hat{\Phi}(\theta) := \frac{n}{k} \hat{P}_n \left(\frac{k}{n} C_\theta \right)$$

$$= \frac{1}{k} \sum_{i=1}^{n} \mathbf{1}_{\{(n+1-R_i^{\mathscr{X}}) \wedge (n+1-R_i^{\mathscr{Y}}) \leq k, \, n+1-R_i^{\mathscr{Y}} \leq (n+1-R_i^{\mathscr{X}}) \tan \theta\}}$$

$$= \frac{1}{k} \sum_{i=1}^{n} \mathbf{1}_{\{R_i^{\mathscr{X}} \vee R_i^{\mathscr{Y}} \geq n+1-k, \, n+1-R_i^{\mathscr{Y}} \leq (n+1-R_i^{\mathscr{X}}) \tan \theta\}}$$

where $R_i^{\mathscr{X}}$ is the rank of $\mathscr{X}_i$ among $\mathscr{X}_1, \ldots, \mathscr{X}_n$, $R_i^{\mathscr{Y}}$ is the rank of $\mathscr{Y}_i$ among $\mathscr{Y}_1, \ldots, \mathscr{Y}_n$ and for any Borel set $C \subset [0, 1]^2$,

$$\hat{P}_n(C) := \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_C(\hat{X}_i, \hat{Y}_i),$$

where

$$(\hat{X}_i, \hat{Y}_i) := \frac{1}{n}(n + 1 - R_i^{\mathscr{X}}, n + 1 - R_i^{\mathscr{Y}}).$$

We shall prove that $\hat{\Phi}$ is weakly consistent for Φ provided $k = k(n) \to \infty$, $k(n) = o(n)$, $n \to \infty$, and strongly consistent if moreover $k(n)/\log \log n \to \infty$, $n \to \infty$. We shall give further conditions on Φ and the sequence $k(n)$ that ensure asymptotic normality.

The estimator seems natural, since it is essentially the empirical distribution function. Although the mathematical details of the derivation are delicate,

the asymptotic results are rather simple and valid for *all* $\gamma_j \in R$, $j = 1, 2$. The non-parametric estimator seems to perform well in applications, better than the semi-parametric one described above [cf. de Haan and de Ronde (1998) or the reports on the Neptune project, Draisma et al. (1996, 1997)].

Apart from Einmahl, de Haan and Huang (1993) and Einmahl, de Haan and Sinha (1997), we are not aware of other work on estimating the spectral measure starting from observations in the domain of attraction. There are several proposals for estimating the measure Λ starting from such observations: Huang (1992) [cf. Drees and Huang (1998)], de Haan and Resnick (1993), Abdous, Ghoudi and Khoudraji (1999) and in a restricted parametric context, Tawn (1988), Coles and Tawn (1991), Joe, Smith and Weissman (1992). The paper by Deheuvels and Martynov (1996) considers observations taken from the limit distribution itself.

If one takes any of the mentioned estimators for Λ and one uses it to estimate the extreme-value distribution G via (2) and (7):

$$G\left(\frac{x^{-\gamma_1} - 1}{\gamma_1}, \frac{y^{-\gamma_2} - 1}{\gamma_2}\right) = \exp\left\{-\Lambda\left\{([x, \infty] \times [y, \infty])^c\right\}\right\},$$

this leads to an estimator of G that is itself not necessarily an extreme value distribution (only max-infinitely divisible). If one estimates G via (2) using Φ, one *does* get an extreme value distribution. Observe that formula (10) indicates that the estimation of Φ is like the estimation of Λ, locally.

Apart from this, $\hat{\Phi}$ is directly useful for assessing the amount of dependence in the tail of F; see, for example, Section 4 and Hauksson et al. (2001). In Theorem 4.1 of de Haan and Sinha (1999), $\hat{\Phi}$ is also necessary to estimate the variance in the analysis of probabilities of rare sets in an extreme value context.

The writeup is for the two dimensional situation. The higher dimensional case can be dealt with in a similar way, but the technical details are much more involved.

The results are presented in Section 2. The proof of the main theorem (Theorem 2) is given in Section 3. Section 4 contains an application and a small simulation study.

2. Main results. Our point of departure is now (6) or (8), that is, we consider a probability measure P on $[0, 1]^2$ with distribution function Γ which has uniform-$[0, 1]$ marginals and assume there exists a measure Λ such that

$$\lim_{t\downarrow 0} \frac{1}{t} P(tA) = \Lambda(A)$$

for all measurable $A \subset [0, \infty]^2 \setminus \{(\infty, \infty)\}$ with $\Lambda(\partial A) = 0$. Note that $\Lambda([0, tx] \times [0, ty]) = t\Lambda([0, x] \times [0, y])$ and that $0 \le \Lambda([0, x] \times [0, y]) \le x \wedge y$. Furthermore $\Lambda([0, x] \times [0, \infty]) = \Lambda([0, \infty] \times [0, x]) = x$. Set

$$C_\theta = \left\{(x, y) \in [0, \infty]^2 : x \wedge y \le 1, \; y \le x \tan\theta\right\}, \qquad \theta \in \left[0, \frac{\pi}{2}\right].$$

So we have i.i.d. random vectors $(X_1, Y_1), \ldots, (X_n, Y_n)$, obtained from $(\mathscr{X}_1, \mathscr{Y}_1), \ldots, (\mathscr{X}_n, \mathscr{Y}_n)$ of Section 1, in the following way: $(X_i, Y_i) = (1 - F_1(\mathscr{X}_i), 1 - F_2(\mathscr{Y}_i))$, $i = 1, \ldots, n$. We denote the marginal empirical distribution functions of $(X_1, Y_1), \ldots, (X_n, Y_n)$ with Γ_{1n} and Γ_{2n}, so, for example, $\Gamma_{1n}(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(-\infty, x]}(X_i)$, for $x \le 1$; for $x > 1$ we set $\Gamma_{jn}(x) = x$, for $j = 1, 2$. Now we transform the data by Γ_{1n} and Γ_{2n} as follows: $(\hat{X}_i, \hat{Y}_i) = (\Gamma_{1n}(X_i), \Gamma_{2n}(Y_i))$, $i = 1, \ldots, n$. Observe that the thus obtained data are no longer independent (with respect to i). This dependence is non-negligible and creates a major technical problem. Denote the empirical measures of the (X_i, Y_i) and $(\hat{X}_i, \hat{Y}_i)$, $i = 1, \ldots, n$ by P_n and $\hat{P}_n$, respectively, so

$$\hat{P}_n(C) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_C(\hat{X}_i, \hat{Y}_i).$$

Let $k = k(n) \le n$ be a sequence of positive numbers such that

(11) $k \to \infty$ and $k/n \to 0$ as $n \to \infty$.

Set $T_\Phi := \{\theta \in [0, \pi/2] : \Phi \text{ is continuous at } \theta\} \cup \{\pi/2\}$. Recall $\Phi(\theta) = \Lambda(C_\theta)$ and $\hat{\Phi}(\theta) = \frac{n}{k} \hat{P}_n(\frac{k}{n} C_\theta)$.

THEOREM 1.   (i) *Suppose* (1) *and* (11) *hold. Then for all* $\theta \in T_\Phi$,

$$\hat{\Phi}(\theta) \overset{\mathbf{P}}{\to} \Phi(\theta).$$

(ii) *Suppose in addition that*

$$k/\log \log n \to \infty,$$

as $n \to \infty$. *Then for all* $\theta \in T_\Phi$,

$$\hat{\Phi}(\theta) \to \Phi(\theta) \qquad a.s.$$

Note that the statements in Theorem 1 imply convergence of $\hat{\Phi}$ to Φ ("in probability" and "almost surely," respectively) in the vague topology on the space $M_+([0, \pi/2])$ of nonnegative Radon measures on $[0, \pi/2]$.

PROOF.   (i) Set

$$\hat{\Lambda}(A) = \frac{n}{k} \hat{P}_n\left(\frac{k}{n} A\right).$$

From Huang [(1992), Chapter 2, Theorem 1], for each $x, y \ge 0$,

$$\hat{\Lambda}\left\{([x, \infty] \times [y, \infty])^c\right\} \overset{\mathbf{P}}{\to} \Lambda\left\{([x, \infty] \times [y, \infty])^c\right\}.$$

For $\varepsilon > 0$ choose two finite unions of (bounded or unbounded) rectangles L_ε and U_ε such that $L_\varepsilon \subset C_\theta \subset U_\varepsilon$ and

$$\Lambda(U_\varepsilon) - \varepsilon < \Lambda(C_\theta) < \Lambda(L_\varepsilon) + \varepsilon.$$

Also we have

$$\hat{\Lambda}(L_\varepsilon) \leq \hat{\Phi}(\theta) = \hat{\Lambda}(C_\theta) \leq \hat{\Lambda}(U_\varepsilon).$$

Since now clearly $\hat{\Lambda}(L_\varepsilon)$ and $\hat{\Lambda}(U_\varepsilon)$ are weakly consistent estimators of $\Lambda(L_\varepsilon)$ and $\Lambda(U_\varepsilon)$, respectively, and since $\Lambda(U_\varepsilon) - \Lambda(L_\varepsilon) \to 0$ as $\varepsilon \downarrow 0$, if $\theta \in T_\Phi$, we conclude that

$$\hat{\Phi}(\theta) \overset{\mathbf{P}}{\to} \Phi(\theta),$$

for $\theta \in T_\Phi$.

