The Annals of Statistics
1999, Vol. 27, No. 2, 732–759

ESTIMATING THE PROBABILITY OF A RARE EVENT[1]

BY LAURENS DE HAAN AND ASHOKE KUMAR SINHA

Erasmus University Rotterdam

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be a random sample from a bivariate distribution function F which is in the domain of attraction of a bivariate extreme value distribution function G. A subset C of $\mathbb{R}^2$ is given, which contains none of the observations. We shall give an asymptotic confidence interval for $\Pr((X_i, Y_i) \in C)$ under certain conditions.

1. Introduction. Extreme wave height and still water level are two very important factors for causing floods along a seacoast. Figure 1 shows the wave height and still water level during 828 independent storm events, recorded along the Dutch coast. The shaded area in the figure represents a possible failure area; any observation falling in this area is dangerous for the "Pettemer zeedijk," a sea-dike near the town of Petten. Our main problem is to estimate the probability that a future storm can cause a wave height and still water level combination which falls in this failure region. We also want to construct a confidence interval for the failure probability. For further details on the data and the failure region see, for example, reports on the Neptune Project in http://www.few.eur.nl/few/people/ldehaan/neptune.htm. An expository paper about this problem and the underlying theory is de Haan and de Ronde (1998).

We can formulate the problem mathematically as follows: let $\{(X_i, Y_i); i = 1, \ldots, n\}$ be a sample from the bivariate distribution function F. Suppose $C \in \mathscr{B}(\mathbb{R}^2)$. On the basis of the sample, we want to estimate $p := \Pr((X_1, Y_1) \in C)$, the failure probability.

The set C is such that none of the sample points falls into the region C, so we cannot use the empirical distribution function to estimate p. To estimate, we shall use some extra condition on the distribution function F in the field of extreme value theory. Since none of the observations fall into C, in first approximation the probability p must be less than $1/n$. Now the fact that none of the observations is close to the failure region is an essential feature of the problem, which we want to retain when applying asymptotic theory as we will do. Thus the inequality $n \Pr((X, Y) \in C) < 1$ forces us to assume that in fact, when applying asymptotic theory, the set C depends on n (notation: C_n) and that the sequence $p_n := \Pr((X, Y) \in C_n)$ tends to zero as $n \to \infty$. We go

Received September 1997; revised September 1998.
[1] Supported in part by the Neptune Project MAS2-CT94-0081.
AMS 1991 *subject classifications*. 62H10, 62P99
Key words and phrases. Failure region, failure chance, empirical process, estimation, functional central limit theorem, multivariate extremes, Vapnik–Cervonenkis class.

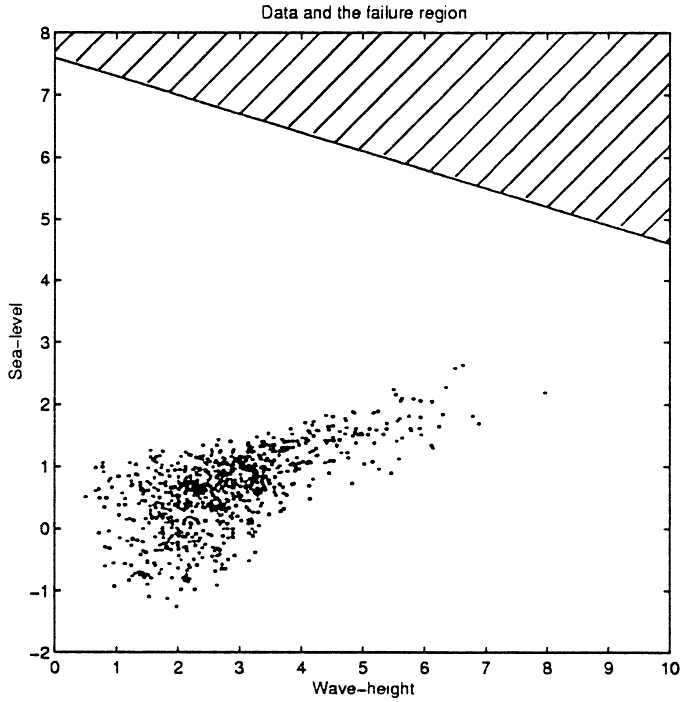

FIG. 1. *Data and failure region.*

one step further. The failure region in Figure 1 is in fact determined by a function: its functional form is

$$\{(s,t)\colon 0.3s + t \geq 7.6\}.$$

Thus in the rest of the paper we shall assume that

$$C_n = \{(s,t)\colon f_n(s,t) \geq 1\}$$

and, more specifically, that

$$f_n(s,t) = f\left(\frac{s}{x_n}, \frac{t}{y_n}\right),$$

where x_n and y_n are some positive numbers and f is some fixed known function.

We now explain the mathematical framework in which we work and give a sketch of how we shall solve the problem.

Assume that F is in the domain of attraction of some bivariate extreme value distribution function G [notation: $F \in \mathscr{D}(G)$], that is, there exist sequences $a_j(n) > 0$ and $b_j(n) \in \mathbb{R}$; $j = 1, 2$, so that

$$\lim_{n\to\infty} F^n(a_1(n)x + b_1(n), a_2(n)y + b_2(n)) = G(x,y) \tag{1.1}$$

for all continuity points $(x, y) \in \mathbb{R}^2$. We shall choose normalizing constants $a_j(n)$ and $b_j(n)$ in such a way that the jth marginal distribution of G will have the form $\exp\{-(1+\gamma_j x)^{-1/\gamma_j}\}$ whenever $1 + \gamma_j x > 0$ where $\gamma_j \in \mathbb{R}$, $j = 1, 2$.

It is known that G satisfying (1.1) is a *max-stable distribution function*, that is, for all $n \geq 1$, there exist constants $\tilde{a}_1(n), \tilde{a}_2(n) > 0, \tilde{b}_1(n), \tilde{b}_2(n) \in \mathbb{R}$ such that,

$$G^n\left(\tilde{a}_1(n)x + \tilde{b}_1(n), \tilde{a}_2(n)y + \tilde{b}_2(n)\right) = G(x, y)$$

[see de Haan and Resnick (1977) for this as well as other results up to (1.3) below]. A max-stable distribution in $\mathbb{R}^2$ is called a *simple extreme value distribution* if each marginal is equal to the extreme value distribution function $\Phi_1(x) = e^{-1/x}\mathbf{I}(x > 0)$ where $\mathbf{I}(\cdot)$ stands for the indicator function. We denote $G((x^{\gamma_1} - 1)/\gamma_1, (y^{\gamma_2} - 1)/\gamma_2)$ by $G_0(x, y)$. Note that G_0 is a *simple extreme value distribution function*. Hence [see de Haan and Resnick (1977)], there is a measure ν concentrating on $[0, \infty]^2 \setminus \{(0, 0)\}$ such that,

1. ν is a finite measure on compact subsets of $[0, \infty]^2 \setminus \{(0, 0)\}$;
2. $G_0(x, y) = \exp\{-\nu(\{(0, x] \times (0, y]\}^c)\}$.

Moreover,

$$s\nu(sB) = \nu(B); \qquad \forall B \in \mathscr{B}\left([0, \infty]^2 \setminus \{(0, 0)\}\right) \tag{1.2}$$

for $s > 0$, provided B is bounded away from the origin. This measure ν is called the *exponent measure* of the distribution G_0.

Let $\mathbf{T}$: $[0, \infty]^2 \setminus \{(0, 0)\} \to (0, \infty] \times [0, \pi/2]$ be the transformation such that

$$\mathbf{T}(x, y) = (r, \omega) \quad \text{with } r = \left(x^2 + y^2\right)^{1/2}, \omega = \arctan\left(\frac{y}{x}\right).$$

Now from here we can show that G_0 is a simple stable distribution function with exponent measure ν iff there exists a finite measure Φ on $[0, \pi/2]$ such that

$$\nu \circ \mathbf{T}^{-1}(dr, d\omega) = r^{-2}\, dr\, \Phi(d\omega),$$

and

$$\int_{[0, \pi/2]} \cos\omega\, \Phi(d\omega) = 1 = \int_{[0, \pi/2]} \sin\omega\, \Phi(d\omega).$$

The measure Φ is called the *spectral measure* or *angular measure* of the distribution G_0.

Now (1.1) is equivalent to

$$\lim_{n\to\infty} n\left[1 - F(a_1(n)x + b_1(n), a_2(n)y + b_2(n))\right] = -\log G(x, y).$$

We can replace n by a continuous variable $t > 0$ and get

$$\lim_{t\to\infty} t\left[1 - F(a_1(t)x + b_1(t), a_2(t)y + b_2(t))\right] = -\log G(x, y),$$

that is,

$$\lim_{t\to\infty} t \Pr\left(\frac{X_1 - b_1(t)}{a_1(t)} > x \text{ or } \frac{Y_1 - b_2(t)}{a_2(t)} > y \right) = -\log G(x, y).$$

Therefore for all $(x, y) \in \mathbb{R}^2_+$,

$$\lim_{t\to\infty} t \Pr\left(\left[1 + \gamma_1 \frac{X_1 - b_1(t)}{a_1(t)} \right]^{1/\gamma_1} > x, \text{ or } \left[1 + \gamma_2 \frac{Y_1 - b_2(t)}{a_2(t)} \right]^{1/\gamma_2} > y \right)$$
$$= -\log G\left(\frac{x^{\gamma_1} - 1}{\gamma_1}, \frac{y^{\gamma_2} - 1}{\gamma_2} \right) = -\log G_0(x, y) = \nu\left(\{[0, x] \times [0, y]\}^c \right).$$

Thus we have that (1.1) is equivalent to the vague convergence of measures on $[0,\infty]^2 \setminus \{(0,0)\}$, that is,

$$\lim_{t\to\infty} t \Pr\left(\left(\left[1 + \gamma_1 \frac{X_1 - b_1(t)}{a_1(t)} \right]^{1/\gamma_1}, \left[1 + \gamma_2 \frac{Y_1 - b_2(t)}{a_2(t)} \right]^{1/\gamma_2} \right) \in \cdot \right) = \nu(\cdot). \tag{1.3}$$

We denote for all $i = 1, 2, \dots, n$,

$$\tilde{X}_i(t) := \left[1 + \gamma_1 \frac{X_i - b_1(t)}{a_1(t)} \right]^{1/\gamma_1}, \qquad \tilde{Y}_i(t) := \left[1 + \gamma_2 \frac{Y_i - b_2(t)}{a_2(t)} \right]^{1/\gamma_2}.$$

So for $0 < k \le n$, (1.3) gives us

$$\lim_{n\to\infty} \frac{n}{k} \Pr\left(\left(\tilde{X}_1\left(\frac{n}{k}\right), \tilde{Y}_1\left(\frac{n}{k}\right) \right) \in B \right) = \nu(B) \tag{1.4}$$

as $n \to \infty$, $k \to \infty$ and $k/n \to 0$; $\forall B \in \mathscr{B}([0,\infty]^2 \setminus \{(0,0)\})$ provided B is bounded away from the origin and $\nu(\partial B) = 0$. We shall assume $k \to \infty$ and $k/n \to 0$ throughout the paper.

In view of the theory just developed involving the transformation from (X, Y) to $(\tilde{X}(n/k), \tilde{Y}(n/k))$, it seems natural to apply the same transformation to the failure region C_n. In fact the problem becomes tractable with our method only if we assume that the set C_n is of the following form: there exist a positive sequence $\{c_n\}$ (with $c_n \to \infty$) and a measurable set A in $[0,\infty]^2 \setminus \{(0,0)\}$, bounded away from the origin so that

$$C_n = a\left(\frac{n}{k}\right) \frac{(c_n A)^{\gamma} - \mathbf{1}}{\gamma} + \mathbf{b}\left(\frac{n}{k}\right). \tag{1.5}$$

Here $\mathbf{a}(t) := (a_1(t), a_2(t)), \mathbf{b}(t) := (b_1(t), b_2(t))$, $\gamma = (\gamma_1, \gamma_2)$ and all operations (addition, multiplication and taking powers) are componentwise.

