Report 98-01

Nonparametric estimation of the spectral measure of an extreme value distribution

J.H.J. Einmahl
L. de Haan
V.I. Piterbarg

Nonparametric estimation of the spectral measure of an extreme value distribution

John H.J. Einmahl*
Eindhoven University of Technology and EURANDOM,
Laurens de Haan†
Erasmus University Rotterdam and EURANDOM,
Vladimir I. Piterbarg‡
Moscow Lomonosov State University.

October 15, 1998

Abstract

Let $(\mathcal{X}_1, \mathcal{Y}_1), ..., (\mathcal{X}_n, \mathcal{Y}_n)$ 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. An application is given.

AMS 1991 subject classifications. Primary 62G05, 62G30; secondary 60G70, 60F15, 60F17.
Key words and phrases. Dependence structure, empirical process, functional central limit theorem, multivariate extremes, nonparametric estimation.

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, i.e. 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 an arbitrary

*Research partially supported by a Fulbright Grant and by European Union HCM grant ERB CHRX-CT 940693. J. Einmahl thanks the Dept. of Statistics of Florida State University for kind hospitality.
†L. de Haan and V. Piterbarg were partially supported by the Stochastics Centre in Gothenburg.
‡Research partially supported by Russian Federation RFFI Grants 95-01-01232, 97-01-00648, 98-01-00524 and Netherlands organization of scientific research though the program Computer intensive methods in Statistics.

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 non-parametric 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 $(\mathcal{X},\mathcal{Y}),(\mathcal{X}_1,\mathcal{Y}_1),(\mathcal{X}_2,\mathcal{Y}_2),...,(\mathcal{X}_n,\mathcal{Y}_n)$ be i.i.d. with common 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

$$\mathrm{P}\left\{\frac{\max_{1\le i\le n}\mathcal{X}_i - b_n}{a_n}\le x, \frac{\max_{1\le i\le n}\mathcal{Y}_i - d_n}{c_n}\le y\right\}$$

converges to a limit distribution function, say $G(x,y)$, with non-degenerate marginals, i.e.

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

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

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

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).

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,

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

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

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

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

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

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

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

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

This interpretation is the basis for estimating Φ.

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

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

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

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

provided $\Lambda(\partial A) = 0$ with the measure Λ on $[0, \infty]^2 \setminus \{(\infty, \infty)\}$ defined by

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

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

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

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

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

The result is

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

for all but countably many θ. Note that we simplified the notation: $\Phi(\theta) := \Phi([0, \theta])$.

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 necessity to estimate six parameters causes mathematical problems: asymptotic normality was only proved for $\gamma_1, \gamma_2 > 0$.

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):

$$\begin{aligned}\hat{\Phi}(\theta) &:= \frac{n}{k}\hat{P}_n\left(\frac{k}{n}C_\theta\right)\\ &= \frac{1}{k}\sum_{i=1}^{n} 1_{\{(n+1-R_i^{\mathcal{X}})\wedge(n+1-R_i^{\mathcal{Y}})\le k, n+1-R_i^{\mathcal{Y}}\le(n+1-R_i^{\mathcal{X}})\tan\theta\}}\\ &= \frac{1}{k}\sum_{i=1}^{n} 1_{\{R_i^{\mathcal{X}}\vee R_i^{\mathcal{Y}}\ge n+1-k, n+1-R_i^{\mathcal{Y}}\le(n+1-R_i^{\mathcal{X}})\tan\theta\}}\end{aligned}$$

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

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

where

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

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* $\gamma_i \in R, i = 1, 2$. 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 (1997) and in a restricted parametric context, Tawn (1988), Coles and Tawn (1991), Joe, Smith and Weissman (1992). The paper 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 $\hat{\Phi}$, one *does* get an extreme value distribution.

Apart from this, $\hat{\Phi}$ is useful for assessing the amount of independence in the tail of F. (Note that G has independent marginals if and only if Φ is concentrated on $\{0, \pi/2\}$). $\hat{\Phi}$ is also necessary as a building block for the analysis of probabilities of rare sets in an extreme value context (de Haan and Sinha (1997)).

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 is given in Section 3. Section 4 contains an application.

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 F 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 $\Lambda(\partial A)=0$, where $tA=\{(tx,ty):(x,y)\in A\}$. Note that $\Lambda([0,tx]\times[0,ty])=t\Lambda([0,x]\times[0,y])$ and that $0\leq\Lambda([0,x]\times[0,y])\leq x\wedge y$. Furthermore $\Lambda([0,\infty]\times[0,x])=\Lambda([0,x]\times[0,\infty])=x$. Set

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

We consider $(X_1,Y_1),...,(X_n,Y_n)$ where $(X_i,Y_i)=(1-F_1(\mathcal{X}_i),1-F_1(\mathcal{Y}_i))$, $i=1,...,n$. We denote the marginal empirical distribution functions of $(X_1,Y_1),...,(X_n,Y_n)$ with F_{1n} and F_{2n}, so, e.g., $F_{1n}(x)=\frac{1}{n}\sum_{i=1}^n 1_{(-\infty,x]}(X_i)$. Now we transform the data by F_{1n} and F_{2n} in the following way: $(\hat{X}_i,\hat{Y}_i)=(F_{1n}(X_i),F_{2n}(Y_i))$, $i=1,...,n$. 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 $(\hat{X}_i,\hat{Y}_i)$, $i=1,...,n$ with P_n and $\hat{P}_n$, respectively, so

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

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

$$k\to\infty \text{ and } k/n\to 0 \text{ as } n\to\infty. \tag{10}$$

Recall $\Phi(\theta)=\Lambda(C_\theta)$ and $\hat{\Phi}(\theta)=\frac{n}{k}\hat{P}_n(\frac{k}{n}C_\theta)$.