(ii) We have from Qi (1997) that under the stated conditions for $x, y \geq 0$,

$$\hat{\Lambda}\left\{([x, \infty] \times [y, \infty])^c\right\} \to \Lambda\left\{([x, \infty] \times [y, \infty])^c\right\} \quad \text{a.s.}$$

The rest of the proof is the same as in the first part. $\square$

We will now consider the process

$$\sqrt{k}\left(\hat{\Phi}(\theta) - \Phi(\theta)\right), \quad \theta \in [0, \pi/2].$$

We will assume that the density λ of Λ exists, that it is continuous on $[0, \infty)^2 \setminus \{(0, 0)\}$, and that $\Lambda([0, \infty) \times \{\infty\}) = \Lambda(\{\infty\} \times [0, \infty)) = 0$. Note that asymptotic independence in the tail, that is, G has independent marginals, is excluded now. Observe that $\lambda(tx, ty) = \frac{1}{t}\lambda(x, y)$. Define

$$\hat{C}_\theta = \frac{n}{k}\left\{(x, y) \in [0, \infty]^2 \setminus \{(\infty, \infty)\} : (\Gamma_{1n}(x), \Gamma_{2n}(y)) \in \frac{k}{n}C_\theta\right\}.$$

Then we have [note $\hat{P}_n(\frac{k}{n}C_\theta) = P_n(\frac{k}{n}\hat{C}_\theta)$]

$$\sqrt{k}\left(\hat{\Phi}(\theta) - \Phi(\theta)\right) = \sqrt{k}\left(\frac{n}{k}P_n\left(\frac{k}{n}\hat{C}_\theta\right) - \frac{n}{k}P\left(\frac{k}{n}\hat{C}_\theta\right)\right)$$

(12)
$$+ \sqrt{k}\left(\frac{n}{k}P\left(\frac{k}{n}\hat{C}_\theta\right) - \Lambda\left(\hat{C}_\theta\right)\right)$$

$$+ \sqrt{k}\left(\Lambda\left(\hat{C}_\theta\right) - \Lambda\left(C_\theta\right)\right)$$

$$:= V_1(\theta) + r(\theta) + V_2(\theta), \qquad \theta \in [0, \pi/2].$$

Define W_Λ to be a Wiener process with "time" Λ, that is, a centred Gaussian process with $\mathbf{E}W_\Lambda(C)W_\Lambda(\tilde{C}) = \Lambda(C \cap \tilde{C})$. Note that

$$\{W_\Lambda(C_\theta), \ \theta \in [0, \pi/2]\} \overset{d}{=} \{W(\Phi(\theta)), \ \theta \in [0, \pi/2]\},$$

with W a standard Wiener process on $[0, \infty)$. Define $W_1(x) = W_\Lambda([0, x] \times [0, \infty])$ and $W_2(y) = W_\Lambda([0, \infty] \times [0, y])$. Note that W_1 and W_2 are also standard Wiener processes. Define the process Z by

$$Z(\theta) = \int_0^{1 \vee \frac{1}{\tan\theta}} \lambda(x, x\tan\theta)\left\{W_1(x)\tan\theta - W_2(x\tan\theta)\right\} dx$$

$$-W_2(1) \int_{1 \vee \frac{1}{\tan \theta}}^{\infty} \lambda(x, 1) \, dx$$

$$-\mathbf{1}_{(\pi/4, \pi/2]}(\theta) W_1(1) \int_1^{\tan \theta} \lambda(1, y) \, dy, \qquad \theta \in [0, \pi/2),$$

$$Z(\pi/2) = -W_2(1) \int_1^{\infty} \lambda(x, 1) \, dx - W_1(1) \int_1^{\infty} \lambda(1, y) \, dy.$$

Our aim is to show that

$$\{(V_1(\theta), r(\theta), V_2(\theta)) : \theta \in [0, \pi/2]\} \xrightarrow{d} \{(W_\Lambda(C_\theta), 0, Z(\theta)) : \theta \in [0, \pi/2]\},$$

where "$\xrightarrow{d}$" denotes weak convergence in $\left(D\left[0, \frac{\pi}{2}\right]\right)^3$, with the supremum norm on $D\left[0, \frac{\pi}{2}\right]$ (and the Euclidian norm on R^3).

Now we are almost ready to present the theorem on the weak convergence of $\sqrt{k}(\hat{\Phi} - \Phi)$, but we need two conditions. First, we need the domain of attraction condition (8) *uniformly* on a class of sets $\mathscr{A}' = \{A \cap A' : A, A' \in \mathscr{A}\}$, where $\mathscr{A} = \mathscr{A}(\Delta, M)$ is a Vapnik-Chervonenkis (VC) class of sets defined as follows. Let $\Delta \in \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, $p = 0, 1, 2, \ldots, \frac{1}{\Delta} - 1$, and define $I_\Delta(p) = \left[p\frac{\Delta}{\tan \theta}, (p+1)\frac{\Delta}{\tan \theta}\right]$, $\theta \in [0, \pi/4]$. Set $\mathscr{A}$ to be the class containing all the following sets:

- $\bigcup_{p=0}^{\frac{1}{\Delta}-1} \left\{(x, y) : x \in I_\Delta(p), \ 0 \le y \le x \tan \theta + C_p(x \tan \theta)^{\frac{1}{16}}\right\}$,
  for some $\theta \in [0, \pi/4]$ and $C_0, C_1 \ldots, C_{\frac{1}{\Delta}-1} \in [-1, 1]$ and

- $\{(x, y) : y \le b\}$, for some $b \le 2$, and

- $\{(x, y) : x \le a\}$, $\{(x, y) : x \le M, y \le 2\}$, for some $a \le M$ (later on M will be taken large) and

- $\left\{(x, y) : x \ge \frac{1}{\tan \theta}, y \le b\right\}$, for some $\theta \in [0, \pi/4]$ and $b \le 2$.

Next define $\tilde{\mathscr{A}}_s = \{A_s : A \in \tilde{\mathscr{A}}\}$, where, for $A \in \mathscr{A}$, $A_s = \{(x, y) : (y, x) \in A\}$. Finally define $\mathscr{A} = \tilde{\mathscr{A}} \cup \tilde{\mathscr{A}}_s$.

CONDITION 1.   *For all* $\Delta \in \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ *and* $M > 1$

$$\sup_{A \in \mathscr{A}'} \left|t^{-1} P(tA) - \Lambda(A)\right| \to 0 \quad (t \downarrow 0).$$

The second condition is partly similar to Condition 1, but it also specifies a rate of convergence in the domain of attraction condition, which can be seen as a condition on the sequence k. Consider the class of sets $\mathscr{C} = \mathscr{C}(\beta)$ defined by

$$\mathscr{C} = \{\{(x, y) : 0 \le y \le b(x) \text{ for some non-decreasing function } b\}$$

(13)
$$\in \mathscr{B}([0, \infty]^2 \setminus \{(\infty, \infty)\}) :$$

$$\sup_{0 < x \le \frac{2}{\tan \theta}} (x \tan \theta)^{-1/16} |b(x) - ((x \tan \theta) \wedge 1)| \le \beta$$

$$\text{for some } \theta \in [0, \pi/4], \text{ and } b(x) = b(2/\tan \theta) \text{ for } x > 2/\tan \theta\},$$

where $\mathscr{B}([0,\infty]^2\backslash\{(\infty,\infty)\})$ denotes the set of Borel sets on $[0,\infty]^2\backslash\{(\infty,\infty)\}$.

