SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853-3801

TECHNICAL REPORT NO. 1114

February 1995

SMOOTHING THE HILL ESTIMATOR[1]

by

Sidney Resnick[1] and
Cătălin Stărică[2]

[1]Parially supported by the National Science Foundation grant DMS-9400535 at Cornell University.
[2]Supported by the National Science Foundation grant DMS-9400535 at Cornell University.

Smoothing the Hill estimator

Sidney Resnick* Cătălin Stărică†
School of ORIE, Cornell University, Ithaca, NY 14853

February 9, 1995

Abstract

For sequences of iid random variables whose common tail 1-F is regularly varying at infinity with an unknown index $-\alpha < 0$, it is well known that Hill's estimator is consistent for α^{-1} and usually asymptotically normally distributed. However, because the Hill estimator is a function of $k = k(n)$, the number of upper order statistics used and which is only subject to the conditions $k \to \infty$, $k/n \to 0$, its use in practice is problematic since there are few reliable guidelines about how to choose k. The purpose of this paper is to make the use of Hill's estimator more reliable through an averaging technique which reduces the asymptotic variance. As a direct result the range in which the smoothed estimator varies as a function of k decreases and the successful use of the estimator is made less dependent on the choice of k. A tail empirical process approach is used to prove the weak convergence of a process closely related to Hill's estimator. The smoothed version of Hill's estimator is a functional of the tail empirical process.

1 Introduction.

Hill's estimator is one of the most popular tools for detecting the presence of heavy tails of the marginal distribution of stationary sequences of random variables. Sets of data displaying characteristics of heavy tails are encountered in diverse fields such as hydrology, finance, reliability or teletraffic engineering. Introduced by Hill (1975), the large sample properties of this estimator have been thoroughly studied. In practice however, the estimator is difficult to use, the main reasons being bias and high sensitivity to the choice of the number of upper order statistics used in estimation. The purpose of this paper is to show that a simple averaging technique and an alternative graphical tool can minimize these difficulties. The averaging technique is theoretically justifiable and can overcome to a certain extent the high volatility of the estimator. Of course, averaging will not correct bias and we hope to address bias in the future.

*Partially supported by NSF Grant DMS-DMS-9400535 at Cornell University.
†Supported by NSF Grant DMS-9400535 at Cornell University.

Let $\{X_n\}$ be a sequence of positive random variables having the same marginal distribution function F, where $\overline{F}$:= 1 - F is $(-\alpha)$ regularly varying at ∞ (written $\overline{F} \in RV_{-\alpha}$); that is, for $\alpha > 0$ we have

$$(1.1) \qquad \lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha}, \quad x > 0.$$

We are interested in estimating α based on observing $X_1, X_2, \ldots, X_n$. Set

$$F^{\leftarrow}(y) = \inf\{x : F(x) \geq y\}, \qquad 0 < y < 1$$

and

$$(1.2) \qquad b(t) := \left(\frac{1}{1-F}\right)^{\leftarrow}(t) = F^{\leftarrow}(1 - \frac{1}{t}), \qquad t > 1,$$

and regular variation implies

$$(1.3) \qquad \overline{F}\,(b(t)) \sim t^{-1}, \quad (t \to \infty).$$

For $1 \leq$ i $\leq$ n, write $X_{(i)}$ for the ith largest value of $X_1, X_2, \ldots, X_n$. Then, with this notation, Hill's estimator based on k upper order statistics is defined as

$$(1.4) \qquad H_{k,n} := \frac{1}{k}\sum_{i=1}^{k} \log\, X_{(i)} - \log\, X_{(k+1)}.$$

When X_n is iid, it is known that Hill's estimator is a consistent estimator of α^{-1}, a fact which is equivalent to regular variation of 1-F in a way made precise by Mason (1982). The consistency requires k=k(n), the number of upper order statistics used in estimation to satisfy $k \to \infty$, $k/n \to 0$. Asymptotic normality of $H_{k,n}$ has also been considered and under additional assumptions $H_{k,n}$ is asymptotically normal with asymptotic variance α^{-2}. Considerable interest has been shown in understanding the asymptotic behaviour of the estimator in depth (see for example Hall (1982), Csörgő and Mason (1985), Häusler and Teugels (1985), Csörgő, Deheuvels and Mason (1985)).