Theorem 1 *1. Suppose (1) and (10) hold. Then*

$$\hat{\Phi}\xrightarrow{P}\Phi$$

in the vague topology on the space $M_+([0,\pi/2])$ of nonnegative Radon measures on $[0,\pi/2]$.

2. Suppose in addition that

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

as $n\to\infty$. Then

$$\hat{\Phi}\to\Phi,\ a.s.$$

in the vague topology.

Proof. 1. We know from Huang (1992), Chapter 2, Theorem 1, that in the vague topology on $M_+([0,\infty]^2 \setminus \{(\infty,\infty)\}$ the measure $\hat{\Lambda}$ represented by

$$\hat{\Lambda}\left\{([x,\infty) \times [y,\infty))^c\right\} := \frac{1}{k}\sum_{i=1}^{n} 1_{\{R_i^X > n+1-kx \text{ or } R_i^Y > n+1-ky\}}$$

satisfies

$$\hat{\Lambda} \xrightarrow{P} \Lambda.$$

Next consider the transformation $\bar{T}: M_+\left([0,\infty]^2 \setminus \{\infty,\infty\}\right) \to M_+\left([0,\infty] \times [0,\pi/2]\right)$ defined by

$$\bar{T}\Lambda := \Lambda \circ T^{\leftarrow}$$

where $T := ([0,\infty]^2 \setminus \{(\infty,\infty)\} \to ([0,\infty) \times [0,\pi/2]$ is defined by $T(x,y) = (r,w))$ with $r = x \wedge y$, $w = \arctan(y/x)$. Note that $T^{\leftarrow}$ is a continuous function. So if $K \subset [0,\infty] \times [0,\pi/2]$ is a compact set, then $T^{\leftarrow}(K)$ is a compact set in $[0,\infty]^2 \setminus \{\infty,\infty\}$. So by Resnick (1987, Proposition 3.18, page 148), we get that $\bar{T}$ is a continuous map in the vague topology. Thus we get that

$$\hat{\Lambda} \circ T^{\leftarrow} \xrightarrow{P} \Lambda \circ T^{\leftarrow}.$$

But since $\hat{\Phi}(\cdot) = \hat{\Lambda} \circ T^{\leftarrow}([0,1] \times \cdot)$ and $\Phi(\cdot) = \Lambda \circ T^{\leftarrow}([0,1] \times \cdot)$, we conclude that

$$\hat{\Phi} \xrightarrow{P} \Phi$$

in the vague topology.

2. We have from Qi (1997) that under the stated conditions

$$\hat{\Lambda} \to \Lambda, \text{ a.s.}$$

in the vague topology. The rest of the proof is the same as in the first part. □

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 and that it is continuous on $[0,\infty)^2 \setminus \{(0,0)\}$. Observe that $\lambda(tx,ty) = \frac{1}{t}\lambda(x,y)$. Define

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

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

$$\begin{aligned}
\sqrt{k}\left(\hat{\Phi}(\theta) - \Phi(\theta)\right) &= \sqrt{k}\left(\frac{n}{k}P_n\left(\frac{k}{n}\hat{C}_\theta\right) - \frac{n}{k}P\left(\frac{k}{n}\hat{C}_\theta\right)\right) \\
&\quad + \sqrt{k}\left(\frac{n}{k}P_n\left(\frac{k}{n}\hat{C}_\theta\right) - \Lambda\left(\hat{C}_\theta\right)\right) \\
&\quad + \sqrt{k}\left(\Lambda\left(\hat{C}_\theta\right) - \Lambda\left(C_\theta\right)\right) \\
&:= V_1(\theta) + r(\theta) + V_2(\theta), \qquad \theta \in [0,\pi/2]. \qquad (11)
\end{aligned}$$

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

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

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

$$\begin{aligned} Z(\theta) &= \int_0^{1\vee\frac{1}{\tan\theta}} \lambda(x, x\tan\theta)\,\{W_1(x)\tan\theta - W_2(x\tan\theta)\}\,dx \\ &\quad -W_2(1)\int_{1\vee\frac{1}{\tan\theta}}^{\infty} \lambda(x,1)\,dx \\ &\quad -1_{(\pi/4,\pi/2]}(\theta)W_1(1)\int_1^{\tan\theta} \lambda(1,y)\,dy, \quad \theta\in[0,\pi/2]. \end{aligned}$$

Our aim is to show that $V_1(\theta) \xrightarrow{d} W_\Lambda(C_\theta)$, $r(\theta) \xrightarrow{d} 0$ and $V_2(\theta) \xrightarrow{d} Z(\theta)$, where '$\xrightarrow{d}$' denotes weak convergence in $D\left[0,\frac{\pi}{2}\right]$, with the supremum norm.