Thus we have

$$\begin{aligned}
p_n &:= \Pr((X,Y) \in C_n) \\
&= \Pr\left((X,Y) \in \mathbf{a}\left(\frac{n}{k}\right)\frac{(c_n A)^{\gamma} - \mathbf{1}}{\gamma} + \mathbf{b}\left(\frac{n}{k}\right)\right) \\
&= \Pr\left(\left(\tilde{X}\left(\frac{n}{k}\right), \tilde{Y}\left(\frac{n}{k}\right)\right) \in c_n A\right) \\
&\approx \frac{k}{n}\nu(c_n A) \qquad [\text{by } (1.4)] \\
&= \frac{k}{nc_n}\nu(A) \qquad [\text{by } (1.2)].
\end{aligned}$$

So we shall try to estimate p_n by $(k/nc_n)\nu(A)$. However, we do now know c_n, A and ν. Therefore we propose

$$\hat{p}_n := \frac{k}{n\hat{c}_n}\hat{\nu}_n(\hat{A}) \tag{1.6}$$

as the estimator of p_n where $\hat{\nu}_n$, $\hat{c}_n$ and $\hat{A}$ are the estimators of ν, c_n and A, respectively, which are defined as follows.

We shall use some specific estimators of $a_1(\cdot)$, $a_2(\cdot)$, $b_1(\cdot)$, $b_2(\cdot)$, γ_1 and γ_2 studied by Dekkers, Einmahl and de Haan (1989). Let $X_{(1,n)} \le X_{(2,n)} \le \cdots \le X_{(n,n)}$ and $Y_{(1,n)} \le Y_{(2,n)} \le \cdots \le Y_{(n,n)}$ be the order statistics. We define the estimators as follows: define functions, $\bar{t} = t \wedge 0$ and

$$\rho_1(t) = \frac{1}{1-\bar{t}}; \qquad \rho_2(t) = \frac{2}{(1-\bar{t})(1-2\bar{t})},$$

and for $r = 1, 2$,

$$M_r(X) := \frac{1}{k}\sum_{i=0}^{k-1}\left\{\log X_{(n-i,n)} - \log X_{(n-k,n)}\right\}^r.$$

Now we define estimators as

$$\hat{\gamma}_1 := M_1(X) + 1 - \frac{1}{2}\left(1 - \frac{M_1(X)^2}{M_2(X)}\right)^{-1}, \tag{1.7}$$

$$\hat{b}_1\left(\frac{n}{k}\right) := X_{(n-k,n)}, \tag{1.8}$$

$$\hat{a}_1\left(\frac{n}{k}\right) := \frac{X_{(n-k,n)}\sqrt{3M_1(X)^2 - M_2(X)}}{\sqrt{3(\rho_1(\hat{\gamma}_1))^2 - \rho_2(\hat{\gamma}_1)}}. \tag{1.9}$$

Then $\hat{\gamma}_2$, $\hat{b}_2$ and $\hat{a}_2$ are defined in the same way by replacing X with Y.

Next we denote for $i = 1, 2, \dots, n$,

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

$$\hat{Y}_i\left(\frac{n}{k}\right) := \left[1 + \hat{\gamma}_2 \frac{Y_i - \hat{b}_2(n/k)}{\hat{a}_2(n/k)}\right]^{1/\hat{\gamma}_2}.$$

We shall use the estimator $\hat{\nu}_n$ as suggested by de Haan and Resnick (1993), which is defined as

$$\hat{\nu}_n(\cdot) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\left\{\left(\hat{X}_i\left(\frac{n}{k}\right), \hat{Y}_i\left(\frac{n}{k}\right)\right) \in \cdot\right\}. \tag{1.10}$$

Next we consider c_n and A. The role of the sequence $\{c_n\}$ is illustrated later in Figures 2 and 3: $\nu(A)$ can be estimated via the empirical measure but not $\nu(c_n A)$. Since c_n and A are not uniquely determined, we have to make a choice; our choice is to determine A such that $\min\{s\colon (s,s) \in A\} = 1$, that is, the point $(1,1)$ is on the boundary of A. Before proceeding further, we would like to introduce some notation in order to avoid cumbersome expressions. We denote

$$\tilde{f}_n(s,t) = f\left(\frac{1}{x_n}\left\{a_1\left(\frac{n}{k}\right)\frac{s^{\gamma_1}-1}{\gamma_1} + b_1\left(\frac{n}{k}\right)\right\}, \frac{1}{y_n}\left\{a_2\left(\frac{n}{k}\right)\frac{t^{\gamma_2}-1}{\gamma_2} + b_2\left(\frac{n}{k}\right)\right\}\right),$$

$$\hat{f}_n(s,t) = f\left(\frac{1}{x_n}\left\{\hat{a}_1\left(\frac{n}{k}\right)\frac{s^{\hat{\gamma}_1}-1}{\hat{\gamma}_1} + \hat{b}_1\left(\frac{n}{k}\right)\right\}, \frac{1}{y_n}\left\{\hat{a}_2\left(\frac{n}{k}\right)\frac{t^{\hat{\gamma}_2}-1}{\hat{\gamma}_2} + \hat{b}_2\left(\frac{n}{k}\right)\right\}\right).$$

So from (1.5) we get that

$$A = \frac{1}{c_n}\left[\mathbf{1} + \gamma\frac{C_n - \mathbf{b}}{\mathbf{a}}\right]^{\mathbf{1}/\gamma} = \left\{(s,t) \mid \tilde{f}_n(c_n s, c_n t) \ge 1\right\}. \tag{1.11}$$

According to our choice of A, the point $(1,1)$ must be on the boundary of A. So c_n must be the solution of $\tilde{f}_n(s,s) = 1$. This determines $c_n = s(\mathbf{a}, \mathbf{b}, \gamma)$. However, we do not know $(\mathbf{a}, \mathbf{b}, \gamma)$. So we can only find $\hat{c}_n = s(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\gamma})$, which is the solution of

$$\hat{f}_n(s,s) = 1. \tag{1.12}$$

Geometrically we can explain $\hat{c}_n$ as the intersecting point of $\hat{f}_n(s,t) = 1$ and the straight line $s = t$.

Finally, we define $\hat{A}$ as follows:

$$\hat{A} := \frac{1}{\hat{c}_n}\left[\mathbf{1} + \hat{\gamma}\frac{C_n - \hat{\mathbf{b}}}{\hat{\mathbf{a}}}\right]^{\mathbf{1}/\hat{\gamma}}.$$

To construct the confidence interval for p_n, we also need an estimator of the spectral measure. There are two consistent estimators available for the spectral measure which are defined as follows.

Semiparametric estimator:

$$\hat{\Phi}_n(\cdot) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\left\{\left(\hat{X}_i\left(\frac{n}{k}\right), \hat{Y}_i\left(\frac{n}{k}\right)\right):\ \left(\hat{X}_i\left(\frac{n}{k}\right)\right)^2 + \left(\hat{Y}_i\left(\frac{n}{k}\right)\right)^2 > 1, \right. \tag{1.13}$$
$$\left. \arctan\left(\hat{Y}_i\left(\frac{n}{k}\right)\Big/\hat{X}_i\left(\frac{n}{k}\right)\right) \in \cdot\right\}$$

[Einmahl, de Haan and Sinha (1997)].

Nonparametric estimator:

$$\hat{\Phi}_n(\cdot) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i):\ (n - R_{X_i})^{-2} + (n - R_{Y_i})^{-2} > k^{-2}, \right. \tag{1.14}$$
$$\left. \arctan\left(\frac{n - R_{X_i}}{n - R_{Y_i}}\right) \in \cdot\right\}$$

[Einmahl, de Haan and Piterbarg (1998)], where R_{X_i} and R_{Y_i} stand for $\text{rank}(X_i)$ and $\text{rank}(Y_i)$, respectively.

Without any ambiguity, we shall denote the distribution function generated by the *spectral measure* Φ, that is, $\Phi([0, \theta])$, by $\Phi(\theta)$. Similarly $\hat{\Phi}_n([0, \theta])$ will be denoted by $\hat{\Phi}_n(\theta)$.

To prove the asymptotic normality, we need to assume certain conditions apart from (1.1) and (1.5). In Section 2, we mention these conditions. One of the conditions is the Vapnik–Cervonenkis (VC) property. Some examples of collections of sets which satisfy this property are mentioned in Appendix B.

Section 3 states some lemmas with proofs which are necessary to prove the main theorem. The main theorem is stated and proved in Section 4.

Section 5 deals with the construction of the asymptotic confidence interval for p_n. Section 6 discusses an application of the main result to the wave height and still water level data near the Pettemer zeedijk.

We have used only the two-dimensional set-up for the proof. An extension to higher dimension is immediate.

The estimation of failure probability is actually a generalization of estimating the exceedance probability in higher dimension, where we have sets of the form $C_n = \{(s, t):\ s/x_n \vee t/y_n \geq 1\}$. The estimation procedure is more or less the same but the proof of asymptotic normality is quite different. It is much more difficult in the case of failure probability than in the exceedance probability case. The estimation of exceedance probability and its asymptotic normality have been discussed in de Haan (1994), and Sinha (1997), page 11. The one-dimensional version of the exceedance probability has been discussed in Dijk and de Haan (1992).

An alternative approach to the problem uses a much more restrictive, in fact, purely parametric, model [cf. for example, Smith, Tawn and Yuen (1990), Coles and Tawn (1991), Joe, Smith and Weissman (1992), Coles and Tawn (1994)]. It is assumed that the underlying distribution function on an interval of the form $(x_0, \infty) \times (y_0, \infty)$ actually coincides with a multidimensional extreme value distribution (i.e., the underlying distribution function equals the limit distribution function on a tail set). Moreover the exponent measure is modelled as

$$\nu\big(\{(0,x]\times(0,y]\}^c\big) = \big(x^{-1/\alpha}+y^{-1/\alpha}\big)^{\alpha},$$

where $0 \le \alpha \le 1$ or variants of this formula. Since the model is completely specified, the probability of any set can be calculated, including the failure set. Due to the parametric context, the parameters can be estimated by the maximum likelihood method.

In contrast, our conditions, as detailed in Section 2, rather than requiring equality of the two sides of (1.1) for any n and some set of (x, y)-values, require mainly a polynomial convergence rate in (1.1) [cf., e.g., (2.3)] as well as various conditions restricting the sequence $\{k(n)\}$ [(2.5) and (2.6)] in order to avoid bias and restrictions on the boundary of the failure region [(2.7), (2.9) and (2.10)].

Finally, the methods we have sketched above are valid only when the limiting extreme value distribution does not have independent components. Otherwise the methods have to be refined [see Ledford and Tawn (1997)]. In the problem under study there does not seem to be asymptotic independence [cf. de Haan and de Ronde (1998), Figures 4 and 5].

2. Conditions. Let us consider the following class of sets in $\mathbb{R}^2$: $\mathscr{C} := \{C_n\colon n \ge 1\}$ and $\mathscr{G} := \{l_1 C_n + l_2 \mid l_1, l_2 \in \mathbb{R}^2, C_n \in \mathscr{C}\}$.

However, we need to restrict our $\mathscr{G}$ so that the class $\mathscr{A} = \{[\mathbf{1}+\gamma S]^{\mathbf{1}/\gamma}\colon S \in \mathscr{G}\}$ should not contain any element for which the ν-measure is infinite. Hence we assume that the class

$$\mathscr{G}_* = \left\{S \in \mathscr{G}\colon [1+\gamma S]^{1/\gamma} \subset [0,\infty]^2 \setminus \left[0, \frac{1}{2}\right]^2\right\} \text{ is a VC class.} \tag{2.1}$$

We also assume that the sets of $\mathscr{G}_*$ satisfy condition (SE) of Gaenssler (1983).