CONDITION 2a(i). *For some* $\beta > 0$,

$$D(t) := \sup_{C \in \mathscr{C}} \left| t^{-1} P(tC) - \Lambda(C) \right| \to 0 \quad (t \downarrow 0).$$

CONDITION 2a(ii). *Let* k *be such that (for the same* β)

$$\sqrt{k} D\left(\frac{k}{n}\right) \to 0 \qquad (n \to \infty).$$

CONDITION 2b. *This is similar to Conditions* 2a(i) *and* 2a(ii), *but with* x *and* y *interchanged.*

THEOREM 2. *Assume the framework of Section* 1 *and suppose* Λ *has a continuous density* λ *on* $[0,\infty)^2 \setminus \{(0,0)\}$. *Under Conditions* 1 *and* 2 *we have, as* $n \to \infty$,

$$\sqrt{k}\left(\hat{\Phi}(\theta) - \Phi(\theta)\right) \xrightarrow{d} W_\Lambda(C_\theta) + Z(\theta) \qquad in \ D[0,\pi/2].$$

Note that $W_\Lambda(C_0) + Z(0) = 0$ a.s.

EXAMPLE. Let us consider the bivariate Cauchy distribution on $(0,\infty)^2$ with density

$$\frac{2}{\pi(1+x^2+y^2)^{\frac{3}{2}}}, \qquad x,\,y > 0.$$

We will show now that this distribution satisfies Conditions 1 and 2. It is straightforward to compute the density of $\frac{\pi}{2t}P((0,\frac{2t}{\pi}x] \times (0,\frac{2t}{\pi}y])$ [cf. (8)]; it is equal to

(14)
$$t\frac{(1+\cot^2(tx))(1+\cot^2(ty))}{(1+\cot^2(tx)+\cot^2(ty))^{\frac{3}{2}}}.$$

Relations (4.3.70) and (23.1.15) in Abramowitz and Stegun (1966) imply

(15)
$$\frac{1}{u} - \frac{8u}{\pi^2} < \cot u < \frac{1}{u} \qquad \text{for } 0 < u < \frac{\pi}{2}.$$

Note that (14) is correct only for $tx, ty < \frac{\pi}{2}$. For either x large, y bounded or the other way around, the reasoning below should be slightly adapted. Using (15) we find that (14) can be written, for $t \downarrow 0$, as

$$t\frac{\left(\frac{1}{t^2x^2} + O(1)\right)\left(\frac{1}{t^2y^2} + O(1)\right)}{\left(\frac{1}{t^2}\left(\frac{1}{x^2} + \frac{1}{y^2}\right) + O(1)\right)^{\frac{3}{2}}} = \frac{\left(\frac{1}{x^2} + O(t^2)\right)\left(\frac{1}{y^2} + O(t^2)\right)}{\left(\frac{1}{x^2} + \frac{1}{y^2} + O(t^2)\right)^{\frac{3}{2}}},$$

with the O-terms not depending on x or y. Now for $u > 0$, $|h| < \frac{u}{2}$,

$$
\left| u^{-\frac{3}{2}} - (u+h)^{-\frac{3}{2}} \right| = \left| \frac{3}{2} \int_u^{u+h} s^{-\frac{5}{2}} ds \right| < 2^{\frac{5}{2}} \frac{3}{2} h u^{-\frac{5}{2}},
$$

hence

$$
\frac{1}{\left(\frac{1}{x^2} + \frac{1}{y^2} + O(t^2) \right)^{\frac{3}{2}}} = \frac{1}{\left(\frac{1}{x^2} + \frac{1}{y^2} \right)^{\frac{3}{2}}} + \frac{1}{\left(\frac{1}{x^2} + \frac{1}{y^2} \right)^{\frac{5}{2}}} O(t^2).
$$

So the main term is

$$
\lambda(x, y) = \frac{\frac{1}{x^2 y^2}}{\left(\frac{1}{x^2} + \frac{1}{y^2} \right)^{\frac{3}{2}}}
$$

and the error term is

$$
\left(\frac{\frac{1}{x^2 y^2}}{\left(\frac{1}{x^2} + \frac{1}{y^2} \right)^{\frac{5}{2}}} + \frac{\frac{1}{x^2} + \frac{1}{y^2}}{\left(\frac{1}{x^2} + \frac{1}{y^2} \right)^{\frac{3}{2}}} + \frac{\frac{1}{x^2} + \frac{1}{y^2}}{\left(\frac{1}{x^2} + \frac{1}{y^2} \right)^{\frac{5}{2}}} \right) O(t^2).
$$

The first factor of this expression has a finite integral over a strip $[0, u] \times [0, \infty)$ or $[0, \infty) \times [0, u]$. We conclude that for the present probability distribution $t^{-1} P(tA)$ converges to $\Lambda(A)$ uniformly over the family of all Borel sets contained in a set of the form $\{(x, y) : x \leq u \text{ or } y \leq u\}$. This is sufficient for Conditions 1 and 2.

3. Proof of Theorem 2.

3.1.   We first prove weak convergence of $\sqrt{k} \left(\hat{\Phi}(\theta) - \Phi(\theta) \right)$ in $D[0, \pi/4]$. More precisely, we will show that for probabilistically equivalent versions of the processes involved and any $\varepsilon > 0$

$$
(16) \quad \lim_{\Delta \downarrow 0} \limsup_{n \to \infty} \mathbf{P} \left\{ \sup_{\theta \in \left[0, \frac{\pi}{4} \right]} \left| \sqrt{k} \left(\hat{\Phi}(\theta) - \Phi(\theta) \right) - (W_\Lambda(C_\theta) + Z(\theta)) \right| \geq 3\varepsilon \right\} = 0,
$$

where $\hat{\Phi} = \hat{\Phi}_\Delta$ and $W_\Lambda = W_{\Lambda,\Delta}$ ($\Delta \in \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$).

In the sequel we will replace $\hat{C}_\theta$, $\theta \in \left[0, \frac{\pi}{4} \right]$, by

$$
\left\{ (x, y) : y \leq \frac{n}{k} Q_{2n} \left((\tan \theta) \Gamma_{1n} \left(x \frac{k}{n} \right) \right), y \leq \frac{n}{k} Q_{2n} \left(\frac{k}{n} \right) \right\},
$$

where Q_{jn}, is the quantile function corresponding to Γ_{jn} (set $Q_{jn}(y) = 0$ for $0 \leq y \leq \frac{1}{2n}$) $j = 1, 2$, and call it $\hat{C}_\theta$ again. Both sets are not exactly equal, but the difference is negligible for our purposes. Define the marginal tail empirical processes by

$$
w_{jn}(x) = \frac{n}{\sqrt{k}} \left(\Gamma_{jn} \left(x \frac{k}{n} \right) - x \frac{k}{n} \right), \qquad x \geq 0, \; j = 1, 2,
$$

and the marginal tail quantile process by

$$v_{jn}(x) = \frac{n}{\sqrt{k}}\left(Q_{jn}\left(x\frac{k}{n}\right) - x\frac{k}{n}\right), \qquad x \geq 0, \ j = 1, 2.$$

Note that for $x \geq 0$,

(17)
$$\frac{n}{k}Q_{2n}\left((\tan\theta)\Gamma_{1n}\left(x\frac{k}{n}\right)\right)$$
$$= x\tan\theta + \frac{1}{\sqrt{k}}\left\{(\tan\theta)w_{1n}(x) + v_{2n}\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)\right\}.$$

3.1.1.   First we deal with $V_1(\theta)$ in (12). Write

$$z_{n,\theta}(x) = (\tan\theta)w_{1n}(x) + v_{2n}\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)$$

and note that $\frac{n}{k}Q_{2n}\left(\frac{k}{n}\right) = 1 + \frac{1}{\sqrt{k}}v_{2n}(1)$. Recall the definitions of $\mathscr{A}$ and $\tilde{\mathscr{A}}$, just above Condition 1. Define (with the convention that $0/0=0$)

$$V^+_{p,\Delta,\theta} = \sup_{x\in I_\Delta(p)} \left\{z_{n,\theta}(x) \wedge (v_{2n}(1) + \sqrt{k}(1 - x\tan\theta))\right\} / (x\tan\theta)^{\frac{1}{16}}$$

and

$$V^-_{p,\Delta,\theta} = \inf_{x\in I_\Delta(p)} \left\{z_{n,\theta}(x) \wedge (v_{2n}(1) + \sqrt{k}(1 - x\tan\theta))\right\} / (x\tan\theta)^{\frac{1}{16}}.$$

Set, for either choice of sign,

$$H^\pm_{p,\Delta,\theta} = \left\{(x, y) : \ x \in I_\Delta(p), 0 \leq y \leq x\tan\theta + \frac{1}{\sqrt{k}}(x\tan\theta)^{\frac{1}{16}}V^\pm_{p,\Delta,\theta}\right\}$$

and

$$M^\pm_{\Delta,\theta} = \bigcup_{p=0}^{\frac{1}{\Delta}-1} H^\pm_{p,\Delta,\theta}.$$

Here it should be noted, especially for $p = 0$, that $V^\pm_{p,\Delta,\theta} = O_{\mathbf{P}}(1)$ as $n \to \infty$. In particular, it is useful to write

$$\frac{v_{2n}\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)}{(x\tan\theta)^{\frac{1}{16}}}$$

$$= \frac{v_{2n}\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)}{\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)^{\frac{1}{4}}} \left((x\tan\theta)^{\frac{3}{4}} + \frac{1}{\sqrt{k}}(\tan\theta)^{\frac{3}{4}}w_{1n}(x)/x^{\frac{1}{4}}\right)^{\frac{1}{4}}$$

and to use the fact that $w_{1n}/I^{\frac{1}{4}}$ (on $[0, 1/\tan\theta]$) and $v_{2n}/I^{\frac{1}{4}}$ (on $[0, 2]$) are bounded in distribution (I is the identity function).