Now we are almost ready to present the theorem on the weak convergence of $\sqrt{k}\left(\hat{\Phi}-\Phi\right)$, but we need two conditions. Let $\mathcal{A} = \mathcal{A}(\Delta, M)$ be as in the proof of the theorem and write $\mathcal{A}' = \{A\cap A' : A, A' \in \mathcal{A}\}$.

Condition 1

$$\sup_{A\in\mathcal{A}'}\left|\frac{n}{k}P\left(\frac{k}{n}A\right) - \Lambda(A)\right| \to 0.$$

Let $C_n \subset [0,\infty]^2 \setminus (\infty,\infty)$ be a sequence of sets. Write $C_n(x) = C_n \cap (\{x\}\times[0,\infty])$ and assume C_n is such that $C_n(\mathrm{x})$ is of the form $\{x\}\times[0,b_n(x)]$, ($[0,b_n(x)]=\emptyset$ possibly) for some $b_n(x)$.

Condition 2a *For all $L>0$ we have: if for some $\theta\in[0,\pi/4]$,*

$$\sqrt{k}\sup_{0\le x\le\frac{2}{\tan\theta}} |b_n(x) - ((x\tan\theta)\wedge 1)| \le L(x\tan\theta)^{\frac{1}{16}}, \tag{12}$$

for all $n\in\mathrm{N}$, then

$$\lim_{n\to\infty}\sup\sqrt{k}\left|\frac{n}{k}P\left(\frac{k}{n}C_n\right) - \Lambda(C_n)\right| = 0,$$

where the 'sup' is taken over all C_n satisfying (12) for some $\theta\in[0,\pi/4]$.

Condition 2b *is similar to Condition 2a, 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, 2a and 2b we have, as $n\to\infty$,*

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

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

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

3 Proof of Theorem 2

A.

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, as $n\to\infty$,

$$\sup_{\theta\in[0,\frac{\pi}{4}]}\left|\sqrt{k}\left(\hat{\Phi}(\theta)-\Phi(\theta)\right)-\left(W_\Lambda(C_\theta)+Z(\theta)\right)\right|\xrightarrow{P}0 \tag{13}$$

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

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

where Q_{jn}, is the quantile function corresponding to F_{jn}, $j=1,2$, and call it $\hat{C}_\theta$ again. Both sets are not exactly equal due to the fact that F_{2n} is a step-function, but the difference is negligible for our purposes. Define the marginal tail empirical processes by

$$w_{jn}=\begin{cases}\frac{n}{\sqrt{k}}\left(F_{jn}\left(x\frac{k}{n}\right)-x\frac{k}{n}\right), & 0\le x\le\frac{n}{k},\\ 0, & x>\frac{n}{k},\ j=1,2,\end{cases}$$

and the marginal tail quantile process by

$$v_{jn}=\begin{cases}\frac{n}{\sqrt{k}}\left(Q_{jn}\left(x\frac{k}{n}\right)-x\frac{k}{n}\right), & 0\le x\le\frac{n}{k},\\ 0, & x>\frac{n}{k},\ j=1,2.\end{cases}$$

Note that

$$\frac{n}{k}Q_{2n}\left((\tan\theta)F_{1n}\left(x\frac{k}{n}\right)\right)=x\tan\theta+\frac{1}{\sqrt{k}}\left\{(\tan\theta)w_{1n}(x)+v_{2n}\left(x\tan\theta+\frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)\right\}. \tag{14}$$

A.1.

First we deal with $V_1(\theta)$ in (11). Let $\Delta>0$, such that $1/\Delta\in\mathrm{N}$. Let $p=0,1,2,...,\frac{1}{\Delta}-1$, and define $I_\Delta(p)=\left[p\frac{\Delta}{\tan\theta},(p+1)\frac{\Delta}{\tan\theta}\right]$, $\theta\in[0,\pi/4]$. We set $\mathcal{A}$ 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 $C_0, C_1 ..., C_{\frac{1}{\Delta}-1}\in\mathrm{R}$, 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 $b\le 2$.