Assume that

$$\sup_{S \in \mathscr{G}'} \left| \frac{n}{k} \Pr\left(\left(\tilde{X}_i\left(\frac{n}{k}\right), \tilde{Y}_i\left(\frac{n}{k}\right)\right) \in [\mathbf{1}+\gamma \mathbf{S}]^{\mathbf{1}/\gamma}\right) - \nu\big([\mathbf{1}+\gamma\mathbf{S}]^{\mathbf{1}/\gamma}\big)\right| \to 0, \tag{2.2}$$

where $\mathscr{G}' = \{S \cap S'\colon S, S' \in \mathscr{G}_*\}$.

Suppose that

$$\begin{aligned} &t\left[1 - F\left(a_1(t)\frac{x^{\gamma_1}-1}{\gamma_1} + b_1(t), a_2(t)\frac{y^{\gamma_2}-1}{\gamma_2} + b_2(t)\right)\right] \\ &\qquad - \nu\big(\{[0,x]\times[0,y]\}^c\big) = O(d(t)) \end{aligned} \tag{2.3}$$

locally uniformly for $(x, y) \in (0, \infty]^2$, where $d(t)$ is a regularly varying function and $d(t) \to 0$ as $t \to \infty$. In fact, often there exists a function d such that the left-hand side of (2.3) divided by $d(t)$ converges for each x and y as $t \to \infty$ [cf. de Haan and Resnick (1993)].

Next we define a transformation Q_γ: $\mathbb{R}_+^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathscr{B}([0,\infty]^2 \setminus \{(0,0)\}) \to [0,1]$,

$$Q_\gamma(\alpha,\beta,\eta,B) = \Pr\left(\left(\tilde{X}_i\left(\frac{n}{k}\right), \tilde{Y}_i\left(\frac{n}{k}\right)\right) \in \left[\mathbf{1} + \gamma\left(\alpha\frac{B^\eta - \mathbf{1}}{\eta} + \beta\right)\right]^{\mathbf{1}/\gamma}\right),$$

and we impose the following condition:

$$\left\{\frac{n}{k}Q_\gamma(\alpha_n,\beta_n,\eta_n,B) - \nu\left(\left[\mathbf{1} + \gamma\left(\alpha_n\frac{B^{\eta_n} - 1}{\eta_n} + \beta_n\right)\right]^{\mathbf{1}/\gamma}\right)\right\} = O\left(d\left(\frac{n}{k}\right)\right) \tag{2.4}$$

for all sequences $(\alpha_n, \beta_n, \eta_n) = (\mathbf{1}, \mathbf{0}, \gamma) + O(k^{-1/2})$ uniformly in $B \in \mathscr{B}([0,\infty]^2 \setminus \{(0,0)\})$.

Next we define for $x > 0$, the function

$$q_\gamma(x) := x^{-\gamma}\left(\int_1^x u^{\gamma-1}(\log u)\,du\right)$$

(see Appendix A for more details about the function q_γ). We assume the following conditions about the growth of the sequence c_n:

$$\lim_{n\to\infty}\frac{\sqrt{k}}{q_{\gamma_1\wedge\gamma_2}(c_n)} = \infty \tag{2.5}$$

and

$$\lim_{n\to\infty}\sqrt{k}\,\frac{c_n d(n/k)}{q_{\gamma_1\wedge\gamma_2}(c_n)} = 0. \tag{2.6}$$

Assume that f has first-order partial derivatives $f^{(1)}$ and $f^{(2)}$. We also assume that

$$\frac{f^{(1)}}{f^{(2)}} \ge 0. \tag{2.7}$$

This actually implies that if (x,y) satisfies $f(x,y) = 1$, then y is a nonincreasing function of x, because $dy/dx = -f^{(1)}/f^{(2)} \le 0$. Or in other words, the boundary of the failure region C_n, that is, the set $\{(s,t)\colon f(s/x_n, t/y_n) = 1\}$, is a decreasing function of the first variable.

We define

$$\xi_n(x,y) := \frac{a_2(n)}{a_1(n)}\,\frac{x_n}{y_n}(c_n)^{\gamma_2-\gamma_1}\frac{\tilde{f}_n^{(2)}(c_n x, c_n y)}{\tilde{f}_n^{(1)}(c_n x, c_n y)},$$

where

$$\tilde{f}_n^{(j)}(s,t) := f^{(j)}\left(\frac{1}{x_n}\left\{a_1\frac{s^{\gamma_1}-1}{\gamma_1} + b_1\right\}, \frac{1}{y_n}\left\{a_2\frac{s^{\gamma_2}-1}{\gamma_2} + b_2\right\}\right); \qquad j = 1, 2.$$

Suppose that

$$\lim_{n\to\infty} \xi_n(x, y) = \xi(x, y) \tag{2.8}$$

exists uniformly in $[0, \infty]^2$. Furthermore, (2.7) implies that $\xi(x, y) \geq 0$.

Let $\{(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)\colon \theta \in [0, \pi/2]\}$ be the boundary of A, that is, $\rho(\theta) = s(\mathbf{a}, \mathbf{b}, \gamma)$ is a solution of $\tilde{f}_n(c_n s\cos\theta, c_n s\sin\theta) = 1$. We assume that, given $\varepsilon > 0$,

$$\begin{aligned}\sup_{\theta\in[0,\pi/2]} \sup_{|s-\rho(\theta)|<\varepsilon} \bigg| \tilde{f}_n(c_n s\cos\theta, c_n s\sin\theta) \\ -f\Bigg(\frac{1}{x_n}\left\{\tilde{a}_1\frac{(c_n s\cos\theta)^{\tilde{\gamma}_1}-1}{\tilde{\gamma}_1}+\tilde{b}_1\right\}, \\ \frac{1}{y_n}\left\{\tilde{a}_2\frac{(c_n s\sin\theta)^{\tilde{\gamma}_2}-1}{\tilde{\gamma}_2}+\tilde{b}_2\right\}\Bigg)\bigg| \to 0,\end{aligned} \tag{2.9}$$

where $(\tilde{\mathbf{a}}/\mathbf{a}, (\tilde{\mathbf{b}}-\mathbf{b})/\mathbf{a}, \tilde{\gamma}) = (\mathbf{1}, \mathbf{0}, \gamma) + O(k^{-1/2})$. This assumption actually ensures that $\hat{\rho}(\theta) \to \rho(\theta)$ uniformly for $\theta \in [0, \pi/2]$, where $\{(\hat{\rho}(\theta)\cos\theta, \hat{\rho}(\theta)\sin\theta)\colon \theta\in[0,\pi/2]\}$ is the boundary of $(\hat{c}_n/c_n)\hat{A}$, that is, $\hat{\rho}(\theta) = s(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\gamma})$ is a solution of $\hat{f}_n(c_n s\cos\theta, c_n s\sin\theta) = 1$.

However, we need another assumption similar to (2.9). We assume that, given $\varepsilon > 0$,

$$\begin{aligned}\sup_{\theta\in[0,\pi/2]} \sup_{|s-\rho(\theta)|<\varepsilon} \Big| \tilde{f}_n(c_n s\cos\theta, c_n s\sin\theta) \\ -\tilde{f}_n(c_n R_1(s)\cos\theta, c_n R_2(s)\sin\theta)\Big| \to 0,\end{aligned} \tag{2.10}$$

where $R_j(x) = [1 + (\eta_j/\alpha_j)\{(x^{\gamma_j}-1)/\gamma_j - \beta_j\}]^{1/\eta_j}$ and $(\alpha, \beta, \gamma) = (\mathbf{1}, \mathbf{0}, \gamma) + O(k^{-1/2})$. This, along with (2.9), ensures that $\sup_{\theta\in[0,\pi/2]}|\hat{\hat{\rho}}(\theta) - \hat{\rho}(\theta)| \to 0$, where $\{(\hat{\rho}(\theta)\cos\theta, \hat{\rho}(\theta)\sin\theta)\colon \theta\in[0,\pi/2]\}$ is the boundary of $\hat{A}$ and $\{(\hat{\hat{\rho}}(\theta)\cos\theta, \hat{\hat{\rho}}(\theta)\sin\theta)\colon \theta\in[0,\pi/2]\}$ is the boundary of $[\mathbf{1} + \gamma\{(\hat{\mathbf{a}}/\mathbf{a})(\hat{A}^{\gamma} - \mathbf{1})/\gamma + (\hat{\mathbf{b}}-\mathbf{b})/\mathbf{a}\}]^{1/\gamma}$.

Suppose $\Phi(\cdot)$ is the spectral measure of ν and $\{(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)\colon \theta \in [0, \pi/2]\}$ is the boundary of A. We assume that there exists a δ $(0 < \delta < \pi/4)$ such that for $\gamma_1, \gamma_2 \neq 0$,

$$\begin{aligned}&\int_I \frac{(\rho(\theta)\cos\theta)^{-\gamma_1}}{\rho(\theta)}\Phi(d\theta) < \infty, &&\int_I \frac{(\rho(\theta)\sin\theta)^{-\gamma_2}}{\rho(\theta)}\Phi(d\theta) < \infty,\\ &\int_I \frac{|\log(\rho(\theta)\cos\theta)|}{\rho(\theta)}\Phi(d\theta) < \infty, &&\int_I \frac{|\log(\rho(\theta)\sin\theta)|}{\rho(\theta)}\Phi(d\theta) < \infty,\end{aligned} \tag{2.11}$$

and

$$(2.12)\qquad \begin{aligned} &\text{if } \gamma_1 = 0, \qquad \int_I \frac{(\log(\rho(\theta)\cos\theta))^2}{\rho(\theta)}\Phi(d\theta) < \infty,\\ &\text{if } \gamma_2 = 0, \qquad \int_I \frac{(\log(\rho(\theta)\sin\theta))^2}{\rho(\theta)}\Phi(d\theta) < \infty, \end{aligned}$$

where $I \in \{[0, \delta), (\pi/2 - \delta, \pi/2]\}$.

REMARK 2.1. Note that as $x \to \infty$, for $\gamma < 0$ we have $q_\gamma(x) \to x^{-\gamma}/\gamma^2$. Hence if $\gamma_1, \gamma_2 < 0$, the condition (2.5) can be written as $\lim_{n\to\infty} \sqrt{k}\, c_n^{\gamma_1 \wedge \gamma_2} = \infty$. Hence the conditions (2.3), (2.5) and (2.6) will be compatible when $\gamma_1, \gamma_2 > -\frac{1}{2}$.