Now we apply Theorem 3.1 of Einmahl (1997); see also Einmahl, de Haan and Sinha (1997). Then using that $\mathscr{A}$ is a VC class and Condition 1, we have for a special construction (but keeping the same notation), as $n \to \infty$,

$$(18) \qquad \sup_{A \in \mathscr{A}} \left| \sqrt{k} \left(\frac{n}{k} P_n \left(\frac{k}{n} A \right) - \frac{n}{k} P \left(\frac{k}{n} A \right) \right) - W_\Lambda(A) \right| \overset{\text{a.s.}}{\to} 0.$$

Set $\hat{C}_{\theta,1} = \left\{ (x, y) \in \hat{C}_\theta : x \le \frac{1}{\tan \theta} \right\}$, $\hat{C}_{\theta,2} = \hat{C}_\theta \setminus \hat{C}_{\theta,1}$, and define for $j = 1, 2$,

$$V_{1,j}(\theta) = \sqrt{k} \left(\frac{n}{k} P_n \left(\frac{k}{n} \hat{C}_{\theta,j} \right) - \frac{n}{k} P \left(\frac{k}{n} \hat{C}_{\theta,j} \right) \right), \qquad j = 1, 2.$$

Then

$$V_{1,1}(\theta) \le \sqrt{k} \left(\frac{n}{k} P_n \left(\frac{k}{n} M_{\Delta,\theta}^+ \right) - \frac{n}{k} P \left(\frac{k}{n} M_{\Delta,\theta}^+ \right) \right)$$

$$(19) \qquad\qquad + \sqrt{k} \frac{n}{k} P \left(\frac{k}{n} \left(M_{\Delta,\theta}^+ \setminus M_{\Delta,\theta}^- \right) \right)$$

$$=: V_{1,1}^+(\theta) + r_1(\theta);$$

similarly

$$V_{1,1}(\theta) \ge \sqrt{k} \left(\frac{n}{k} P_n \left(\frac{k}{n} M_{\Delta,\theta}^- \right) - \frac{n}{k} P \left(\frac{k}{n} M_{\Delta,\theta}^- \right) \right)$$

$$(20) \qquad\qquad - \sqrt{k} \frac{n}{k} P \left(\frac{k}{n} \left(M_{\Delta,\theta}^+ \setminus M_{\Delta,\theta}^- \right) \right)$$

$$=: V_{1,1}^-(\theta) - r_1(\theta).$$

We now first deal with $r_1(\theta)$ and next with $V_{1,1}^\pm(\theta)$.

Using Condition 2a and the results on the behavior of weighted tail empirical and quantile processes [see Einmahl (1992, 1997)] we can show that, as $n \to \infty$,

$$(21) \qquad \sup_{\theta \in [0,\pi/4]} \left| r_1(\theta) - \sqrt{k} \Lambda \left(M_{\Delta,\theta}^+ \setminus M_{\Delta,\theta}^- \right) \right| \overset{\mathbf{P}}{\to} 0.$$

Now consider

$$(22) \qquad \sup_{\theta \in [0,\pi/4]} \sqrt{k} \Lambda \left(M_{\Delta,\theta}^+ \setminus M_{\Delta,\theta}^- \right).$$

Note that

$$\sqrt{k} \Lambda \left(M_{\Delta,\theta}^+ \setminus M_{\Delta,\theta}^- \right)$$

$$= \sqrt{k} \sum_{p=0}^{\frac{1}{\Delta}-1} \int_{p \frac{\Delta}{\tan \theta}}^{(p+1) \frac{\Delta}{\tan \theta}} \int_{x \tan \theta + \frac{1}{\sqrt{k}} (x \tan \theta)^{\frac{1}{16}} V_{p,\Delta,\theta}^-}^{x \tan \theta + \frac{1}{\sqrt{k}} (x \tan \theta)^{\frac{1}{16}} V_{p,\Delta,\theta}^+} \lambda(x, y) \, dy \, dx.$$

Setting $y = x\tan\theta + \frac{1}{\sqrt{k}}(x\tan\theta)^{\frac{1}{16}}z$ we obtain

$$\sum_{p=0}^{\frac{1}{\Delta}-1}\int_{p\frac{\Delta}{\tan\theta}}^{(p+1)\frac{\Delta}{\tan\theta}}\int_{V_{p,\Delta,\theta}^{-}}^{V_{p,\Delta,\theta}^{+}}\lambda\left(x, x\tan\theta + \frac{1}{\sqrt{k}}(x\tan\theta)^{\frac{1}{16}}z\right)(x\tan\theta)^{\frac{1}{16}}\,dz\,dx$$

$$\leq 16\sup_{y\geq 0}\lambda(1, y)\max_{p\in\left\{0,1,\ldots,\frac{1}{\Delta}-1\right\}}\left(V_{p,\Delta,\theta}^{+} - V_{p,\Delta,\theta}^{-}\right).$$

Since $\lambda(1, y) = y^{-1}\lambda(1/y, 1)$ and by the continuity of λ on $[0,\infty)^2\setminus\{(0,0)\}$ we have $\lim_{y\to\infty}\lambda(1, y) = 0$. Hence $\sup_{y\geq 0}\lambda(1, y) < \infty$. Also because of the tightness of w_{jn}/I^δ and v_{jn}/I^δ, $j = 1, 2$, $0 < \delta < 1/2$, on $[0, M]$, we see that for any $\tilde{\varepsilon} > 0$,

$$\lim_{\Delta\downarrow 0}\limsup_{n\to\infty}\mathbf{P}\left\{\sup_{\theta\in[0,\pi/4]}\max_{p\in\left\{0,1,\ldots,\frac{1}{\Delta}-1\right\}}\left(V_{p,\Delta,\theta}^{+} - V_{p,\Delta,\theta}^{-}\right) \geq \frac{\tilde{\varepsilon}}{2}\right\} = 0,$$

and hence, using (21),

$$(23)\qquad\qquad \lim_{\Delta\downarrow 0}\limsup_{n\to\infty}\mathbf{P}\left\{\sup_{\theta\in[0,\pi/4]}r_1(\theta) \geq \frac{\varepsilon}{2}\right\} = 0.$$

Now consider for either choice of sign $V_{1,1}^{\pm}(\theta)$. Since $\mathbf{P}(M_{\Delta,\theta}^{\pm} \in \tilde{\mathscr{A}}$, for all $\theta \in [0, \pi/4]) \to 1$ $(n \to \infty)$, we have, using (18), that

$$(24)\qquad\qquad \sup_{\theta\in[0,\pi/4]}\left|V_{1,1}^{\pm}(\theta) - W_\Lambda\left(M_{\Delta,\theta}^{\pm}\right)\right| \xrightarrow{\mathbf{P}} 0.$$

But with similar calculations as for (22) we obtain that

$$\Lambda\left(M_{\Delta,\theta}^{\pm}\,\triangle\,C_{\theta,1}\right) \leq \frac{16}{\sqrt{k}}\sup_{y\geq 0}\lambda(1, y)\max_{p\in\left\{0,1,\ldots,\frac{1}{\Delta}-1\right\}}\left|V_{p,\Delta,\theta}^{\pm}\right|$$

with

$$C_{\theta,1} = \left\{(x, y) \in C_\theta : x \leq \frac{1}{\tan\theta}\right\}$$

$$= \left\{(x, y) \in [0,\infty)^2 : 0 \leq x \leq \frac{1}{\tan\theta}, 0 \leq y \leq x\tan\theta\right\}.$$

Since

$$\sup_{\theta\in[0,\pi/4]}\max_{p\in\left\{0,1,\ldots,\frac{1}{\Delta}-1\right\}}\left|V_{p,\Delta,\theta}^{\pm}\right| = O_{\mathbf{P}}(1),$$

we have that for any $\Delta \in \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$,

$$\sup_{\theta\in[0,\pi/4]}\Lambda\left(M_{\Delta,\theta}^{\pm}\,\triangle\,C_{\theta,1}\right) \xrightarrow{\mathbf{P}} 0.$$