Then $\mathcal{A} = \mathcal{A}(\Delta, M)$ is a Vapnik-Chervonenkis (VC) class. 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 $\frac{n}{k}Q_{2n}\left(\frac{k}{n}\right) = 1 + \frac{1}{\sqrt{k}}v_{2n}(1)$. Define

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

and

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

Set, for either choice of sign,

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

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 the $V^{\pm}_{p,\Delta,\theta}$ do not "blow up" as $n\to\infty$. In particular, it is useful to write

$$\begin{aligned}
&\frac{v_{2n}\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)}{(x\tan\theta)^{\frac{1}{16}}}\\
&= \frac{v_{2n}\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)}{\left(x\tan\theta + \frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)^{\frac{1}{4}}}\left((x\tan\theta)^{\frac{3}{4}} + \frac{1}{\sqrt{k}}(\tan\theta)^{\frac{3}{4}}w_{1n}(x)/x^{\frac{1}{4}}\right)
\end{aligned}$$

and to use the fact that $v_{2n}/I^{\frac{1}{4}}$ and $w_{2n}/I^{\frac{1}{4}}$ 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 $\mathcal{A}$ is a VC class and Condition 1, we have for a special construction (but keeping the same notation), as $n\to\infty$,

$$\sup_{A\in\mathcal{A}}\left|\sqrt{k}\left(\frac{n}{k}P_n\left(\frac{k}{n}A\right) - \frac{n}{k}P\left(\frac{k}{n}A\right)\right) - W_\Lambda(A)\right| \xrightarrow{a.s.} 0. \tag{15}$$

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 $j = 1,2$,

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

Then

$$\begin{aligned}
V_{1,1}(\theta) &\le \sqrt{k}\left(\frac{n}{k}P_n\left(\frac{k}{n}M^{+}_{\Delta,\theta}\right) - \frac{n}{k}P\left(\frac{k}{n}M^{+}_{\Delta,\theta}\right)\right)\\
&+ \sqrt{k}\frac{n}{k}P\left(\frac{k}{n}\left(M^{+}_{\Delta,\theta}\setminus M^{-}_{\Delta,\theta}\right)\right)\\
&=: V^{+}_{1,1}(\theta) + r_1(\theta);
\end{aligned} \tag{16}$$

similarly

$$\begin{aligned} V_{1,1}(\theta) &\geq \sqrt{k}\left(\frac{n}{k}P_n\left(\frac{k}{n}M^-_{\Delta,\theta}\right) - \frac{n}{k}P\left(\frac{k}{n}M^-_{\Delta,\theta}\right)\right) \\ &- \sqrt{k}\frac{n}{k}P\left(\frac{k}{n}\left(M^+_{\Delta,\theta}\setminus M^-_{\Delta,\theta}\right)\right) \\ &=: V^-_{1,1}(\theta) - r_1(\theta). \end{aligned} \tag{17}$$

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

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

$$\sup_{\theta\in[0,\pi/4]}\left|r_1(\theta) - \sqrt{k}\Lambda\left(M^+_{\Delta,\theta}\setminus M^-_{\Delta,\theta}\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{18}$$

Now consider

$$\sup_{\theta\in[0,\pi/4]}\sqrt{k}\Lambda\left(M^+_{\Delta,\theta}\setminus M^-_{\Delta,\theta}\right). \tag{19}$$

Note that

$$\begin{aligned} &\sqrt{k}\Lambda\left(M^+_{\Delta,\theta}\setminus M^-_{\Delta,\theta}\right) \\ &= \sqrt{k}\sum_{p=0}^{\frac{1}{\Delta}-1}\int_{p\frac{\Delta}{\tan\theta}}^{(p+1)\frac{\Delta}{\tan\theta}}\int_{x\tan\theta+\frac{1}{\sqrt{k}}(x\tan\theta)^{\frac{1}{16}}V^-_{p,\Delta,\theta}}^{x\tan\theta+\frac{1}{\sqrt{k}}(x\tan\theta)^{\frac{1}{16}}V^+_{p,\Delta,\theta}}\lambda(x,y)\,dy\,dx. \end{aligned}$$

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

$$\begin{aligned} &\sum_{p=0}^{\frac{1}{\Delta}-1}\int_{p\frac{\Delta}{\tan\theta}}^{(p+1)\frac{\Delta}{\tan\theta}}\int_{V^-_{p,\Delta,\theta}}^{V^+_{p,\Delta,\theta}}\lambda(x, x\tan\theta + \frac{1}{\sqrt{k}}(x\tan\theta)^{\frac{1}{16}}z)(x\tan\theta)^{\frac{1}{16}}\,dz\,dx \\ &\leq 16\sup_{y\geq 0}\lambda(1,y)\max_{p\in\{0,1,\ldots,\frac{1}{\Delta}-1\}}\left(V^+_{p,\Delta,\theta} - V^-_{p,\Delta,\theta}\right). \end{aligned}$$

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 $\varepsilon > 0$,

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

and hence, using (18),

$$\lim_{\Delta\downarrow 0}\limsup_{n\to\infty}\mathbf{P}\left\{\sup_{\theta\in[0,\pi/4]} r_1(\theta)\geq\varepsilon\right\} = 0. \tag{20}$$

Now consider for either choice of sign $V^{\pm}_{1,1}(\theta)$. Since $M^{\pm}_{\Delta,\theta}\in\mathcal{A}$, we have, using (15), that

$$\sup_{\theta\in[0,\pi/4]}\left|V^{\pm}_{1,1}(\theta) - W_{\Lambda}\left(M^{\pm}_{\Delta,\theta}\right)\right| \xrightarrow{a.s.} 0. \tag{21}$$