3. Some lemmas.

LEMMA 3.1. *Suppose that* (1.1), (2.1), (2.2) *and the second-order condition* (2.3) *hold for the underlying distribution function F. Moreover, suppose that $k = k(n)$ is so that*

$$\lim_{n\to\infty} \sqrt{k}\, d\left(\frac{n}{k}\right) = 0.$$

Let W_ν be a bounded, uniformly continuous, zero-mean Gaussian process with

$$\operatorname{Cov}(W_\nu(B_1), W_\nu(B_2)) = \nu(B_1 \cap B_2).$$

Then there exist probabilistically equivalent versions of $\Delta_n, \hat{\mathbf{a}}, \hat{\mathbf{b}}$ and $\hat{\gamma}$ (which we denote, without any loss of generality, by the same symbol) such that they are defined on the same probability space as $W_\nu(\cdot)$ and

$$\left\{\Delta_n(\hat{A}), \sqrt{k}\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}} - \mathbf{1}, \frac{\hat{\mathbf{b}} - \mathbf{b}}{\mathbf{a}}, \hat{\gamma} - \gamma\right)\right\} \Rightarrow (0, \mathbf{A}, \mathbf{B}, \Gamma),$$

with

$$\Delta_n(\cdot) := \left|\sqrt{k}\left\{\hat{\nu}_n(\cdot) - \frac{n}{k} Q_\gamma\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}}, \frac{\hat{\mathbf{b}} - \mathbf{b}}{\mathbf{a}}, \hat{\gamma}, \cdot\right)\right\} - W_\nu(\cdot)\right|$$

and $(\mathbf{A}, \mathbf{B}, \Gamma)$ are expressed as functionals of the process W_ν as follows:

$$\begin{aligned} \mathbf{A} &= \gamma W_\nu(\mathbf{1}) + \frac{(\mathbf{1} - \overline{\gamma})^2(\mathbf{1} - 2\overline{\gamma})}{\mathbf{1} - 4\overline{\gamma}}\left(\frac{3\mathbf{P}}{\mathbf{1} - \overline{\gamma}} - \frac{\mathbf{Q}}{2}\right)\\ &\quad - \frac{3(\mathbf{1} - 2\overline{\gamma})^3 + (\mathbf{1} - \overline{\gamma})(\mathbf{1} - 2\overline{\gamma})^2 - 4(\mathbf{1} - \overline{\gamma})^3}{(\mathbf{1} - 4\overline{\gamma})^2(\mathbf{1} - \overline{\gamma})(\mathbf{1} - 2\overline{\gamma})}\Gamma,\\ \mathbf{B} &= W_\nu(\mathbf{1}),\\ \Gamma &= \left\{\gamma - \overline{\gamma} - 2(\mathbf{1} - \overline{\gamma})^2(\mathbf{1} - 2\overline{\gamma})\right\}\mathbf{P} + \frac{(\mathbf{1} - \overline{\gamma})^2(\mathbf{1} - 2\overline{\gamma})^2}{2}\mathbf{Q}, \end{aligned}$$

where $\bar{s} = s \wedge 0$, *and if the jth component of* W_ν *is* $W_{\nu j}$ $(j = 1, 2)$ *then the jth component of* $\mathbf{P}$ *is*

$$\int_1^\infty W_{\nu j}(s) \frac{ds}{s^{1-\bar{\gamma}_j}} - \frac{1}{1-\bar{\gamma}_j} W_{\nu j}(1),$$

and the jth component of $\mathbf{Q}$ *is*

$$2\left\{ \int_1^\infty W_{\nu j}(s) \frac{s^{\bar{\gamma}_j} - 1}{\bar{\gamma}_j} \frac{ds}{s^{1-\bar{\gamma}_j}} - \frac{1}{(1-\bar{\gamma}_j)(1-2\bar{\gamma}_j)} W_{\nu j}(1) \right\}.$$

PROOF. Define $\tilde{\nu}_n(\cdot) := (1/k)\sum_{i=1}^n \mathbf{I}\{(\tilde{X}_i(n/k), \tilde{Y}_i(n/k)) \in \cdot\}$. First, note that

$$\hat{\nu}(\hat{A}) = \tilde{\nu}_n\left(\left[\mathbf{1} + \gamma\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}} \frac{\hat{A}^{\hat{\gamma}} - \mathbf{1}}{\hat{\gamma}} + \frac{\hat{\mathbf{b}} - \mathbf{b}}{\mathbf{a}} \right) \right]^{\mathbf{1}/\gamma} \right) = \tilde{\nu}_n\left([\mathbf{1} + \gamma(l_1 C_n + l_2)]^{\mathbf{1}/\gamma} \right),$$

where $l_1 = (\mathbf{a}\hat{c}_n^{\hat{\gamma}})^{-1}$ and $l_2 = l_1[\hat{\mathbf{a}}(\mathbf{1} - \hat{c}_n^{\hat{\gamma}})/\hat{\gamma} - \hat{\mathbf{b}}] + (\hat{\mathbf{b}} - \mathbf{b})/\mathbf{a}$. Note that by (2.1), $\mathscr{C}$ is a VC class.

Then we proceed in the same way as in Proposition 3.2 of Einmahl, de Haan and Sinha (1997). □

LEMMA 3.2. *Under conditions* (1.1), (2.1), (2.2), (2.3) *and* (2.8),

$$\frac{\sqrt{k}}{q_{\gamma_1 \wedge \gamma_2}(c_n)} \left(\frac{\hat{c}_n}{c_n} - 1 \right) \to -\frac{\tau_1 S_1 + \xi(1,1)\tau_2 S_2}{1 + \xi(1,1)},$$

where τ_1, τ_2 *are as defined in* (A.1) *and with the notation* $\bar{\gamma} = \gamma \wedge 0$,

$$S_j = \Gamma_j - \bar{\gamma}_j A_j + \bar{\gamma}_j^2 B_j; \qquad j = 1, 2$$

and $(\mathbf{A}, \mathbf{B}, \Gamma)$ *are as defined in Lemma* 3.1.

REMARK 3.1. By (2.8) and (2.7), $\xi(1,1) \geq 0$ and so $1 + \xi(1,1) > 0$. Hence the limit is always defined.

PROOF. First recall that $c_n = s(\mathbf{a}, \mathbf{b}, \gamma)$ is the solution of $\tilde{f}_n(s, s) = 1$, and $\hat{c}_n = s(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\gamma})$ is the solution of $\hat{f}_n(s, s) = 1$.

Now

$$\begin{aligned} \hat{c}_n - c_n &= s(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\gamma}) - s(\mathbf{a}, \mathbf{b}, \gamma) \\ &\approx \sum_{j=1}^2 \left[(\hat{a}_j - a_j)\frac{\partial s}{\partial a_j} + (\hat{b}_j - b_j)\frac{\partial s}{\partial b_j} + (\hat{\gamma}_j - \gamma_j)\frac{\partial s}{\partial \gamma_j} \right]_{(\mathbf{a}, \mathbf{b}, \gamma)}. \end{aligned}$$

By the implicit function theorem

$$\frac{\partial s}{\partial \gamma_1} = -\left(\frac{\partial f}{\partial \gamma_1}\right)\Big/\left(\frac{\partial f}{\partial s}\right).$$

Here

$$\frac{\partial}{\partial s}\tilde{f}_n(s,s) = \frac{a_1}{x_n}s^{\gamma_1-1}\tilde{f}_n^{(1)}(s,s) + \frac{a_2}{y_n}s^{\gamma_2-1}\tilde{f}_n^{(2)}(s,s),$$

and

$$\frac{\partial}{\partial \gamma_1}\tilde{f}_n(s,s) = \frac{a_1}{x_n}\left(\int_1^s u^{\gamma_1-1}\log u\,du\right)\tilde{f}_n^{(1)}(s,s) = \frac{a_1}{x_n}s^{\gamma_1}q_{\gamma_1}(s)\tilde{f}_n^{(1)}(s,s).$$

Thus

$$\begin{aligned}\frac{\partial s}{\partial \gamma_1}\bigg|_{(\mathbf{a},\mathbf{b},\gamma)} &= -\frac{(a_1/x_n)s^{\gamma_1}q_{\gamma_1}(s)\tilde{f}_n^{(1)}(s,s)}{(a_1/x_n)s^{\gamma_1-1}\tilde{f}_n^{(1)}(s,s) + (a_2/y_n)s^{\gamma_2-1}\tilde{f}_n^{(2)}(s,s)}\bigg|_{(\mathbf{a},\mathbf{b},\gamma)}\\ &= -\frac{sq_{\gamma_1}(s)}{1+(a_2/a_1)(x_n/y_n)s^{\gamma_2-\gamma_1}\tilde{f}_n^{(2)}(s,s)/\tilde{f}_n^{(1)}(s,s)}\bigg|_{(\mathbf{a},\mathbf{b},\gamma)}\\ &= -\frac{c_n q_{\gamma_1}(c_n)}{1+\xi_n(1,1)}.\end{aligned}$$

Hence

$$\begin{aligned}\frac{\sqrt{k}}{c_n q_{\gamma_1\wedge\gamma_2}(c_n)}(\hat{\gamma}_1-\gamma_1)\frac{\partial s}{\partial \gamma_1}\bigg|_{(\mathbf{a},\mathbf{b},\gamma)} &= -\frac{\sqrt{k}(\hat{\gamma}_1-\gamma_1)}{c_n q_{\gamma_1\wedge\gamma_2}(c_n)}\frac{c_n q_{\gamma_1}(c_n)}{1+\xi_n(1,1)}\\ &\to -\frac{\tau_1\Gamma_1}{1+\xi(1,1)}.\end{aligned}$$

Similarly, one can show that

$$\begin{aligned}\frac{\sqrt{k}}{c_n q_{\gamma_1\wedge\gamma_2}(c_n)}(\hat{a}_1-a_1)\frac{\partial s}{\partial a_1}\bigg|_{(\mathbf{a},\mathbf{b},\gamma)} &= -\frac{\sqrt{k}(\hat{a}_1/a_1-1)}{c_n q_{\gamma_1\wedge\gamma_2}(c_n)}\frac{c_n((1-c_n^{-\gamma_1})/\gamma_1)}{1+\xi_n(1,1)}\\ &\to \frac{\tau_1\bar{\gamma}_1 A_1}{1+\xi(1,1)},\\ \frac{\sqrt{k}}{c_n q_{\gamma_1\wedge\gamma_2}(c_n)}(\hat{b}_1-b_1)\frac{\partial s}{\partial b_1}\bigg|_{(\mathbf{a},\mathbf{b},\gamma)} &= -\frac{\sqrt{k}\left((\hat{b}_1-b_1)/a_1\right)}{c_n q_{\gamma_1\wedge\gamma_2}(c_n)}\frac{c_n^{1-\gamma_1}}{1+\xi_n(1,1)}\\ &\to -\frac{\tau_1\bar{\gamma}_1^2 B_1}{1+\xi(1,1)}.\end{aligned}$$

Proceeding in the same way, we can find the limit of the other terms and combining all these we get the required result. □

COROLLARY 3.1. *Under the conditions of Lemma* 3.2 *and assumption* (2.5),

$$\frac{\hat{c}_n}{c_n} \to_P 1.$$

LEMMA 3.3. *Suppose the conditions* (1.1), (2.1)–(2.4), (2.8)–(2.12) *hold for the underlying distribution function F. Moreover, suppose that* $k = k(n)$ *is so that*

$$\lim_{n \to \infty} \sqrt{k}\, d\left(\frac{n}{k}\right) = 0.$$

Then there exists a probabilistically equivalent version of $\hat{\nu}_n$ (*which we denote, without any loss of generality, by the same symbol*) *such that*

$$\left|\sqrt{k}\left\{\hat{\nu}_n(\hat{A}) - \nu(\hat{A})\right\} - W_\nu(\hat{A}) - R(\mathbf{A}, \mathbf{B}, \Gamma)\right| \longrightarrow 0 \text{ a.s.},$$

where W_ν, $\mathbf{A}, \mathbf{B}, \Gamma$ *are as defined in Lemma* 3.1 *and*

$$R(\mathbf{A}, \mathbf{B}, \Gamma) = \int_0^{\pi/2} \frac{\{h(A_1, B_1, \Gamma_1, \rho(\theta)\cos\theta) + D(\theta) h(A_2, B_2, \Gamma_2, \rho(\theta)\sin\theta)\}}{\rho(\theta)[1 + D(\theta)]} \Phi(d\theta),$$

where for $j = 1, 2$,

$$h(A_j, B_j, \Gamma_j, s) = A_j \frac{s^{-\gamma_j} - 1}{\gamma_j} - B_j s^{-\gamma_j} - \Gamma q_{\gamma_j}(s),$$

$\Phi(\cdot)$ *is the spectral measure of* ν, $\{(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)\colon \theta \in [0, \pi/2]\}$ *is the boundary of A and* $D(\theta) := \xi(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)((\rho(\theta)\sin\theta)^{\gamma_2 + 1})/((\rho(\theta)\cos\theta)^{\gamma_1 + 1})$.

PROOF. First by invoking the Skorohod construction we can obtain a new probability space on which all the random elements of Lemma 3.1 are defined and where the weak convergence can be replaced by strong convergence. Thus we reduce the problem to an analytical problem. Without any ambiguity, we maintain the old notation for these new random elements.