Later developments (Hsing (1991), Rootzen, Leadbetter and de Haan, (1990), Resnick and Stărică (1994)) have shown that Hill's estimator could be successfully applied to a variety of dependent sequences of random variables. Together with accompanying tools, Hill's estimator can be of great help in studying sets of real data (cf. Feigin, Resnick and Stărică (1994)).

When using Hill's estimator in practice, what is frequently done is to produce the so called Hill plot by graphing $\{(k, H_{k,n}^{-1}), 1 \leq k \leq n\}$ and hoping that one can infer a value of α from a stable regime of this plot. Unfortunately, the practical use of the estimator is hampered by the already mentioned high volatility of the plot and bias problems. It is often the case that volatility of the Hill plot prevents a clear choice of α. The estimate is often highly sensitive to the choice of k, the number of upper order statistics used to calculate $H_{k,n}$. A relevant example is shown in the left hand picture of Figure 1 which displays the Hill plot for a sequence of 700 independent random samples from a Pareto distribution, with no region where the graph is sufficiently stable that reliable estimates can be achieved. The range of the plot is rather large and the practitioner is left to the mercy of inspiration when estimating α from the graph.

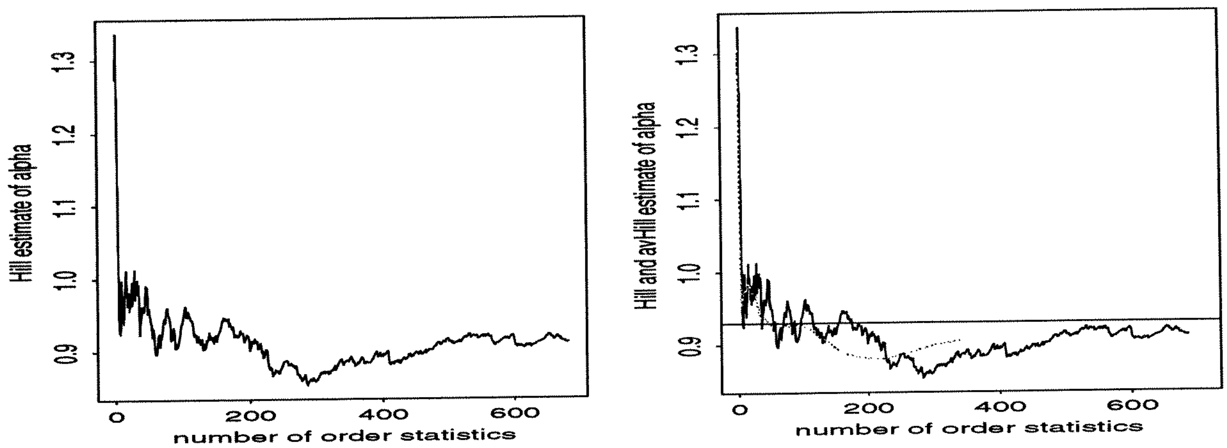

Figure 1

The simple averaging technique proposed in this paper reduces the volatility of the plot. The smoothing procedure consists in averaging the Hill estimator values corresponding to different numbers of order statistics:

$$avH_{k,n} := \frac{1}{(u-1)k} \sum_{p=k+1}^{uk} H_{p,n} \tag{1.5}$$

where $u > 1$. An important conclusion of our paper is that through averaging the variance of the Hill estimator can be considerably reduced and the volatility of the plot thereby tamed. The estimate of α suggested by avHill can be used as the basis for further studies which would, for example, use bootstrap techniques to correct for bias. The averaging technique, though simple and obvious, seems quite useful and it is illustrated in the right hand picture of Figure 1 which shows the result of the smoothing (the dotted graph) in comparison with the original Hill estimate (the solid graph). The real value of α which is 0.93 is also plotted. Due to the fact that u in (1.5) was 2, the averaging stopped when k reached 350. This explains why the smoothed graph (the dotted line) does not go all the way to the right of the picture.