Hence, since W_Λ is uniformly continuous on $\mathscr{A}$ (with respect to the pseudo-metric
$\Lambda(A \triangle A')$, $A, A' \in \mathscr{A}$),

$$(25) \qquad \sup_{\theta \in [0, \pi/4]} \left| W_\Lambda \left(M_{\Delta, \theta}^\pm \right) - W_\Lambda \left(C_{\theta, 1} \right) \right| \overset{\mathbf{P}}{\to} 0.$$

Combining (19), (20), (23), (24) and (25), we now have proven that

$$(26) \qquad \lim_{\Delta \downarrow 0} \limsup_{n \to \infty} \mathbf{P} \left\{ \sup_{\theta \in [0, \pi/4]} \left| V_{1,1}(\theta) - W_\Lambda \left(C_{\theta, 1} \right) \right| \geq \varepsilon \right\} = 0.$$

Observe that $\hat{C}_{\theta, 2}$ is (almost) a rectangle. [Only near $(1/\tan\theta, 1)$ there is a small deviation from the rectangular shape, but with some care it can be shown that this deviation is negligible.] But these rectangles are in the VC class $\tilde{\mathscr{A}}$ and need no approximation like $\hat{C}_{\theta, 1}$. Therefore we can show in a similar but easier way than for $V_{1,1}$ that, with $C_{\theta, 2} = C_\theta \setminus C_{\theta, 1}$,

$$(27) \qquad \lim_{\Delta \downarrow 0} \limsup_{n \to \infty} \mathbf{P} \left\{ \sup_{\theta \in [0, \pi/4]} \left| V_{1,2}(\theta) - W_\Lambda \left(C_{\theta, 2} \right) \right| \geq \varepsilon \right\} = 0.$$

Combining (26) and (27), we now have, as $n \to \infty$,

$$(28) \qquad \lim_{\Delta \downarrow 0} \limsup_{n \to \infty} \mathbf{P} \left\{ \sup_{\theta \in [0, \pi/4]} \left| V_1(\theta) - W_\Lambda \left(C_\theta \right) \right| \geq 2\varepsilon \right\} = 0.$$

3.1.2.   Next we consider $V_2(\theta)$. We show that, as $n \to \infty$,

$$(29) \qquad \sup_{\theta \in [0, \pi/4]} \left| \sqrt{k} \left(\Lambda(\hat{C}_\theta) - \Lambda(C_\theta) \right) - Z(\theta) \right| \overset{\mathbf{P}}{\to} 0.$$

Note that for $\theta \in [0, \pi/4]$:

$$Z(\theta) = \int_0^{\frac{1}{\tan\theta}} \lambda(x, x\tan\theta) \{ W_1(x)\tan\theta - W_2(x\tan\theta) \} \, dx - W_2(1) \int_{\frac{1}{\tan\theta}}^\infty \lambda(x, 1) \, dx.$$

Observe, with $\hat{C}_{\theta, 1}$, $C_{\theta, 1}$ and $z_{n, \theta}$ as before, that

$$\sqrt{k} \left(\Lambda(\hat{C}_{\theta, 1}) - \Lambda(C_{\theta, 1}) \right)$$

$$(30) \qquad = \sqrt{k} \int_0^{\frac{1}{\tan\theta}} \int_{x\tan\theta}^{x\tan\theta + \frac{1}{\sqrt{k}} \left\{ z_{n, \theta}(x) \wedge \left(v_{2n}(1) + \sqrt{k}(1 - x\tan\theta) \right) \right\}} \lambda(x, y) \, dy \, dx.$$

Now for (large) $M > 1$,

$$\sup_{\theta \in [0, \pi/4]} \left| \sqrt{k} \left(\Lambda(\hat{C}_{\theta, 1}) - \Lambda(C_{\theta, 1}) \right) \right.$$

$$\left. - \int_0^{\frac{1}{\tan\theta}} \lambda(x, x\tan\theta) \{ W_1(x)\tan\theta - W_2(x\tan\theta) \} \, dx \right|$$

$$\leq \sup_{\theta \in \left[\arctan \frac{1}{M}, \pi/4 \right]} \left| \sqrt{k} \left(\Lambda(\hat{C}_{\theta,1}) - \Lambda(C_{\theta,1}) \right) \right.$$

$$\left. - \int_0^{\frac{1}{\tan \theta}} \lambda(x, x \tan \theta) \left\{ W_1(x) \tan \theta - W_2(x \tan \theta) \right\} dx \right|$$

$$+ \sup_{\theta \in \left[0, \arctan \frac{1}{M} \right]} \left| \sqrt{k} \left(\Lambda(\hat{C}_{\theta,1}) - \Lambda(C_{\theta,1}) \right) \right|$$

$$+ \sup_{\theta \in \left[0, \arctan \frac{1}{M} \right]} \left| \int_0^{\frac{1}{\tan \theta}} \lambda(x, x \tan \theta) \left\{ W_1(x) \tan \theta - W_2(x \tan \theta) \right\} dx \right|$$

$$=: T_1 + T_2 + T_3.$$

We have

$$T_3 \leq \sup_{\theta \in \left[0, \arctan \frac{1}{M} \right]} \int_0^{\frac{1}{\tan \theta}} \frac{1}{x} \lambda(1, \tan \theta) \left\{ |W_1(x)| \tan \theta + |W_2(x \tan \theta)| \right\} dx$$

$$\leq \sup_{\theta \in \left[0, \arctan \frac{1}{M} \right]} \left\{ (\tan \theta) \lambda(1, \tan \theta) \int_0^{\frac{1}{\tan \theta}} \frac{|W_1(x)|}{x} dx \right.$$

$$\left. + \lambda(1, \tan \theta) \int_0^1 \frac{|W_2(y)|}{y} dy \right\}$$

(31)

$$\leq \sup_{\theta \in \left[0, \arctan \frac{1}{M} \right]} \left\{ (\tan \theta) \lambda(1, \tan \theta) \left(\sup_{x \in [0,1]} \frac{W_1(x)}{x^{1/4}} \int_0^1 \frac{1}{v^{3/4}} dv \right. \right.$$

$$\left. + \sup_{x \geq 1} \frac{|W_1(x)|}{x^{3/4}} \int_1^{\frac{1}{\tan \theta}} \frac{1}{v^{1/4}} dv \right)$$

$$\left. + \lambda(1, \tan \theta) \sup_{y \in [0,1]} \frac{|W_2(y)|}{y^{1/4}} \int_0^1 \frac{1}{v^{3/4}} dv \right\}.$$

Since P has uniform marginals we have

$$\int_1^\infty \lambda(x, 0) \, dx \leq 1.$$

But since $\lambda(x, 0) = \frac{1}{x} \lambda(1, 0)$, this implies $\lambda(1, 0) = 0$. Hence by the continuity of λ: $\lim_{y \downarrow 0} \lambda(1, y) = 0$. Combining this with (31) yields that for any $\eta > 0$,

(32)
$$\lim_{M \to \infty} \mathbf{P} \{ T_3 \geq \eta \} = 0.$$

Let us consider T_2 now. For T_2, and also for T_1, we will replace $z_{n,\theta}(x) \wedge (v_{2n}(1) + \sqrt{k}(1 - x \tan \theta))$ by $z_{n,\theta}(x)$ in the right-hand side of (30), since it can be shown that the difference between these two expressions is negligible. Concerning T_2 we have

$$\left| \sqrt{k} \int_0^1 \int_{x \tan \theta}^{x \tan \theta + \frac{1}{\sqrt{k}} z_{n,\theta}(x)} \lambda(x, y) \, dy \, dx \right|$$

$$\le \sup_{y \ge 0} \lambda(1, y) \left| \int_0^1 \frac{1}{x} z_{n,\theta}(x) \, dx \right|$$

$$\le \sup_{y \ge 0} \lambda(1, y) \left\{ \tan \theta \sup_{x \in [0,1]} \frac{|w_{1n}(x)|}{x^{1/16}} \right.$$

$$+ (\tan \theta)^{\frac{1}{16}} \sup_{x \in [0,1]} \frac{\left| v_{2n} \left(x \tan \theta + \frac{1}{\sqrt{k}} (\tan \theta) w_{1n}(x) \right) \right|}{(x \tan \theta)^{\frac{1}{16}}} \left. \right\}$$

$$\times \int_0^1 \frac{1}{v^{15/16}} \, dv.$$

Also

$$\left| \sqrt{k} \int_1^{\frac{1}{\tan \theta}} \int_{x \tan \theta}^{x \tan \theta + \frac{1}{\sqrt{k}} z_{n,\theta}(x)} \lambda(x, y) \, dy \, dx \right|$$

$$= \left| \int_1^{\frac{1}{\tan \theta}} \frac{1}{x} \int_0^{z_{n,\theta}(x)} \lambda \left(1, \tan \theta + \frac{z}{x\sqrt{k}} \right) dz \, dx \right|$$

$$\le \sup_{|z| \le \sup_{1 \le x \le \frac{1}{\tan \theta}} |z_{n,\theta}(x)|} \lambda \left(1, \tan \theta + \frac{z}{\sqrt{k}} \right)$$

$$\times \sup_{1 \le x \le \frac{1}{\tan \theta}} \frac{|z_{n,\theta}(x)|}{(x \tan \theta)^{1/16}} \int_1^{\frac{1}{\tan \theta}} \frac{1}{v} (v \tan \theta)^{1/16} \, dv.$$