Now

$$\begin{aligned}
\sqrt{k}\left[\hat{\nu}_n(\hat{A}) - \nu(\hat{A})\right]
&= \sqrt{k}\left[\hat{\nu}_n(\hat{A}) - Q_\gamma\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}}, \frac{\hat{\mathbf{b}} - \mathbf{b}}{\mathbf{a}}, \hat{\gamma}, \hat{A}\right)\right] \\
&\quad + \sqrt{k}\left[Q_\gamma\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}}, \frac{\hat{\mathbf{b}} - \mathbf{b}}{\mathbf{a}}, \hat{\gamma}, \hat{A}\right)\right.
\end{aligned}$$

$$-\nu\left(\left[\mathbf{1}+\gamma\left\{\frac{\hat{\mathbf{a}}}{\mathbf{a}}\frac{\hat{A}^{\gamma}-\mathbf{1}}{\gamma}+\frac{\hat{\mathbf{b}}-\mathbf{b}}{\mathbf{a}}\right\}\right]^{1/\gamma}\right)\Bigg]$$

$$+\sqrt{k}\left[\nu\left(\left[\mathbf{1}+\gamma\left\{\frac{\hat{\mathbf{a}}}{\mathbf{a}}\frac{\hat{A}^{\gamma}-\mathbf{1}}{\gamma}+\frac{\hat{\mathbf{b}}-\mathbf{b}}{\mathbf{a}}\right\}\right]^{1/\gamma}\right)-\nu(\hat{A})\right].$$

By Lemma 3.1,

$$\left|\sqrt{k}\left[\hat{\nu}_n(\hat{A})-Q_{\gamma}\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}},\frac{\hat{\mathbf{b}}-\mathbf{b}}{\mathbf{a}},\hat{\gamma},\hat{A}\right)\right]-W_{\nu}(\hat{A})\right|\to 0.$$

The second-order condition (2.4) will give us that

$$\sqrt{k}\left[Q_{\gamma}\left(\frac{\hat{\mathbf{a}}}{\mathbf{a}},\frac{\hat{\mathbf{b}}-\mathbf{b}}{\mathbf{a}},\hat{\gamma},\hat{A}\right)-\nu\left(\left[\mathbf{1}+\gamma\left\{\frac{\hat{\mathbf{a}}}{\mathbf{a}}\frac{\hat{A}^{\gamma}-1}{\gamma}+\frac{\hat{\mathbf{b}}-\mathbf{b}}{\mathbf{a}}\right\}\right]^{1/\gamma}\right)\right]\to 0.$$

Now we consider the third part. Notice $\hat{A} := (1/\hat{c}_n)[\mathbf{1}+\hat{\gamma}((C_n-\hat{\mathbf{b}})/\hat{\mathbf{a}})]^{\mathbf{1}/\hat{\gamma}} = \{(s,t)\colon \hat{f}_n(\hat{c}_n s,\hat{c}_n t)\ge 1\}$. If $\{(\hat{\rho}(\theta)\cos\theta, \hat{\rho}(\theta)\sin\theta)\colon \theta\in[0,\pi/2]\}$ is the boundary of $\hat{A}$, then it satisfies the equation $\hat{f}_n(\hat{c}_n s,\hat{c}_n t)=1$. Similarly, if $\{(\hat{\hat{\rho}}(\theta)\cos\theta, \hat{\hat{\rho}}(\theta)\sin\theta)\colon \theta\in[0,\pi/2]\}$ is the boundary of $[\mathbf{1}+\gamma\{(\hat{\mathbf{a}}/\mathbf{a})\times(\hat{A}^{\gamma}-\mathbf{1})/\gamma+(\hat{\mathbf{b}}-\mathbf{b})/\mathbf{a}\}]^{1/\gamma}$ then it satisfies the equation

$$\hat{f}_n\left(\hat{c}_n\left[1+\hat{\gamma}_1\frac{a_1}{\hat{a}_1}\left\{\frac{s^{\gamma_1}-1}{\gamma_1}-\frac{\hat{b}_1-b_1}{a_1}\right\}\right]^{1/\hat{\gamma}_1},\right.$$
$$\left.\hat{c}_n\left[1+\hat{\gamma}_2\frac{a_2}{\hat{a}_2}\left\{\frac{t^{\gamma_2}-1}{\gamma_2}-\frac{\hat{b}_2-b_2}{a_2}\right\}\right]^{1/\hat{\gamma}_2}\right)=1.$$

We have

$$\sqrt{k}\left[\nu\left(\left[\mathbf{1}+\gamma\left\{\frac{\hat{\mathbf{a}}}{\mathbf{a}}\frac{\hat{A}^{\hat{\gamma}}-\mathbf{1}}{\hat{\gamma}}+\frac{\hat{\mathbf{b}}-\mathbf{b}}{\mathbf{a}}\right\}\right]^{1/\gamma}\right)-\nu(\hat{A})\right]$$
$$=\sqrt{k}\left[\int_0^{\pi/2}\int_{\hat{\hat{\rho}}(\theta)}^{\infty}r^{-2}\,dr\,\Phi(d\theta)-\int_0^{\pi/2}\int_{\hat{\rho}(\theta)}^{\infty}r^{-2}\,dr\,\Phi(d\theta)\right]$$
$$=\sqrt{k}\int_0^{\pi/2}\left(\frac{\hat{\rho}(\theta)-\hat{\hat{\rho}}(\theta)}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\right)\Phi(d\theta)$$
$$=-\sum_{j=1}^{2}\left[\sqrt{k}\,(\hat{\gamma}_j-\gamma_j)\int_0^{\pi/2}\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\eta_j}\Phi(d\theta)\right.$$

$$+\sqrt{k}\left(\frac{\hat{a}_j}{a_j}-1\right)\int_0^{\pi/2}\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\alpha_j}\Phi(d\theta)$$

$$+\sqrt{k}\left(\frac{\hat{b}_j-b_j}{a_j}\right)\int_0^{\pi/2}\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\beta_j}\Phi(d\theta)\Bigg]_{(\alpha,\beta,\eta)},$$

where $(\alpha,\beta,\eta)=t(\hat{\mathbf{a}}/\mathbf{a},(\hat{\mathbf{b}}-\mathbf{b})/\mathbf{a},\hat{\gamma})+(1-t)(\mathbf{1},\mathbf{0},\gamma)$, for some $t\in(0,1)$.

Now proceeding as in Lemma 3.1 and by assumptions (2.8), (2.9), (2.10) we get that

$$\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\eta_j}\to\frac{1}{\rho(\theta)}\frac{q_{\gamma_1}(\rho(\theta)\cos\theta)}{[1+D(\theta)]}.$$

Moreover, for any $0<\delta<\pi/4$,

$$\sup_{\theta\in[\delta,\pi/2-\delta]}\left|\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\eta_j}\right|\le M$$

for some positive constant M. On the other hand, when $\theta\in[0,\delta)$, $\hat{\rho}(\theta)\cos\theta$ is close to 0. As $x\to 0$,

$$q_{\eta_1}\left(\left[1+\frac{\eta_1}{\alpha_1}\left\{\frac{x^{\gamma_1}-1}{\gamma_1}-\beta_1\right\}\right]^{1/\eta_1}\right)\sim\begin{cases}m_1x^{-\gamma_1}, & \text{if } \gamma_1>0,\\ (\log x)^2/2, & \text{if } \gamma_1=0,\\ m_2\log x+m_3, & \text{if } \gamma_1<0,\end{cases}$$

for some positive constants m_1, m_2, m_3. Using this we get, for $\gamma_1>0$,

$$\sup_{\theta\in[0,\delta)}\left|\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\eta_j}\right|\le M_1\frac{1}{\rho(\theta)}\left[1+(\rho(\theta)\cos\theta)^{-\gamma_1}\right],$$

for $\gamma_1=0$,

$$\sup_{\theta\in[0,\delta)}\left|\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\eta_j}\right|\le M_2\frac{1}{\rho(\theta)}(\log(\rho(\theta)\cos\theta))^2$$

and for $\gamma_1<0$,

$$\sup_{\theta\in[0,\delta)}\left|\frac{1}{\hat{\hat{\rho}}(\theta)\hat{\rho}(\theta)}\frac{\partial\hat{\rho}}{\partial\eta_j}\right|\le M_3\frac{1}{\rho(\theta)}(\log(\rho(\theta)\cos\theta)+1),$$

where M_1, M_2, M_3 are positive constants. Similarly, we can get uniform bounds for $\theta\in(\pi/2-\delta,\pi/2]$.

According to assumptions (2.11) and (2.12), these bounds are integrable. So, applying Lebesgue's dominated convergence theorem,

$$\sqrt{k}\,(\hat{\gamma}_1 - \gamma_1)\int_0^{\pi/2} \frac{1}{\hat{\hat{\rho}}(\theta)\,\hat{\rho}(\theta)}\,\frac{\partial\hat{\rho}}{\partial\eta_1}\Phi(d\theta)$$
$$\to -\Gamma_1\int_0^{\pi/2} \frac{q_{\gamma_1}(\rho(\theta)\cos\theta)}{\rho(\theta)(\rho(\theta)\cos\theta)^{\gamma_1}[1 + D(\theta)]}\Phi(d\theta).$$

Similarly, we get the other terms. □

LEMMA 3.4. *Suppose* (1.1), (2.1)–(2.3), (2.8), (2.10)–(2.12) *hold*, *then*

$$\frac{\sqrt{k}}{q_{\gamma_1\wedge\gamma_2}(c_n)}\left[\nu(\hat{A}) - \frac{\hat{c}_n}{c_n}\nu(A)\right]$$
$$\to -\tau_1 S_1\int_0^{\pi/2} (\rho(\theta)\cos\theta)^{(\gamma_1\vee 0)}$$
$$\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1}$$
$$+\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta)$$
$$-\tau_2 S_2\int_0^{\pi/2}\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{(\gamma_2\vee 0)}$$
$$\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1}$$
$$+\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta),$$

where $\{(\rho(\theta)\cos\theta,\ \rho(\theta)\sin\theta)\colon\ \theta\in[0,\pi/2]\}$ *is the boundary of* A.

The proof is almost the same as for Lemma 3.3.

4. The main theorem.

THEOREM 4.1. *Under conditions* (1.1), (1.5) *and* (2.1)–(2.12) *and if* $\lim_{n\to\infty}\sqrt{k}\,d(n/k) = 0$, *then*

$$\frac{\sqrt{k}}{q_{\gamma_1\wedge\gamma_2}(c_n)}\left(\frac{\hat{p}_n}{p_n} - 1\right)$$
$$\to -\frac{\tau_1 S_1}{\nu(A)}\int_0^{\pi/2}(\rho(\theta)\cos\theta)^{(\gamma_1\vee 0)}$$
$$\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1}$$
$$+\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta)$$

$$
\begin{aligned}
-\frac{\tau_2 S_2}{\nu(A)} \int_0^{\pi/2} &\xi(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{(\gamma_2 \vee 0)} \\
&\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1} \\
&\qquad + \xi(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1} \Phi(d\theta).
\end{aligned}
$$

In particular, the limit random variable is normal with mean zero.