As an alternative to the Hill plot, we found it useful to display the information provided by the Hill estimation as

$$\{(\theta, H^{-1}_{\lceil n^\theta \rceil, n}), 0 \leq \theta \leq 1, \}$$

where we write $\lceil y \rceil$ for the smallest integer greater or equal to $y \geq 0$. We call such a plot the *alternative Hill plot*. As in the case of the traditional Hill plot one tries to read the estimated value from a stable portion of the graph. The reason the alternative display is helpful is that the significant part of the graph, i.e. the part corresponding to a relatively small number of order statistics gets to be shown more clearly, covering a bigger portion of the displayed space. On the other hand, the part of the graph corresponding to a high number of order statistics, which covers a disproportionately large part in the traditional Hill plot gets rescaled. Thus, the interpretation of the graph is easier and more accurate. More detailed discussion of this procedure follows in Section 5.

The improvement that the smoothed estimator brings can be better assessed if one uses the alternative Hill plot which is displayed in Figure 2. The solid graph is the alternative Hill plot $\{(\theta, H^{-1}_{\lceil n^\theta \rceil,n}), 0 \leq \theta \leq 1\}$, the dotted graph is the alternative avHill plot $\{(\theta, avH^{-1}_{\lceil n^\theta \rceil,n}), 0 \leq \theta \leq 1\}$. Due to the fact that u in (1.5) was 2, the averaging stopped when k reached $700^{0.87}$. This explains why the smoothed graph (the dotted one) does not go all the way to the right of the picture.

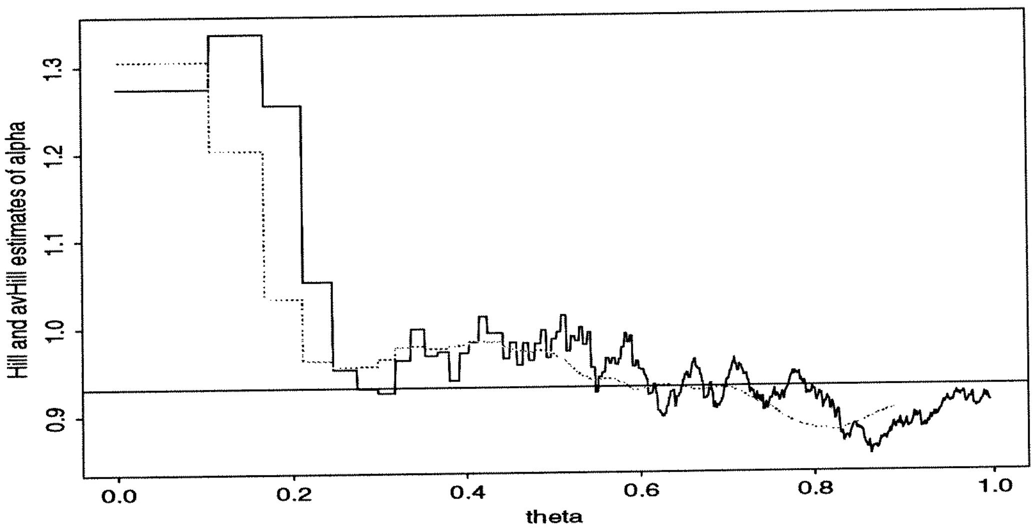

Figure 2

Looking at Figure 2 one sees that the practitioner is now left with less doubts. The smoothed graph, drawn by the dotted curve, has a narrower range over its stable regime, with less sensitivity to the value of k.