Hence, since

$$\sup_{\theta \in [0, \pi/4]} \sup_{1 \le x \le \frac{1}{\tan \theta}} \frac{|z_{n,\theta}(x)|}{(x \tan \theta)^{1/16}} = O_{\mathbf{P}}(1),$$

we see, somewhat similar as for T_3, that

(33) $$\lim_{M \to \infty} \limsup_{n \to \infty} \mathbf{P} \{ T_2 \ge \eta \} = 0.$$

Finally consider T_1. Write $z_\theta(x) = W_1(x) \tan \theta - W_2(x \tan \theta)$. Then we have

$$T_1 \le \sup_{\theta \in [\arctan \frac{1}{M}, \frac{\pi}{4}]} \left| \sqrt{k} \int_0^{\frac{1}{\tan \theta}} \int_{x \tan \theta + \frac{1}{\sqrt{k}} z_\theta(x)}^{x \tan \theta + \frac{1}{\sqrt{k}} z_{n,\theta}(x)} \lambda(x, y) \, dy \, dx \right|$$

$$+ \sup_{\theta \in [\arctan \frac{1}{M}, \frac{\pi}{4}]} \left| \sqrt{k} \int_0^{\frac{1}{\tan \theta}} \int_{x \tan \theta}^{x \tan \theta + \frac{1}{\sqrt{k}} z_\theta(x)} \lambda(x, y) \, dy \, dx \right.$$

$$\left. - \int_0^{\frac{1}{\tan \theta}} \lambda(x, x \tan \theta) z_\theta(x) \, dx \right|$$

$$=: T_{1,1} + T_{1,2}.$$

For handling $T_{1,1}$, note that it can be easily shown that

$$\sup_{\theta\in[0,\pi/4]} \sup_{0\le x\le M} \frac{|z_{n,\theta}(x)-z_\theta(x)|}{(x\tan\theta)^{1/16}} \xrightarrow{\mathbf{P}} 0.$$

We have

$$T_{1,1} \le \sup_{y\ge 0} \lambda(1,y) \int_0^{\frac{1}{\tan\theta}} \frac{1}{x} \frac{|z_{n,\theta}(x)-z_\theta(x)|}{(x\tan\theta)^{1/16}} (x\tan\theta)^{1/16}\, dx.$$

Hence, for any $M>1$,

$$(34) \qquad\qquad T_{1,1} \xrightarrow{\mathbf{P}} 0.$$

In the term $T_{1,2}$ we split up outer integral in the integral from 0 to δ ($0 < \delta < 1$), and from δ to $\frac{1}{\tan\theta}$, and denote the corresponding expressions by $T_{1,2,1}$ and $T_{1,2,2}$, respectively. Then

$$(35) \quad T_{1,2,1} \le 2 \sup_{y\ge 0} \lambda(1,y) \sup_{\theta\in\left[\arctan\frac{1}{M},\frac{\pi}{4}\right]} \sup_{x\in[0,\delta]} \frac{|z_\theta(x)|}{(x\tan\theta)^{1/4}} \int_0^\delta \frac{1}{v}(v\tan\theta)^{1/4}\, dv$$

and

$$(36) \qquad T_{1,2,2} \le \sup_{\theta\in\left[\arctan\frac{1}{M},\frac{\pi}{4}\right]} \left| \int_\delta^{\frac{1}{\tan\theta}} \frac{1}{x} \int_0^{z_\theta(x)} \left(\lambda\left(1,\tan\theta + \frac{z}{x\sqrt{k}}\right) \right. \right.$$
$$\left.\left. -\lambda(1,\tan\theta)\right) dz\, dx \right|.$$

Now noting that

$$\sup_{\theta\in\left[\arctan\frac{1}{M},\pi/4\right]} \sup_{0\le x\le \frac{1}{\tan\theta}} |z_\theta(x)| < \infty \qquad \text{a.s.}$$

and that $\frac{1}{\tan\theta} \le M$, we obtain from (35) and (36) that for any $M>1$,

$$(37) \qquad\qquad T_{1,2} \xrightarrow{\mathbf{P}} 0.$$

Combining (32)-(34) and (37) yields that, as $n\to\infty$,

$$(38) \quad \sup_{\theta\in[0,\pi/4]} \left| \sqrt{k}\left(\Lambda(\hat{C}_{\theta,1}) - \Lambda(C_{\theta,1})\right) - \int_0^{\frac{1}{\tan\theta}} \lambda(x, x\tan\theta)z_\theta(x)\, dx \right| \xrightarrow{\mathbf{P}} 0.$$

Similarly, but much more easily, we obtain,

$$(39) \quad \sup_{\theta\in[0,\pi/4]} \left| \sqrt{k}\left(\Lambda(\hat{C}_{\theta,2}) - \Lambda(C_{\theta,2})\right) - W_2(1)\int_{\frac{1}{\tan\theta}}^\infty \lambda(x,1)\, dx \right| \xrightarrow{\mathbf{P}} 0.$$

Combining (38) and (39) yields (29).

3.1.3.   We now consider $r(\theta)$ in (12). From (17), Condition 2a and the well-known behavior of weighted tail empirical and quantile processes, it now follows that

$$(40) \qquad \sup_{\theta \in [0, \pi/4]} |r(\theta)| \overset{\mathbf{P}}{\to} 0 \text{ as } n \to \infty.$$

Combining (28),(29) and (40) yields (16). So actually we proved the theorem for $\theta \in [0, \pi/4]$.

3.2.   Next note that it rather easy to show that, as $n \to \infty$,

$$\left| \sqrt{k} \left(\hat{\Phi} \left(\frac{\pi}{2} \right) - \Phi \left(\frac{\pi}{2} \right) \right) - \left(W_\Lambda \left(C_{\frac{\pi}{2}} \right) + Z \left(\frac{\pi}{2} \right) \right) \right| \overset{\mathbf{P}}{\to} 0.$$

Hence it follows by a symmetry argument, observing that for $\theta \in (\pi/4, \pi/2]$ (the closure of) $C_{\frac{\pi}{2}} \setminus C_\theta$ is the mirror image (with respect to the line $y = x$) of $C_{\frac{\pi}{2} - \theta}$, that, as $n \to \infty$,

$$(41) \quad \lim_{\Delta \downarrow 0} \limsup_{n \to \infty} \mathbf{P} \left\{ \sup_{\theta \in \left(\frac{\pi}{4}, \frac{\pi}{2} \right]} \left| \sqrt{k} \left(\hat{\Phi}(\theta) - \Phi(\theta) \right) - (W_\Lambda(C_\theta) + Z(\theta)) \right| \geq 4\varepsilon \right\} = 0.$$

Combining (16) and (41) completes the proof. □

4. An application and simulations.   The National Institute for Coastal and Marine management of The Netherlands provided a data set consisting of wave heights and still water levels during 828 storm events spread over 13 years in front of the Dutch coast near the town of Petten. They can be considered independent and all following the same probability distribution. These observations are relevant for a small stretch of sea dike near Petten that protects a gap in the natural coast protection formed by sand dunes. The dike is called "Pettemer zeedijk," but it is also known as the "Hondsbossche Zeewering." Figure 1 displays the estimated spectral measure

$$\hat{\Phi}(\theta) = \frac{1}{k} \sum_{i=1}^n \mathbf{1}_{\{R_i^{\mathscr{X}} \vee R_i^{\mathscr{Y}} \geq n+1-k, \arctan \frac{n+1-R_i^{\mathscr{Y}}}{n+1-R_i^{\mathscr{X}}} \leq \theta\}}$$

$(0 \leq \theta/\pi \leq 1/2)$, for various values of k: $k = 20, 28, 30$ and 40. The solid line represents $k = 28$. The graph seems to be rather robust against changes of k. Since the graph has sufficient increase away from the endpoints, we may conclude that there seems to be asymptotic dependence between the variables "wave height" and "still water level."