PROOF. Using Lemmas 3.3 and 3.4 and Corollary 3.1, we get that $\hat{p}_n := (k/n\hat{c}_n)\hat{\nu}_n(\hat{A}) \sim (k/n\hat{c}_n)\nu(A)$.

$$
\begin{aligned}
\nu(A)\left(1 - \frac{p_n}{\hat{p}_n}\right) &= \frac{\nu(A)}{\hat{p}_n}(\hat{p}_n - p_n) \\
&\sim \frac{n\hat{c}_n}{k}\left(\frac{k}{n\hat{c}_n}\hat{\nu}_n(\hat{A}) - \Pr((X,Y) \in C_n)\right) \\
&= \left[\hat{\nu}_n(\hat{A}) - \nu(\hat{A})\right] + \left[\nu(\hat{A}) - \frac{\hat{c}_n}{c_n}\nu(A)\right] \\
&\quad - \hat{c}_n\left[\frac{n}{k}\Pr((X,Y) \in C_n) - \frac{1}{c_n}\nu(A)\right].
\end{aligned}
$$

Now by Lemma 3.3 we get that

$$
\sqrt{k}\left[\hat{\nu}_n(\hat{A}) - \nu(\hat{A})\right] \to W_\nu(A) + R(\mathbf{A}, \mathbf{B}, \Gamma).
$$

By conditions (2.4) and (2.6) we obtain

$$
\hat{c}_n\left[\frac{n}{k}\Pr((X,Y) \in C_n) - \frac{1}{c_n}\nu(A)\right] \to 0.
$$

Therefore, combining Lemmas 3.3 and 3.4, we get

$$
\begin{aligned}
&\frac{\sqrt{k}}{q_{\gamma_1 \wedge \gamma_2}(c_n)}\left(\frac{\hat{p}_n}{p_n} - 1\right) \\
&\quad = \left(\frac{\hat{p}_n}{p_n}\right)\frac{1}{\nu(A)}\frac{\sqrt{k}}{q_{\gamma_1 \wedge \gamma_2}(c_n)}\left[\frac{\nu(A)}{\hat{p}_n}(\hat{p}_n - p_n)\right] \\
&\quad \to -\frac{\tau_1 S_1}{\nu(A)}\int_0^{\pi/2}(\rho(\theta)\cos\theta)^{(\gamma_1 \vee 0)} \\
&\qquad\qquad \times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1} \\
&\qquad\qquad\qquad + \xi(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta)
\end{aligned}
$$

$$-\frac{\tau_2 S_2}{\nu(A)}\int_0^{\pi/2}\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{(\gamma_2\vee 0)}$$
$$\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1}$$
$$+\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta).$$

5. Confidence interval. Let us denote

$$\zeta_1 = -\frac{\tau_1}{\nu(A)}\int_0^{\pi/2}(\rho(\theta)\cos\theta)^{(\gamma_1\vee 0)}$$
$$\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1}$$
$$+\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta),$$

$$\zeta_2 = -\frac{\tau_2}{\nu(A)}\int_0^{\pi/2}\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{(\gamma_2\vee 0)}$$
$$\times\Big\{\rho(\theta)\big[(\rho(\theta)\cos\theta)^{\gamma_1}$$
$$+\xi(\rho(\theta)\cos\theta,\rho(\theta)\sin\theta)(\rho(\theta)\sin\theta)^{\gamma_2}\big]\Big\}^{-1}\Phi(d\theta).$$

So by our main theorem we get that

$$\frac{\sqrt{k}}{q_{\gamma_1\wedge\gamma_2}(c_n)}\left(\frac{\hat{p}_n}{p_n}-1\right)\Rightarrow \zeta_1 S_1+\zeta_2 S_2.$$

Notice that by assumption (2.5), $k^{-1/2}q_{\gamma_1\wedge\gamma_2}(c_n)\to 0$, so $(\hat{p}_n/p_n-1)\to 0$. Since $x\sim\log(1+x)$ as $x\to 0$, we have

$$\frac{\sqrt{k}}{q_{\gamma_1\wedge\gamma_2}(c_n)}\log\left(\frac{\hat{p}_n}{p_n}\right)\Rightarrow \zeta_1 S_1+\zeta_2 S_2.$$

Now S_1, S_2 are linear combinations of $(\mathbf{A},\mathbf{B},\Gamma)$, which are again some functionals of the bounded, uniformly continuous, zero-mean Gaussian process W_ν (see Lemma 3.1). Thus one can compute the variance of $k^{-1/2}q_{\gamma_1\wedge\gamma_2}(c_n)(\zeta_1 S_1+\zeta_2 S_2)$ and that will determine the asymptotic variance of $\log(\hat{p}_n/p_n)$, which we denote by $v_n(\gamma_1,\gamma_2,a_1,a_2,b_1,b_2,c_n,\Phi)$. However, this depends on unknown parameters $\gamma_1,\gamma_2,a_1,a_2,b_1,b_2,c_n$ and Φ. Nevertheless, $v_n(\gamma_1,\gamma_2,a_1,a_2,b_1,b_2,c_n,\Phi)$ can be estimated consistently by replacing the unknown parameters with their respective consistent estimators, that is, by $v_n(\hat{\gamma}_1,\hat{\gamma}_2,\hat{a}_1,\hat{a}_2,\hat{b}_1,\hat{b}_2,\hat{c}_n,\hat{\Phi}_n)=:\hat{v}_n$.

Thus we get

$$\lim_{n\to\infty}\Pr\Big[\hat{p}_n\exp\big(-\hat{v}_n^{-1/2}z_{\alpha/2}\big)<p_n<\hat{p}_n\exp\big(\hat{v}_n^{-1/2}z_{\alpha/2}\big)\Big]=1-\alpha,$$

where for $\alpha \in (0,1)$, z_α represents the $(1-\alpha)$th quantile of the standard normal distribution.

5.1. *Computation of* $\hat{v}_n$. First we compute $\mathrm{Var}(\zeta_1 S_1 + \zeta_2 S_2)$. We denote for $j = 1,2$,

$$I_j = \int_1^\infty W_{\nu j}(s)\frac{ds}{s^{1-\overline{\gamma}_j}}, \qquad I_{2+j} = \int_1^\infty W_{\nu j}(s)\frac{s^{\overline{\gamma}_j}-1}{\overline{\gamma}_j}\frac{ds}{s^{1-\overline{\gamma}_j}}.$$

So $(\mathbf{B},\mathbf{P},\mathbf{Q})' = \mathbf{M}_1(W_{\nu 1}(1), W_{\nu 2}(1), I_1, I_2, I_3, I_4)'$, where $\mathbf{M}_1 = ((m_1(i,j)))$ is a 6×6 matrix with $m_1(i,i) = 1$ if $1 \le i \le 4$ (2 if $i = 5,6$); $m_1(3,1) = -(1-\overline{\gamma}_1)^{-1}$, $m_1(4,2) = -(1-\overline{\gamma}_2)^{-1}$, $m_1(5,1) = -2[(1-\overline{\gamma}_1)(1-2\overline{\gamma}_1)]^{-1}$, $m_1(6,2) = -2[(1-\overline{\gamma}_2)(1-2\overline{\gamma}_2)]^{-1}$ and $m_1(i,j) = 0$ otherwise.

Now $(\mathbf{A},\mathbf{B},\Gamma)' = \mathbf{M}_2(\mathbf{B},\mathbf{P},\mathbf{Q})'$, where $\mathbf{M}_2 = ((m_2(i,j)))$ is a 6×6 matrix which is defined as follows: for $j=1,2$, $\vartheta_j = (1-\overline{\gamma}_j)^2(1-2\overline{\gamma}_j)(1-4\overline{\gamma}_j)^{-1}$; $t_j = [3(1-2\overline{\gamma}_j)^3 + (1-\overline{\gamma}_j)(1-2\overline{\gamma}_j)^2 - 4(1-\overline{\gamma}_j)^3][(1-4\overline{\gamma}_j)^2(1-\overline{\gamma}_j)(1-2\overline{\gamma}_j)]^{-1}$; $l_j = \gamma_j - \overline{\gamma}_j - 2(1-\overline{\gamma}_j)^2(1-2\overline{\gamma}_j)$, $\tilde{l}_j = \frac{1}{2}(1-\overline{\gamma}_j)^2(1-2\overline{\gamma}_j)^2$.

Then (for $j=1,2$), $m_2(j,j) = \gamma_j$; $m_2(4+j,4+j) = \tilde{l}_j$; $m_2(4+j,2+j) = l_j$; $m_2(j,2+j) = 3\vartheta_j(1-\overline{\gamma}_j)^{-1} - t_j l_j$; $m_2(j,4+j) = -\frac{1}{2}\vartheta_j - t_j\tilde{l}_j$; $m_2(2+j,j) = 1$. All the other terms are 0.

Finally, $(S_1,S_2)' = \mathbf{M}_3(\mathbf{A},\mathbf{B},\Gamma)'$ where

$$\mathbf{M}_3 = \begin{pmatrix} -\overline{\gamma}_1 & 0 & \overline{\gamma}^2 & 0 & 1 & 0 \\ 0 & -\overline{\gamma}_2 & 0 & \overline{\gamma}_2^2 & 0 & 1 \end{pmatrix}.$$

So

$$\mathrm{Var}(\zeta_1 S_1 + \zeta_2 S_2) = (\zeta_1,\zeta_2)\mathrm{Var}(S_1,S_2)(\zeta_1,\zeta_2)' \tag{5.1}$$

and $\mathrm{Var}(S_1,S_2) = \mathbf{M}_3\mathbf{M}_2\mathbf{M}_1\Sigma(\mathbf{M}_3\mathbf{M}_2\mathbf{M}_1)'$ where $\Sigma = \mathrm{Var}(W_{\nu 1}(1), W_{\nu 2}(1), I_1, I_2, I_3, I_4) := ((\sigma_{ij}))_{6\times 6}$.

Now

$$\sigma_{11} = \nu((1,\infty)\times[0,\infty)),$$
$$\sigma_{22} = \nu([0,\infty)\times(1,\infty)),$$
$$\sigma_{33} = \frac{2}{(1-\overline{\gamma}_1)(1-2\overline{\gamma}_1)}\nu((1,\infty)\times[0,\infty)),$$
$$\sigma_{44} = \frac{2}{(1-\overline{\gamma}_2)(1-2\overline{\gamma}_2)}\nu([0,\infty)\times(1,\infty)),$$
$$\sigma_{55} = \frac{6}{(1-\overline{\gamma}_1)(1-2\overline{\gamma}_1)(1-3\overline{\gamma}_1)(1-4\overline{\gamma}_1)}\nu((1,\infty)\times[0,\infty)),$$

$$\sigma_{66} = \frac{6}{(1-\bar{\gamma}_2)(1-2\bar{\gamma}_2)(1-3\bar{\gamma}_2)(1-4\bar{\gamma}_2)}\nu([0,\infty)\times(1,\infty)),$$

$$\sigma_{12} = \nu((1,\infty)\times(1,\infty)),$$

$$\sigma_{13} = \frac{1}{(1-\bar{\gamma}_1)}\nu((1,\infty)\times[0,\infty)),$$

$$\sigma_{14} = \int_1^\infty \nu((1,\infty)\times(s,\infty))s^{\bar{\gamma}_2-1}\,ds,$$

$$\sigma_{15} = \frac{1}{(1-\bar{\gamma}_1)(1-2\bar{\gamma}_1)}\nu((1,\infty)\times[0,\infty)),$$

$$\sigma_{16} = \int_1^\infty \nu((1,\infty)\times(s,\infty))\left(\frac{s^{\bar{\gamma}_2}-1}{\bar{\gamma}_2}\right)s^{\bar{\gamma}_2-1}\,ds,$$

$$\sigma_{23} = \int_1^\infty \nu((s,\infty)\times(1,\infty))s^{\bar{\gamma}_1-1}\,ds,$$

$$\sigma_{24} = \frac{1}{(1-\bar{\gamma}_2)}\nu([0,\infty)\times(1,\infty)),$$

$$\sigma_{25} = \int_1^\infty \nu((s,\infty)\times(1,\infty))\left(\frac{s^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right)s^{\bar{\gamma}_1-1}\,ds,$$

$$\sigma_{26} = \frac{1}{(1-\bar{\gamma}_2)(1-2\bar{\gamma}_2)}\nu([0,\infty)\times(1,\infty)),$$

$$\sigma_{34} = \int_1^\infty\int_1^\infty \nu((s,\infty)\times(t,\infty))s^{\bar{\gamma}_1-1}t^{\bar{\gamma}_2-1}\,ds\,dt,$$

$$\sigma_{35} = \frac{3}{(1-\bar{\gamma}_1)(1-2\bar{\gamma}_1)(1-3\bar{\gamma}_1)}\nu((1,\infty)\times[0,\infty)),$$

$$\sigma_{36} = \int_1^\infty\int_1^\infty \nu((s,\infty)\times(t,\infty))s^{\bar{\gamma}_1-1}\left(\frac{t^{\bar{\gamma}_2}-1}{\bar{\gamma}_2}\right)t^{\bar{\gamma}_2-1}\,ds\,dt,$$

$$\sigma_{45} = \int_1^\infty\int_1^\infty \nu((s,\infty)\times(t,\infty))\left(\frac{s^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right)s^{\bar{\gamma}_1-1}t^{\bar{\gamma}_2-1}\,ds\,dt,$$

$$\sigma_{46} = \frac{3}{(1-\bar{\gamma}_2)(1-2\bar{\gamma}_2)(1-3\bar{\gamma}_2)}\nu([0,\infty)\times(1,\infty)),$$

$$\sigma_{56} = \int_1^\infty\int_1^\infty \nu((s,\infty)\times(t,\infty))\left(\frac{s^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right)s^{\bar{\gamma}_1-1}\left(\frac{t^{\bar{\gamma}_2}-1}{\bar{\gamma}_2}\right)t^{\bar{\gamma}_2-1}\,ds\,dt.$$