In practice one quickly notices that Hill's estimator can be subject to considerable bias even for big sample sizes. An example where the sample size is big (1,000,000 observations) and where the bias displayed by the estimator is considerable follows. Consider the distribution of a random variable X obtained by applying the function $U(x) = x \log x$ to a random variable with a Pareto distribution with parameter 5; that is

$$X = U(Y) = Y \log Y, \; P(Y > x) = x^{-5}, \; x > 1. \tag{1.6}$$

A sample of 1,000,000 observations is drawn from this distribution.

Note that $U^{\leftarrow}(x) \sim x / \log x$ for $x \to \infty$. Thus

$$P(X > x) = P(U(Y) > x) = P(Y > U^{\leftarrow}(x)) = (U^{\leftarrow}(x))^{-5} \sim x^{-5} \log^5 x$$

when $x \to \infty$. Therefore the tail of X belongs to RV_{-5} and the Hill plot should give an estimate of 5. Figure 3 shows such an alternative Hill and alternative avHill plot where the range of the graph stays far from value 5. The smoothed estimator points toward an estimate of 3.3. Obviously when considerable bias is present, averaging technique offers no improvement.

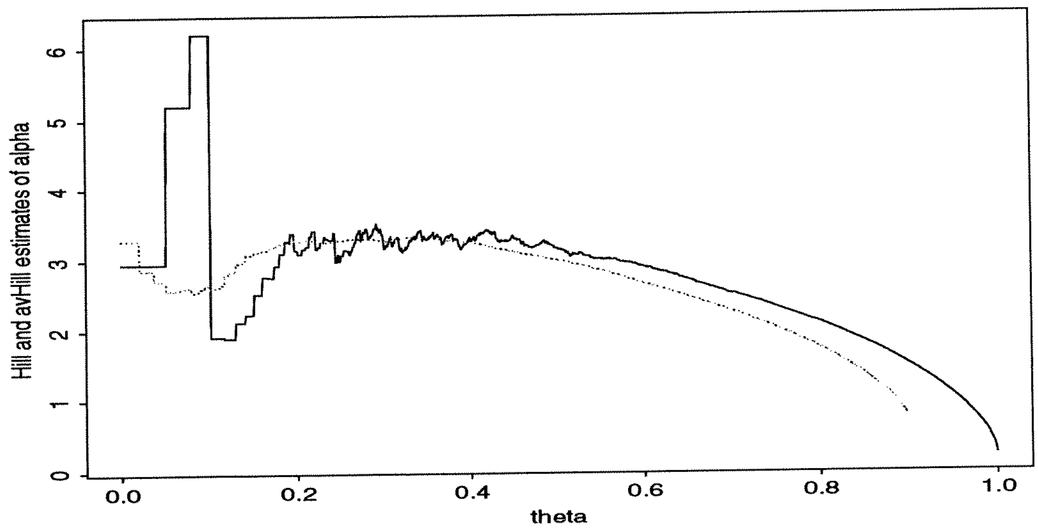

Figure 3

We now introduce some notation: For $x \in R$ and $A \subset R$ define

$$\epsilon_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \in A^c. \end{cases}$$

On $[0, \infty)$ define the tail empirical process and its normalization as

$$E_{k,n}(y) := \frac{1}{k} \sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[y^{-1/\alpha}, \infty] \tag{1.7}$$

$$\mathcal{E}_{k,n}(y) := \sqrt{k}(E_{k,n}(y) - y) = \sqrt{k}\left(\frac{1}{k} \sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[y^{-1/\alpha}, \infty] - y\right) \tag{1.8}$$

and the Hill process and its normalization on $(0, \infty)$, by

$$H_{k,n}(y) := H_{\lceil ky \rceil, n} \tag{1.9}$$

$$\mathcal{H}_{k,n}(y) := \sqrt{k}\left(H_{k,n}(y) - \frac{1}{\alpha}\right) = \sqrt{k}\left(H_{\lceil ky \rceil, n} - \frac{1}{\alpha}\right). \tag{1.10}$$