But with similar calculations as for (19) we obtain that

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

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

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

we have that for any $\Delta > 0$ ($\frac{1}{\Delta} \in \mathbb{N}$),

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

Hence, since W_Λ is uniformly continuous on $\mathcal{A}$,

$$\sup_{\theta\in[0,\pi/4]} \left|W_\Lambda\left(M^{\pm}_{\Delta,\theta}\right) - W_\Lambda\left(C_{\theta,1}\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{22}$$

Combining (16), (17), (20), (21) and (22), we now have proven that

$$\sup_{\theta\in[0,\pi/4]} \left|V_{1,1}(\theta) - W_\Lambda\left(C_{\theta,1}\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{23}$$

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 $\mathcal{A}$ 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

$$\sup_{\theta\in[0,\pi/4]} \left|V_{1,2}(\theta) - W_\Lambda\left(C_{\theta,2}\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{24}$$

Combining (23) and (24), we now have, as $n \to \infty$,

$$\sup_{\theta\in[0,\pi/4]} \left|V_1(\theta) - W_\Lambda\left(C_\theta\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{25}$$

A.2.

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

$$\sup_{\theta\in[0,\pi/4]} \left|\sqrt{k}\left(\Lambda(\hat{C}_\theta) - \Lambda(C_\theta)\right) - Z(\theta)\right| \xrightarrow{\mathbf{P}} 0. \tag{26}$$

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

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

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

$$\begin{aligned}&\sqrt{k}\left(\Lambda(\hat{C}_{\theta,1})-\Lambda(C_{\theta,1})\right)\\&\quad=\sqrt{k}\int_0^{\frac{1}{\tan\theta}}\int_{x\tan\theta}^{x\tan\theta+\frac{1}{\sqrt{k}}\{z_{n,\theta}(x)\wedge(v_{2n}(1)+\sqrt{k}(1-x\tan\theta))\}}\lambda(x,y)\,dy\,dx.\end{aligned}\tag{27}$$

Now for (large) $M>1$,

$$\begin{aligned}&\sup_{\theta\in[0,\pi/4]}\Big|\sqrt{k}\left(\Lambda(\hat{C}_{\theta,1})-\Lambda(C_{\theta,1})\right)\\&\qquad-\int_0^{\frac{1}{\tan\theta}}\lambda(x,x\tan\theta)\,\{W_1(x)\tan\theta-W_2(x\tan\theta)\}\,dx\Big|\\&\le\sup_{\theta\in\left[\arctan\frac{1}{M},\pi/4\right]}\Big|\sqrt{k}\left(\Lambda(\hat{C}_{\theta,1})-\Lambda(C_{\theta,1})\right)\\&\qquad-\int_0^{\frac{1}{\tan\theta}}\lambda(x,x\tan\theta)\,\{W_1(x)\tan\theta-W_2(x\tan\theta)\}\,dx\Big|\\&+\sup_{\theta\in\left[0,\arctan\frac{1}{M}\right]}\left|\sqrt{k}\left(\Lambda(\hat{C}_{\theta,1})-\Lambda(C_{\theta,1})\right)\right|\\&+\sup_{\theta\in\left[0,\arctan\frac{1}{M}\right]}\left|\int_0^{\frac{1}{\tan\theta}}\lambda(x,x\tan\theta)\,\{W_1(x)\tan\theta-W_2(x\tan\theta)\}\,dx\right|\\&=:T_1+T_2+T_3.\end{aligned}$$

We have

$$\begin{aligned}T_3&\le\sup\theta\in\left[0,\arctan\frac{1}{M}\right]\int_0^{\frac{1}{\tan\theta}}\frac{1}{x}\lambda(1,\tan\theta)\,\{|W_1(x)|\tan\theta+|W_2(x\tan\theta)|\}\,dx\\&\le\sup_{\theta\in\left[0,\arctan\frac{1}{M}\right]}\left\{(\tan\theta)\lambda(1,\tan\theta)\int_0^{\frac{1}{\tan\theta}}\frac{|W_1(x)|}{x}\,dx+\lambda(1,\tan\theta)\int_0^1\frac{|W_2(y)|}{y}\,dy\right\}\\&\le\sup_{\theta\in\left[0,\arctan\frac{1}{M}\right]}\left\{(\tan\theta)\lambda(1,\tan\theta)\left(\sup_{x\in[0,1]}\frac{W_1(x)}{x^{1/4}}\int_0^1\frac{1}{v^{3/4}}\,dv\right.\right.\\&+\left.\left.\sup_{x\ge1}\frac{|W_1(x)|}{x^{3/4}}\int_1^{\frac{1}{\tan\theta}}\frac{1}{v^{1/4}}\,dv\right)+\lambda(1,\tan\theta)\sup_{y\in[0,1]}\frac{|W_2(y)|}{y^{1/4}}\int_0^1\frac{1}{v^{3/4}}\,dv\right\}.\end{aligned}\tag{28}$$