For comparison the spectral measure has also been estimated as in Einmahl, de Haan and Sinha (1997):

$$\widetilde{\Phi}(\theta) := \frac{1}{k} \sum_{i=1}^n \mathbf{1}_{\{\widetilde{X}_i(\frac{n}{k}) \vee \widetilde{Y}_i(\frac{n}{k}) > 1, \arctan \frac{\widetilde{X}_i(\frac{n}{k})}{\widetilde{Y}_i(\frac{n}{k})} \leq \theta\}}$$
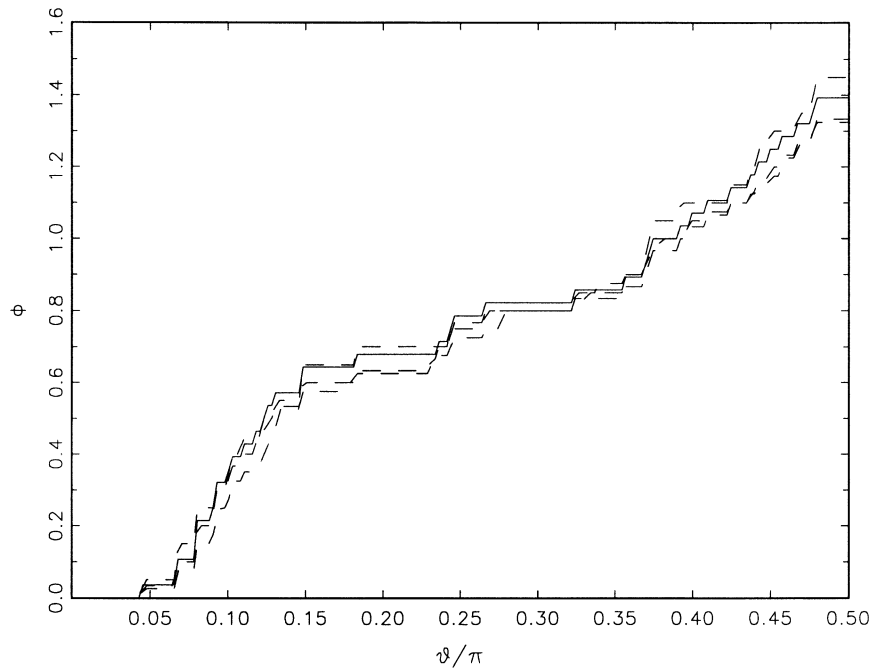
FIG. 1.   *The solid line is the nonparametric estimator* $\hat{\Phi}$ *for* $k = 28$, *the dashed lines correspond to the same estimator for* $k = 20, 30$ *and* 40, *respectively*.

with

$$\widetilde{X}_i \left(\frac{n}{k} \right) := \left(1 + \hat{\gamma}_1 \frac{\mathscr{X}_i - \hat{b}_1(\frac{n}{k})}{\hat{a}_1(\frac{n}{k})} \right)^{1/\hat{\gamma}_1},$$

$$\widetilde{Y}_i \left(\frac{n}{k} \right) := \left(1 + \hat{\gamma}_2 \frac{\mathscr{Y}_i - \hat{b}_2(\frac{n}{k})}{\hat{a}_2(\frac{n}{k})} \right)^{1/\hat{\gamma}_2},$$

and $\hat{a}_j$, $\hat{b}_j$ and $\hat{\gamma}_j$ $(j = 1, 2)$ appropriate estimators of scale, location and shape, respectively. Judging from the marginal distributions [cf. de Haan and de Ronde (1998)] the choice $k = 28$ seems to be natural. But we have again displayed the estimates with $k = 20, 28, 30$ and 40 (Figure 2; the solid line represents $k = 28$). One sees that this semiparametric estimator is less robust with respect to these variations of k.

Simulations have been carried out using a bivariate Cauchy distribution restricted to the first quadrant, but with the marginal distributions transformed so that these are similar in the tail to those in the application above, that is,
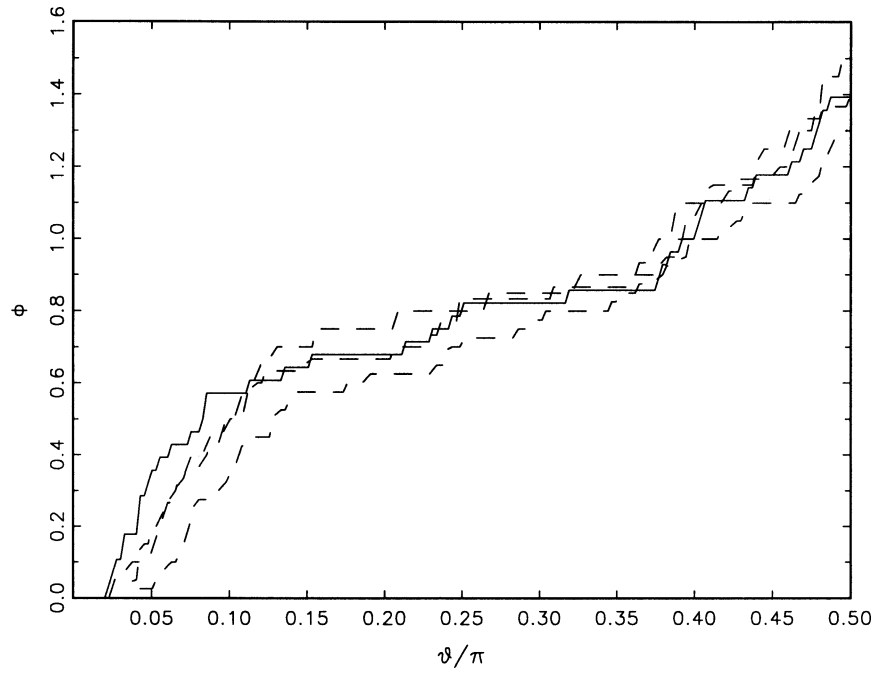
FIG. 2. *The solid line is the semiparametric estimator $\widetilde{\Phi}$ for $k = 28$, the dashed lines correspond to the same estimator for $k = 20$, 30 and 40, respectively.*
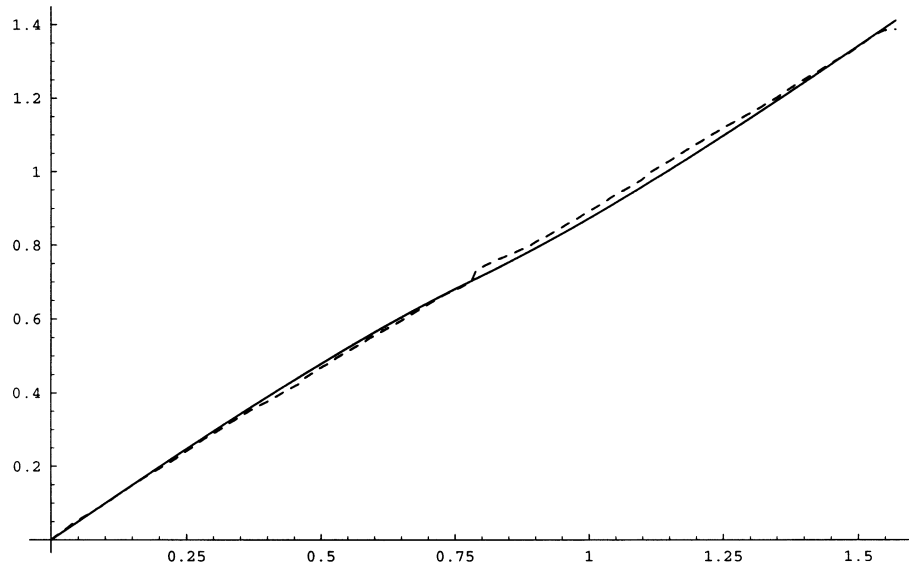

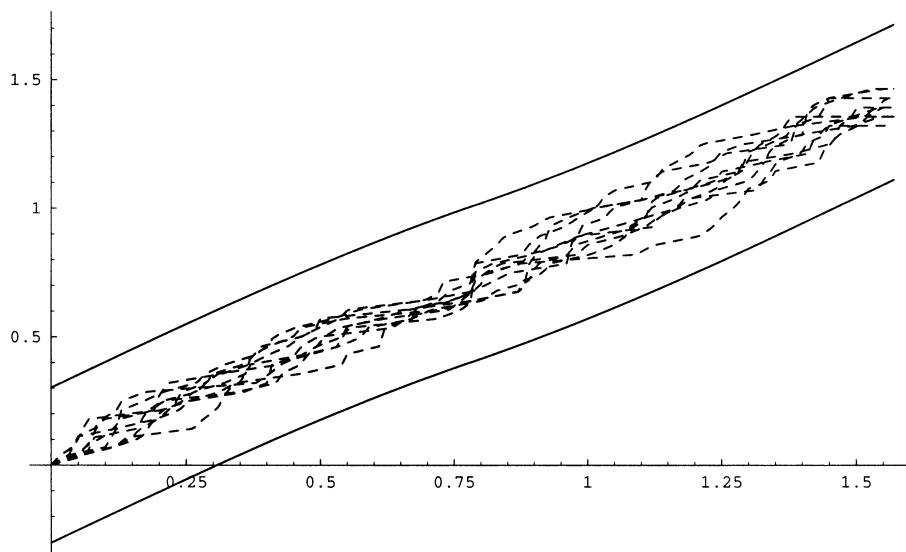FIG. 3. *The true Φ (solid) and the average of 100 estimates $\hat{\Phi}$ (dashed).*