Thus we get $v_n = k^{-1}(q_{\gamma_1\wedge\gamma_2}(c_n))^2\ \mathrm{Var}(\zeta_1 S_1 + \zeta_2 S_2)$. Now to compute $\hat{v}_n$ we just replace the unknown parameters by their consistent estimators. The integrals which appeared in the expressions of σ_{14}, σ_{15}, σ_{23}, σ_{25}, σ_{34}, σ_{36}, σ_{45}, σ_{56} can be computed by replacing ν with $\hat{\nu}_n$. Then the computed integrals are as follows:

$$\int_1^\infty \hat{\nu}_n((1,\infty) \times (s,\infty))\, s^{\bar{\gamma}_2 - 1}\, ds$$

$$= \frac{1}{k} \sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right) > 1, \hat{Y}_i\left(\frac{n}{k}\right) > 1\right\}\left(\frac{\left(\hat{Y}_i(n/k)\right)^{\bar{\gamma}_2} - 1}{\bar{\gamma}_2}\right);$$

$$\int_1^\infty \hat{\nu}_n((s,\infty) \times (1,\infty))\, s^{\bar{\gamma}_1 - 1}\, ds$$

$$= \frac{1}{k} \sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right) > 1, \hat{Y}_i\left(\frac{n}{k}\right) > 1\right\}\left(\frac{\left(\hat{X}_i(n/k)\right)^{\bar{\gamma}_1} - 1}{\bar{\gamma}_1}\right);$$

$$\int_1^\infty \hat{\nu}_n((1,\infty) \times (s,\infty))\left(\frac{s^{\bar{\gamma}_2} - 1}{\bar{\gamma}_2}\right) s^{\bar{\gamma}_2 - 1}\, ds$$

$$= \frac{1}{2k} \sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right) > 1, \hat{Y}_i\left(\frac{n}{k}\right) > 1\right\}\left(\frac{\left(\hat{Y}_i(n/k)\right)^{\bar{\gamma}_2} - 1}{\bar{\gamma}_2}\right)^2;$$

$$\int_1^\infty \hat{\nu}_n((s,\infty) \times (1,\infty))\left(\frac{s^{\bar{\gamma}_1} - 1}{\bar{\gamma}_1}\right) s^{\bar{\gamma}_1 - 1}\, ds$$

$$= \frac{1}{2k} \sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right) > 1, \hat{Y}_i\left(\frac{n}{k}\right) > 1\right\}$$

$$\times \left(\frac{\left(\hat{X}_i(n/k)\right)^{\bar{\gamma}_1} - 1}{\bar{\gamma}_1}\right)^2;$$

$$\int_1^\infty \int_1^\infty \hat{\nu}_n((s,\infty) \times (t,\infty))\, s^{\bar{\gamma}_1 - 1} t^{\bar{\gamma}_2 - 1}\, ds\, dt$$

$$= \frac{1}{k} \sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right) > 1, \hat{Y}_i\left(\frac{n}{k}\right) > 1\right\}$$

$$\times \left(\frac{\left(\hat{X}_i(n/k)\right)^{\bar{\gamma}_1} - 1}{\bar{\gamma}_1}\right)\left(\frac{\left(\hat{Y}_i(n/k)\right)^{\bar{\gamma}_2} - 1}{\bar{\gamma}_2}\right);$$

$$\int_1^\infty \int_1^\infty \hat{\nu}_n((s,\infty) \times (t,\infty))\, s^{\bar{\gamma}_1 - 1}\left(\frac{t^{\bar{\gamma}_2} - 1}{\bar{\gamma}_2}\right) t^{\bar{\gamma}_2 - 1}\, ds\, dt$$

$$= \frac{1}{2k} \sum_{i=1}^{n} \mathbf{I}\left\{(X_i, Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right) > 1, \hat{Y}_i\left(\frac{n}{k}\right) > 1\right\}$$

$$\times \left(\frac{\left(\hat{X}_i(n/k)\right)^{\bar{\gamma}_1} - 1}{\bar{\gamma}_1}\right)\left(\frac{\left(\hat{Y}_i(n/k)\right)^{\bar{\gamma}_2} - 1}{\bar{\gamma}_2}\right)^2;$$

$$\int_1^\infty \int_1^\infty \hat{\nu}_n((s,\infty)\times(t,\infty)) \left(\frac{s^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right) s^{\bar{\gamma}_1-1} t^{\bar{\gamma}_2-1}\,ds\,dt$$
$$= \frac{1}{2k}\sum_{i=1}^n \mathbf{I}\left\{(X_i,Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right)>1, \hat{Y}_i\left(\frac{n}{k}\right)>1\right\}$$
$$\times\left(\frac{\left(\hat{X}_i(n/k)\right)^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right)^2\left(\frac{\left(\hat{Y}_i(n/k)\right)^{\bar{\gamma}_2}-1}{\bar{\gamma}_2}\right);$$

$$\int_1^\infty \int_1^\infty \hat{\nu}_n((s,\infty)\times(t,\infty)) \left(\frac{s^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right) s^{\bar{\gamma}_1-1}\left(\frac{t^{\bar{\gamma}_2}-1}{\bar{\gamma}_2}\right) t^{\bar{\gamma}_2-1}\,ds\,dt$$
$$= \frac{1}{4k}\sum_{i=1}^n \mathbf{I}\left\{(X_i,Y_i)\colon \hat{X}_i\left(\frac{n}{k}\right)>1, \hat{Y}_i\left(\frac{n}{k}\right)>1\right\}$$
$$\times\left(\frac{\left(\hat{X}_i(n/k)\right)^{\bar{\gamma}_1}-1}{\bar{\gamma}_1}\right)^2\left(\frac{\left(\hat{Y}_i(n/k)\right)^{\bar{\gamma}_2}-1}{\bar{\gamma}_2}\right)^2.$$

6. Application. We have used our result to estimate the failure probability of the Pettemer zeedijk. We have a set of 828 data of wave height and sea level which can be considered independent and identically distributed. The failure region is given by the set $C = \{(x, y)\colon 0.3x + y \geq 7.6\}$ where x and y are wave height and sea level, respectively (see Figure 1). So in this case we have, $x_n = (7.6)/(0.3)$ and $y_n = 7.6$. We used the upper 27 order statistics (i.e., $k = 27$). This choice is justified in de Haan and de Ronde (1998), Figures 2 and 3. Our estimates of the unknown parameters are

$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_1$	$\hat{b}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
0.5300	0.2915	5.5300	1.6900	-0.0074	-0.1215

The transformed data $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$ and the transformed failure region $[\mathbf{1} + \hat{\gamma}((C_n - \hat{\mathbf{b}})/\hat{\mathbf{a}})]^{1/\hat{\gamma}} = \hat{c}_n \hat{A}$ are shown in Figure 2 (logarithmic scale). The $\hat{c}_n$ is so chosen that the point $(1, 1)$ falls on the boundary of $\hat{A}$. This gives $\hat{c}_n = 2.9772 \times 10^6$. Such a choice of $\hat{c}_n$ allows 26 data points to fall in the shifted failure region. The estimated failure probability is $\hat{p}_n = 1.0547 \times 10^{-8}$. Figure 3 shows the transformed data and the region $\hat{A}$ in logarithmic scale.

To construct a 95% confidence interval for p_n, we used two different estimators of Φ: semiparametric and nonparametric [for definitions, see (1.13) and (1.14), respectively]. The confidence interval in the first case is $(5.75 \times 10^{-9}, 1.934 \times 10^{-8})$, while the second case gives a confidence interval $(3.41 \times 10^{-9}, 3.263 \times 10^{-8})$.

The above-mentioned $\hat{p}_n$ is the failure probability per storm of the considered type. Since the 828 observations cover a period of 13 years, the estimated failure probability per year is 6.72×10^{-7} and its confidence intervals in the two cases are $(3.66 \times 10^{-7}, 1.23 \times 10^{-6})$ (semiparametric) and $(2.17 \times 10^{-7}, 2.07 \times 10^{-6})$ (nonparametric).

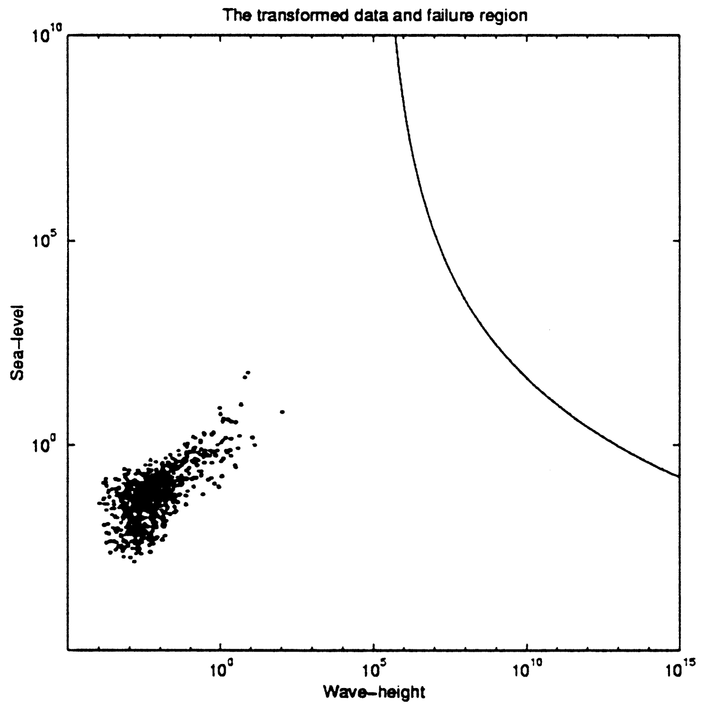

FIG. 2. *Transformed data and failure region on log scale.*

Figure 4 shows how the estimated failure probability and its 95% confidence interval change with different values of the blowup factor c_n (in logarithmic scale) while k is kept fixed at 27. The solid line represents the failure probability, the small circle on this line corresponds to the estimated value $\hat{c}_n = 2.9772 \times 10^6$ as mentioned above, that is, it corresponds to the choice of c_n that follows from the condition: $(1, 1)$ is on the boundary of the failure region. The dash-dotted line represents the 95% semiparametric confidence interval and the dotted line stands for the 95% nonparametric confidence interval.

The results of this paper have been obtained under various restrictions on the sequences $k = k(n)$ and c_n. The optimal choice of k is a function of the marginal distributions and has been discussed in various papers [Hall (1982), Dekkers and de Haan (1993), Drees and Kaufmann (1998), Danielsson, de Haan, Peng and de Vries (1997)]. In the problem at hand, the k was determined in an intuitive way [cf. de Haan and De Ronde (1998), Figures 2 and 3]. The optimal choice of the sequence c_n depends on the convergence rate of the dependence structure and seems to be less critical. This follows from the fact that the uncertainty in the determination of the γ's contributes to the uncertainty in the determination of p_n (via S_1 and S_2 in Theorem 4.1), but the uncertainty in the determination of ν or Φ does not enter. See also Figure 4. The situation regarding the choice of c_n is similar to that of k.