Since P has uniform marginals we have

$$\int_1^\infty\lambda(x,0)\,dx\le1.$$

But since $\lambda(x,0)=\frac{1}{x}\lambda(1,0)$, this implies $\lambda(1,0)=0$. Hence by the continuity of λ: $\lim_{y\downarrow0}\lambda(1,y)=0$. Combining this with (28) yields that for any $\varepsilon>0$,

$$\lim_{M\to\infty}\mathrm{P}\{T_3\ge\varepsilon\}=0.\tag{29}$$

Let us consider T_2 now. For T_2, and also for T_1, we will replace $z_{n,\theta}\wedge(v_{2n}(1)+\sqrt{k}(1-x\tan\theta))$ by $z_{n,\theta}(x)$ in the right-hand side of (27), since it can be shown that the difference between these two expressions is negligible. Concerning T_2 we have

$$\begin{aligned}
&\left|\sqrt{k}\int_0^1 \int_{x\tan\theta}^{x\tan\theta+\frac{1}{\sqrt{k}}z_{n,\theta}(x)} \lambda(x,y)\,dy\,dx\right| \\
&\le \sup_{y\ge 0}\lambda(1,y)\left|\int_0^1 \frac{1}{x} z_{n,\theta}(x)\,dx\right| \\
&\le \sup_{y\ge 0}\lambda(1,y\Bigg\{\tan\theta \sup_{x\in[0,1]}\frac{|w_{1n}(x)|}{x^{1/16}} \\
&\qquad +(\tan\theta)^{\frac{1}{16}}\sup_{x\in[0,1]}\frac{\left|v_{2n}\left(x\tan\theta+\frac{1}{\sqrt{k}}(\tan\theta)w_{1n}(x)\right)\right|}{(x\tan\theta)^{\frac{1}{16}}}\Bigg\}\int_0^1 \frac{1}{v^{15/16}}\,dv.
\end{aligned}$$

Also

$$\begin{aligned}
&\left|\sqrt{k}\int_1^{\frac{1}{\tan\theta}} \int_{x\tan\theta}^{x\tan\theta+\frac{1}{\sqrt{k}}z_{n,\theta}(x)} \lambda(x,y)\,dy\,dx\right| \\
&= \left|\int_1^{\frac{1}{\tan\theta}} \frac{1}{x}\int_0^{z_{n,\theta}(x)} \lambda\left(1,\tan\theta+\frac{z}{x\sqrt{k}}\right)dz\,dx\right| \\
&\le \sup_{z\le \sup_{1\le x\le\frac{1}{\tan\theta}}|z_{n,\theta}(x)|}\lambda\left(1,\tan\theta+\frac{z}{\sqrt{k}}\right)\sup_{1\le x\le\frac{1}{\tan\theta}}\frac{|z_{n,\theta}(x)|}{(x\tan\theta)^{1/16}}\int_1^{\frac{1}{\tan\theta}}\frac{1}{v}(v\tan\theta)^{1/16}\,dv.
\end{aligned}$$

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

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbf{P}\{T_2\ge\varepsilon\}=0. \tag{30}$$

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

$$\begin{aligned}
T_1 \;&\le\; \sup_{\theta\in[\arctan\frac{1}{M},\frac{\pi}{4}]}\left|\sqrt{k}\int_0^{\frac{1}{\tan\theta}}\int_{x\tan\theta+\frac{1}{\sqrt{k}}z_\theta(x)}^{x\tan\theta+\frac{1}{\sqrt{k}}z_{n,\theta}(x)}\lambda(x,y)\,dy\,dx\right| \\
&\quad+\sup_{\theta\in[\arctan\frac{1}{M},\frac{\pi}{4}]}\Bigg|\sqrt{k}\int_0^{\frac{1}{\tan\theta}}\int_{x\tan\theta}^{x\tan\theta+\frac{1}{\sqrt{k}}z_\theta(x)}\lambda(x,y)\,dy\,dx \\
&\qquad\qquad -\int_0^{\frac{1}{\tan\theta}}\lambda(x,x\tan\theta)z_\theta(x)\,dx\Bigg| \\
&=:\; T_{1,1}+T_{1,2}.
\end{aligned}$$

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 $M > 1$,

$$T_{1,1} \xrightarrow{\mathbf{P}} 0. \tag{31}$$

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 with $T_{1,2,1}$ and $T_{1,2,2}$, respectively. Then

$$T_{1,2,1} \le 2\sup_{y\ge 0} \lambda(1,y) \sup_{\theta\in\left[\arctan\frac{1}{M},\frac{\pi}{4}\right]} \sup_{x\in[0,\delta]} \frac{|z_\theta(x)|}{(x\tan\theta)^{1/4}} \int_0^\delta \frac{1}{v}(v\tan\theta)^{1/4}\,dv \tag{32}$$

and

$$T_{1,2,2} \le \sup_{\theta\in\left[\arctan\frac{1}{M},\frac{\pi}{4}\right]} \left| \int_\delta^{\frac{1}{\tan\theta}} \frac{1}{x} \int_0^{z_\theta(x)} \left(\lambda\left(1,\tan\theta + \frac{z}{x\sqrt{k}}\right) - \lambda(1,\tan\theta)\right) dz\,dx \right|. \tag{33}$$