FIG. 4.   10 *estimates* $\hat{\Phi}$ *and asymptotic probability*-0.95 *bands around* Φ.

from the density function

$$\frac{2(1+\gamma_1 x)^{\frac{1}{\gamma_1}-1}(1+\gamma_2 y)^{\frac{1}{\gamma_2}-1}}{\pi\left(1+(1+\gamma_1 x)^{\frac{2}{\gamma_1}}+(1+\gamma_2 y)^{\frac{2}{\gamma_2}}\right)^{\frac{3}{2}}},$$

with x and y such that $1 + \gamma_1 x$, $1 + \gamma_2 y > 0$. Note that we showed in Section 2 that this distribution satisfies Conditions 1 and 2 there. We have chosen $(\gamma_1, \gamma_2) = (-0.0074, -0.1215)$, the same as the estimates for the variables 'wave height' and 'still water level' above. This is similar to what has been done in de Haan and Sinha (1999). Our estimator $\hat{\Phi}$, however, is clearly not affected by strictly increasing transformations of the marginals. It can be computed that for this example $\Phi(\theta) = \sin \theta$, for $\theta \in [0, \pi/4]$, and $\Phi(\theta) = \sqrt{2} - \cos \theta$, for $\theta \in (\pi/4, \pi/2]$. We took, as in the application above, *always* $n = 828$ and $k = 28$. Figure 3 displays the average of 100 (independent) estimates $\hat{\Phi}$ (dashed), as well as the true function Φ (solid). Clearly, they are very close to each other. In Figures 4 and 5 we plotted respectively 10 and 100 estimates $\hat{\Phi}$. In those pictures the solid lines are bands around the true Φ, computed from Theorem 2 and 500 simulations of the limiting Gaussian process appearing there. To be more precise, the (constant) vertical distance between Φ and one of these bands is the simulated 0.95 quantile of the supremum of the absolute value of this limiting process divided by $\sqrt{k}$, with $k = 28$. This shows that for this example the asymptotic theory seems to work well. Maybe the bands are a bit conservative, but this could also be due to the inaccuracy in the simulation of the aforementioned 0.95 quantile.
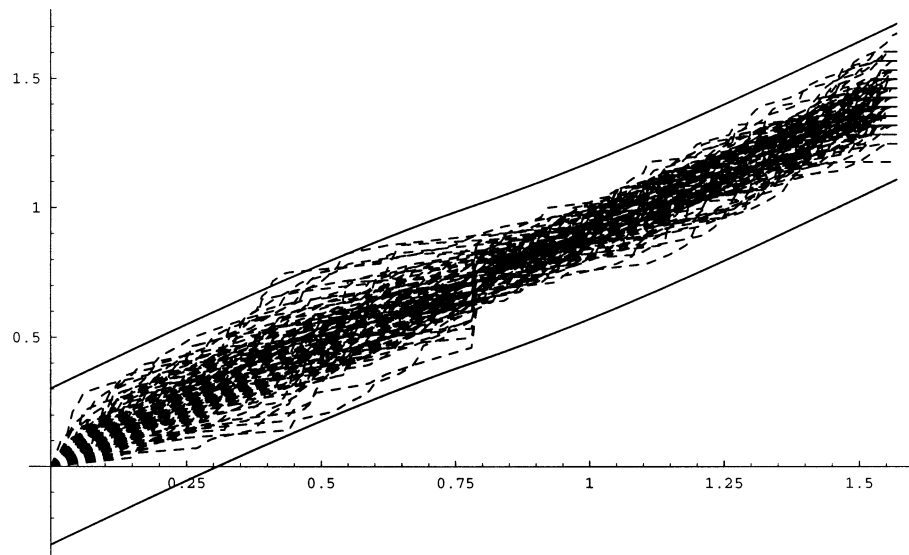
FIG. 5.   100 *estimates* $\hat{\Phi}$ *and asymptotic probability*-0.95 *bands around* Φ.

REFERENCES

ABDOUS, B., GHOUDI, K. and KHOUDRAJI, A. (1999). Non-parametric estimation of the limit dependence function of multivariate extremes. *Extremes* **2** 245–268.

ABRAMOWITZ, M. and STEGUN, I. A. (1966). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

COLES, S. G. and TAWN, J. A. (1991). Modelling extreme multivariate events. *J. Roy. Statist. Soc. Ser. B* **53** 377–392.

HAUKSSON, H. A., DACOROGNA, M. M., DOMENIG, T., MÜLLER, U. and SAMORODNITSKY, G. (2001) Multivariate extremes, aggregation and risk estimation. *Quantitative Finance* **1** 79–95.

DEHEUVELS, P. (1978). Charactérisation complète des lois extrêmes multivariées et de la convergence aux types extrêmes. *Publ. Inst. Statist. Univ. Paris* **23** 1–36.

DEHEUVELS, P. and MARTYNOV, G. V. (1996). Cramér-von Mises-type tests with applications to tests of independence for multivariate extreme-value distributions. *Comm. Statist. Theory Methods* **25** 871–908.

DRAISMA, G., DE HAAN, L., PENG, L. AND SINHA, A. K. (1996, 1997). Report Neptune T400 EUR 4, 5, 6, 7, 8, 9, 10, 12; Reports EUR/RIKZ 96.2 and 96.3. Available at www.few.eur.nl/few/people/ldehaan.

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

EINMAHL, J. H. J. (1992). Limit theorems for tail processes with application to intermediate quantile processes. *J. Statist. Plann. Inference* **32** 137–145.

EINMAHL, J. H. J. (1997). Poisson and Gaussian approximation of weighted local empirical processes. *Stochastic Process. Appl.* **70** 31–58.

EINMAHL, J. H. J., DE HAAN, L. and HUANG, X. (1993). Estimating a multidimensional extreme-value distribution. *J. Multivariate Anal.* **47** 35–47.

EINMAHL, J. H. J., DE HAAN, L. and SINHA, A. K. (1997). Estimating the spectral measure of an extreme value distribution. *Stochastic Process. Appl.* **70** 143–171.

DE HAAN, L. and RESNICK, S. (1977). Limit theory for multidimensional sample extremes. *Z. Wahrsch. Verw. Gebiete* **40** 317–337.

DE HAAN, L. and RESNICK, S. (1993). Estimating the limit distribution of multivariate extremes. *Comm. Statist. Stochastic Models* **9** 275–309.

DE HAAN, L. and DE RONDE, J. (1998) Sea and wind: multivariate extremes at work. *Extremes* **1** 7–45.

DE HAAN, L. and SINHA, A. K. (1999). Estimating the probability of a rare event. *Ann. Statist.* **27** 732–759.

JOE, H., SMITH, R. L. and WEISSMAN, I. (1992). Bivariate threshold methods for extremes. *J. Roy. Statist. Soc. Ser. B* **54** 171–183.

HUANG, X. (1992). Statistics of bivariate extremes. Thesis, Erasmus Univ., Tinbergen Institute Research Series 22.

PICKANDS, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd Session of the International Statistical Institute* **2** 859–878. ISI, Voorburg, Netherlands.

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

RESNICK, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

TAWN, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika* **75** 397–415.

J. H. J. EINMAHL
DEPARTMENT OF ECONOMICS
TILBURG UNIVERSITY
P.O. BOX 90153
5000 LE TILBURG
THE NETHERLANDS
E-MAIL: j.h.j.einmahl@kub.nl

L. DE HAAN
ECONOMETRIC INSTITUTE
ERASMUS UNIVERSITY
P.O. BOX 1738
3000 DR ROTTERDAM
THE NETHERLANDS
E-MAIL: ldhaan@few.eur.nl

V. I. PITERBARG
FACULTY OF MECHANICS AND MATHEMATICS
MOSCOW STATE UNIVERSITY
VOROBJOVY GORY
MOSCOW 119899
RUSSIA
E-MAIL: piter@mech.math.msu.su