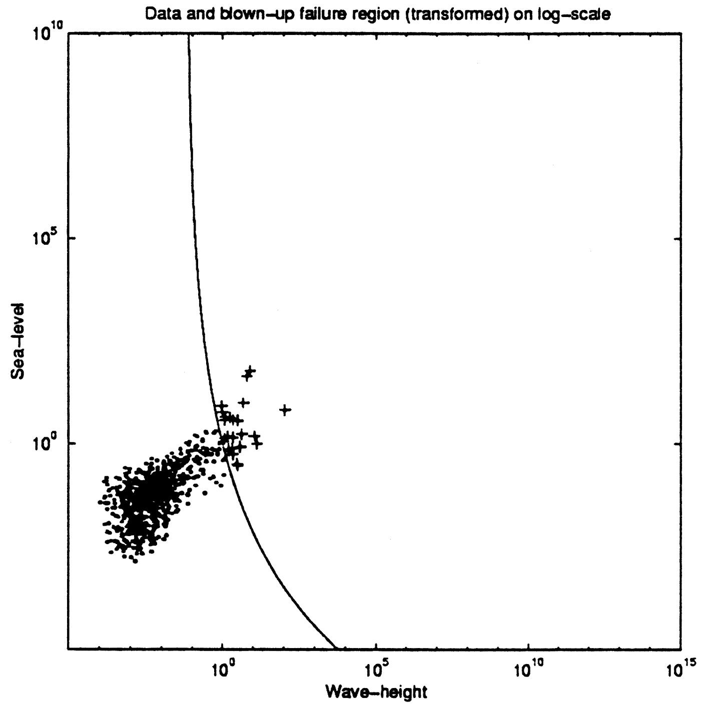

FIG. 3. *Transformed data and blownup failure region on log scale.*

Choosing c_n small could result in higher uncertainty and choosing c_n very big could result in introducing a bias.

6.1. *Simulation study.* We have simulated 50 samples, each of size 1000, from the density function

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

where $1+\gamma_1 x, 1+\gamma_2 y > 0$. We have chosen our $(\gamma_1, \gamma_2) = [-0.0074, -0.1215]$, the same as the estimated value of the extreme value indices from the sea water sample. The exact probability of the failure region $C = \{(x, y)\colon 0.3x + y \geq 7.6\}$ under this distribution is obtained by integrating the density function on the failure region, and that is equal to 1.4224×10^{-4}.

The failure probability for each of the 50 samples is also estimated by our method. For each sample k, the number of order statistics to be used is selected in an intuitive way so that the extreme value indices and henceforth the marginal distributions are estimated as accurately as possible. For each sample, the blowup factor c_n is estimated by solving equation (1.12). On the

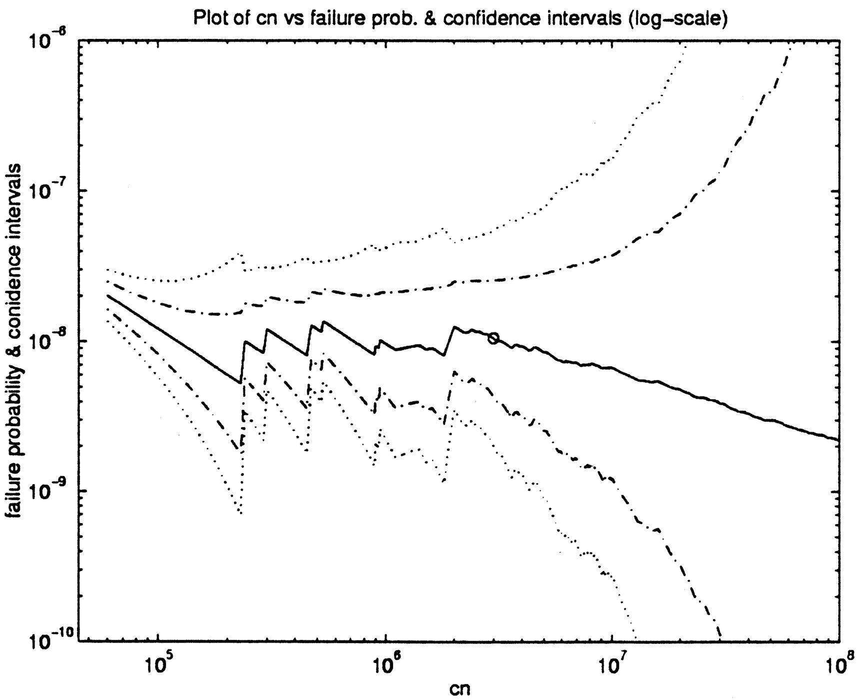

FIG. 4. *Plot of the blowup factor* c_n *against estimated failure probability and confidence intervals.*

basis of the samples drawn, we have that the average estimated failure probability equals 1.6451×10^{-4}.

APPENDIX A

Some properties of q_γ. Here we discuss some important properties of the function $q_\gamma(x) := x^{-\gamma}(\int_1^x u^{\gamma-1} \log u\, du)$. One can easily compute that

$$q_\gamma(x) = \begin{cases} (\gamma \log x - 1 + x^{-\gamma})/\gamma^2, & \text{if } \gamma \neq 0, \\ (\log x)^2/2, & \text{if } \gamma = 0. \end{cases}$$

So it is obvious that, as $x \to \infty$,

$$q_\gamma(x) \sim \begin{cases} (\log x)/\gamma, & \text{if } \gamma > 0, \\ (\log x)^2/2, & \text{if } \gamma = 0, \\ (x^{-\gamma})/\gamma^2, & \text{if } \gamma < 0. \end{cases}$$

Therefore for all $\gamma \in \mathbb{R}$, $\lim_{x \to \infty} q_\gamma(x) = \infty$ and for $a > 0$,

$$\lim_{x \to \infty} \frac{q_\gamma(ax)}{q_\gamma(x)} = a^{-(\gamma \wedge 0)}$$

and

$$\lim_{x \to \infty} \frac{q_{\gamma_1}(x)}{q_{\gamma_2}(x)} = \begin{cases} \gamma_2/\gamma_1, & \text{if } \gamma_1, \gamma_2 > 0, \\ \infty, & \text{otherwise with } \gamma_1 < \gamma_2. \end{cases}$$

Thus we have, as $x \to \infty$,

$$\text{(A.1)} \qquad \frac{q_{\gamma_1}(x)}{q_{\gamma_1 \wedge \gamma_2}(x)} \to \tau_1 := \begin{cases} (\gamma_1 \wedge \gamma_2)/\gamma_1, & \text{if } \gamma_1, \gamma_2 > 0, \\ 1, & \text{otherwise with } \gamma_1 \le \gamma_2, \\ 0, & \text{otherwise with } \gamma_1 > \gamma_2. \end{cases}$$

Similarly we can define τ_2 as the limit of $q_{\gamma_2}(x)/q_{\gamma_1 \wedge \gamma_2}(x)$.

APPENDIX B

Some Vapnik–Cervonenkis classes. To prove the main theorem, we need an assumption that a certain class of sets should satisfy the VC property [see assumption (2.1)]. Here we would like to mention some classes which will satisfy assumptions (2.1). For more details and examples on VC classes, see Dudley (1987) and van der Vaart and Wellner (1996).

First, recall that our set $C_n = \{(s,t)\colon f_n(s,t) \ge 1\} = \{(s,t)\colon f(s/x_n, t/y_n) \ge 1\}$ where f is a certain function and $x_n, y_n > 0$. Now

$$\mathscr{C} := \{C_n\colon n \ge 1\} = \left\{ \left\{ (s,t)\colon f\left(\frac{s}{x_n}, \frac{t}{y_n}\right) \ge 1 \right\} \middle| x_n, y_n > 0, n \ge 1, f \in \mathscr{F} \right\}$$

and

$$\text{(B.1)} \qquad \mathscr{G} := \left\{ l_1 C_n + l_2 \mid l_1, l_2 \in \mathbb{R}^2, C_n \in \mathscr{C} \right\}.$$

However, we need to restrict our $\mathscr{G}$ so that the class $\mathscr{A} = \{[\mathbf{1} + \gamma S]^{\mathbf{1}/\gamma}\colon S \in \mathscr{G}\}$ should not contain any element for which the ν-measure is infinite. Hence we consider the class $\mathscr{G}_* = \{S \in \mathscr{G}\colon [\mathbf{1} + \gamma S]^{1/\gamma} \subset [0,\infty]^2 \setminus [0, \frac{1}{2}]^2\}$, and we have to show $\mathscr{G}_*$ is a VC class. However, if $\mathscr{G}$ is a VC class, then so also is $\mathscr{G}_*$, because it is just a restriction of $\mathscr{G}$.

Now the following classes $\mathscr{C}$ will generate a VC class $\mathscr{G}$ as defined in (B.1):

$\mathscr{C} = \left\{ \frac{x}{a} + \frac{y}{b} \ge 1 \,\middle|\, a, b > 0 \right\}$, the collection of sets bounded below by straight lines;

$\mathscr{C} = \left\{ \frac{x^2}{a} + \frac{y^2}{b} \ge 1 \,\middle|\, a, b > 0 \right\}$, the collection of sets bounded below by ellipses;

$\mathscr{C} = \{cxy \ge 1 \mid c > 0\}$, the collection of sets bounded below by hyperbola;

$\mathscr{C} = \left\{ a_1(b_1 + c_1 x)^{p_1} + a_2(b_2 + c_2 y)^{p_2} \ge 1 \,\middle|\, a_i, c_i > 0, b_i, p_i \in \mathbb{R}, i = 1, 2 \right\}.$

Here we assume that x, y are chosen in such a way that $b_1 + c_1 x$, $b_2 + c_2 y > 0$, so that $f(x,y)$ is defined.

Acknowledgments. Comments by referees and an Associate Editor have helped us to improve several aspects of the paper. The authors gratefully acknowledge the help and technical assistance of Mr. Gerrit Draisma and Dr. Peng Liang.

REFERENCES

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

COLES, S. G. and TAWN, J. A. (1994). Statistical methods for multivariate extremes: an application to structural design. *Appl. Statist.* **43** 1–48.

DANIELSSON, J., DE HAAN, L., PENG, L. AND DE VRIES, C. G. (1997). Using a bootstrap method to choose the sample fraction in the tail index estimation. Technical Report TI97-016/4, Tinbergen Institute, Erasmus Univ. Rotterdam.

DE HAAN, L. (1994). Estimating exceedance probabilities in higher-dimensional space. *Communications in Statistics—Stochastic Models* **10** 765–780.

DE HAAN, L. AND RESNICK, S. (1977). Limit theory for multivariate 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.

DEKKERS, A. and DE HAAN, L. (1993). Optimal choice of sample fraction in extreme-value estimation. *J. Multivariate Anal.* **47** 173–195.

DEKKERS, A., EINMAHL, J. and DE HAAN, L. (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17** 1833–1855.

DIJK, V. and DE HAAN, L. (1992). On the estimation of the exceedance probability of a high level. In *Order Statistics and Nonparametrics: Theory and Applications* (P. K. Sen and I. A. Salama, eds.) 79–92 North-Holland, Amsterdam.

DREES, H. and KAUFMANN, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Process. Appl.* **75** 149–172.

DUDLEY, R. (1987). Universal Donsker classes and metric entropy. *Ann. Probab.* **15** 1306–1326.

EINMAHL, J., DE HAAN, L. and PITERBARG, V. (1998). A non-parametric estimator for the spectral measure of extremes. Technical Report 1998–50, Dept Mathematics Chalmers Univ.

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

GAENSSLER, P. (1983). *Empirical Processes*. IMS, Hayward, CA.

HALL, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc. Ser. B* **44** 37–42.

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

LEDFORD, A. W. and TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83** 169–187.

SINHA, A. K. (1997). Estimating failure probability when failure is rare: multidimensional case. Thesis, Erasmus Univ. Rotterdam. Tinbergen Institute Research Ser. 165.

SMITH, R. L., TAWN, J. A. and YUEN, H. K. (1990). Statistics of multivariate extremes. *Internat. Statist. Rev.* **58** 47–58.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer, New York.

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

TILBURG UNIVERSITY
WARANDELAAN 2
P.O. BOX 90153
5000 LE TILBURG
THE NETHERLANDS
E-MAIL: a.k.sinha@kub.nl