Now noting that

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

and that $\frac{1}{\tan\theta} \le M$, we obtain from (32) and (33) that for any $M > 1$,

$$T_{1,2} \xrightarrow{\mathbf{P}} 0. \tag{34}$$

Combining (29)-(31) and (34) yields that, as $n \to \infty$,

$$\sup_{\theta\in[0,\pi/4]} \left| \sqrt{k}\left(\Lambda(\hat{C}_{\theta,1}) - \Lambda(C_{\theta,1})\right) - \int_0^{\frac{1}{\tan\theta}} \lambda(x, x\tan\theta) z_\theta(x)\,dx \right| \xrightarrow{\mathbf{P}} 0. \tag{35}$$

Similarly, but much easier, we obtain,

$$\sup_{\theta\in[0,\pi/4]} \left| \sqrt{k}\left(\Lambda(\hat{C}_{\theta,2}) - \Lambda(C_{\theta,2})\right) - W_2(1)\int_{\frac{1}{\tan\theta}}^\infty \lambda(x,1)\,dx \right| \xrightarrow{\mathbf{P}} 0. \tag{36}$$

Combining (35) and (36) yields (26).

A.3.

We now consider $r(\theta)$ in (11). From (14), Condition 2a, and the well-known behaviour of weighted tail empirical and quantile processes, it now follows that

$$\sup_{\theta\in[0,\pi/4]} |r(\theta)| \xrightarrow{\mathbf{P}} 0 \quad \text{as } n \to \infty. \tag{37}$$

Combining (25),(26) and (37) yields (13). So actually we proved the theorem for $\theta \in [0, \pi/4]$.

B.

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

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

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 $n \to \infty$,

$$\sup_{\theta\in\left(\frac{\pi}{4},\frac{\pi}{2}\right]} \left| \sqrt{k}\left(\hat{\Phi}(\theta) - \Phi(\theta)\right) - (W_\Lambda(C_\theta) + Z(\theta)) \right| \xrightarrow{\mathbf{P}} 0. \tag{38}$$

(Obviously, the VC class $\mathcal{A}$ has to be extended for this, but that can be done without any problem.) Combining (13) and (38) completes the proof. □

4 An application

The National Institute for Coastal and Marine management of The Netherlands provided a data set consisting of wave heights (HmO) and still water levels (SWL) 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 that protects a gap in the natural coast protection formed by sand dunes near Petten. The dike is called 'Pettemer zeedijk'. Figure 1 displays the estimated spectral measure

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

($0 \leq \theta \leq \pi/2$), based on 28 extreme observations ($k = 28$) along with the points $\arctan \frac{n+1-R_i^Y}{n+1-R_i^X}$, $i = 1, 2, ..., 28$. Asymptotic dependence seems to be present.

In order to see how non-parametric methods compare with semi-parametric ones (see Section 1), we have displayed two estimators, not of Φ but of the measure Λ through the level sets of its estimated distribution function $\Lambda\{([x,\infty] \times [y,\infty])^c\}$ (cf. (7)) in Figure 2. The upper figure displays the level sets for the non-parametric Λ-estimator (cf. Huang (1992)) and the lower figure the level sets for the semi-parametric Λ-estimator (cf. de Haan and Resnick (1993)). It is known that the level sets of the theoretical Λ-function form concave functions and these functions have the same shape for different levels, Huang (1992). These properties are better reflected in the upper figure than in the lower one. An important reason for the poorer showing in the lower figure is the less than optimal fit of the parametric distribution for lower values of HmO. So, although Figure 2 does not refer to an estimation for Φ, it does contain the message that in this area non-parametric methods for estimating the dependence perform better.

For more information about this application see de Haan and de Ronde (1998).

REFERENCES

ABDOUS, B., GHOUDI, K. and KHOUDRAJI, A. (1997) Non-parametric estimation of the limit dependence function of multivariate extremes. Report Univ. du Québec à Trois-Rivières.

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

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. 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 http://www.cs.few.eur.nl/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. Stat. Plann. Inference* **32** 137-145.

EINMAHL, J.H.J. (1997) Poisson and Gaussian approximation of weighted local empirical processes. *Stoch. Proc. Appl.* **70** 31-58.

EINMAHL, J.H.J., DE HAAN, L. and HUANG, XIN (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. *Stoch. Proc. Appl.* **70** 143-171.

DE HAAN, L. and RESNICK, S. (1977) Limit theory for multidimensional sample extremes. *Z. Wahrscheinlichkeitstheorie 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. (1997) Estimation of the probability of the failure probability. Technical report, Erasmus University Rotterdam.

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

HUANG, XIN (1992) *Statistics of bivariate extremes.* Thesis, Erasmus University Rotterdam, Tinbergen Institute Research. series no. 22.

PICKANDS, J. (1981) Multivariate extreme value distributions. *Proceedings 43rd session, International Statistical Institute.* Buenos Aires, Argentina. Book 2, 859-878.

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.

Department of Mathematics
and Computing Science
Eindhoven University
of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands
Email: einmahl@win.tue.nl

Econometric Institute
Erasmus University
P. O. Box 1738
3000 DR Rotterdam
The Netherlands
Email: ldhaan@few.eur.nl

Laboratory of Probability
Faculty of Mechanics
and Mathematics
Moscow State University
Vorobjovy Gory
Moscow 119899 Russia
Email: piter@mech.math.msu.su

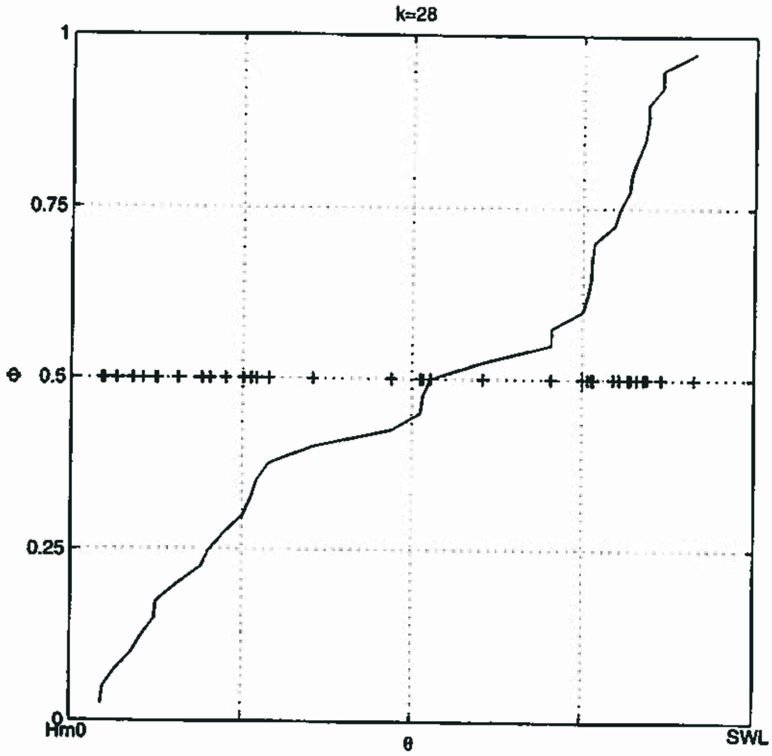
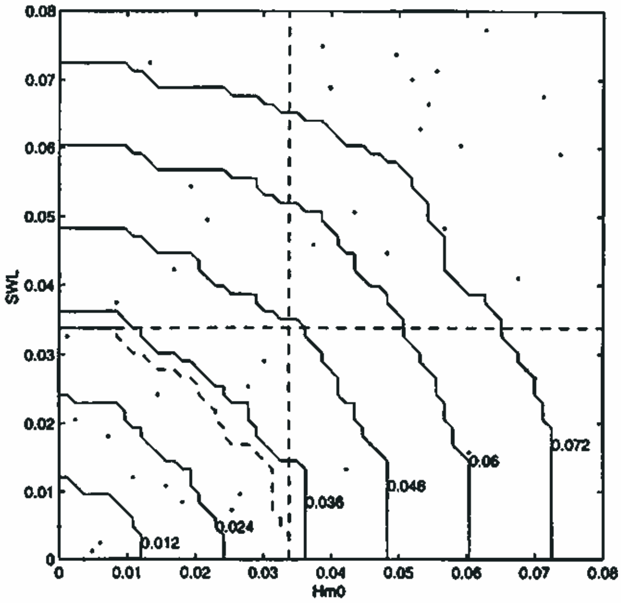


Figure 1: Estimated spectral measure. The solid line represents the estimated distribution function $\Phi(\theta)$, scaled down from 39/28 to 1.



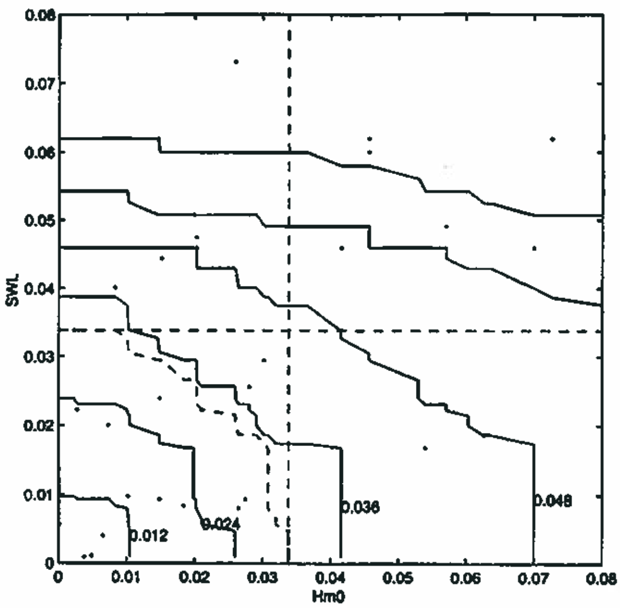


Figure 2: Level sets of the estimated function $\Lambda\{([x,\infty)\times[y,\infty))^c\}$. The top picture shows the level sets estimated in a non-parametric way. The bottom picture the same but estimated in a semi-parametric way.