The Hill process defined here is a modification of the Hill process of Mason and Turova (1994). The reason for the modification was that we wished to focus on statistical estimation of the index of regular variation rather then the study of the general asymptotic behaviour of the Hill process. The centering and scaling sequences in Hill's process of Mason and Turova were functionals of the quantile function of F. Therefore, though theoretically interesting their process was statistically unsuitable as it depended on the unknown quantile function. To render the process more suitable to our purpose we chose to center and scale the process differently. With some effort it is possible to derive our result from the convergence proved by Mason and Turova. However a proof using the tail empirical process provides insight into the phenomena of heavy tails and consolidates the evidence for the importance of the tail empirical process in studying the tail behoavior of heavy tailed distributions. This technique was also used by Resnick and Stărică (1995) to prove the consistency of the Hill estimator for infinite moving averages.

We emphasize the following approach to the problem: Associate the tail empirical process to the sequence $X_1, X_2, \ldots, X_n$, show the weak convergence of the normalized tail empirical process to a Brownian motion and deduce from this the convergence of the normalized Hill process. By a simple process of integration the asymptotic behaviour of the smoothed estimator is obtained.

The paper is organized as follows: Section 2 covers the asymptotic behaviour of the normalized tail empirical process

$$y \to \sqrt{k}(E_{k,n}(y) - y) = \sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[y^{-1/\alpha}, \infty] - y\right)$$

in $D[0, \infty)$. Section 3 studies the behaviour of the normalized Hill process

$$y \to \sqrt{k}\left(H_{k,n}(y) - \frac{1}{\alpha}\right) = \sqrt{k}\left(H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right)$$

in $D(0, \infty)$. The theoretical ground for the smoothing of the Hill estimator is provided in Section 4. Comments on the practical use of the smoothed estimator, more examples with simulated data and one where our technique proves helpful in analyzing real data are contained in Section 5.

We will assume throughout the paper that $X_n, n \geq 1$ is iid. For proving asymptotic normality of the Hill estimator, a second order regular variation condition is needed.

Condition 1 : $\overline{F}$:= 1 - F is second order $(-\alpha, \gamma)$ regularly varying at ∞ (written $\overline{F} \in 2RV(-\alpha, \gamma)$); that is, there exists an $\alpha > 0$, $\gamma \leq 0$, K$\in$ $\mathbf{R}$ such that $\overline{F} = x^{-\alpha}L(x)$ and

$$\lim_{t\to\infty} \frac{\frac{L(tx)}{L(t)} - 1}{g(t)} = K\frac{x^\gamma - 1}{\gamma} \tag{1.11}$$

where $g \in RV_\gamma$, for all $x > 0$. (See, for example, de Haan and Stadtmüller (1994), Geluk and de Haan (1987)). Furthermore, by Goldie and Smith (1987), Theorem 2.1.1, Proposition 2.5.2

(i) the convergence in (1.11) is uniform on $[1, \infty)$,

(ii) $\int_1^\infty \left(\frac{L(xu)}{L(x)} - 1\right) u^{-\alpha-1} du \sim \frac{K}{\alpha(\alpha-\gamma)} g(x)$ as $x \to \infty$.

The function g appearing in (1.11) is sometimes used to further restrict the sequence $k = k(n)$.

Condition 2 : The sequence k(n) satisfies

$$\sqrt{k} g(b(\frac{n}{k})) \to 0 \tag{1.12}$$

for $n \to \infty$ where $b(n/k)$ is defined in (1.2).

Condition 2 has been used by many authors. See for example Hall (1982), Häusler and Teugels (1985), Dekkers and de Haan (1991).

2 Asymptotic behaviour of the tail empirical process

The derivation of the asymptotic behaviour of the tail empirical process is the first step in our endeavour of smoothing the Hill estimator.

Proposition 2.1 *Assume that* **Condition 1** *and* **Condition 2** *hold. Then, as* $n \to \infty$, $k \to \infty$ *and* $k/n \to 0$,

$$\mathcal{E}_{k,n}(y) = \sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[y^{-1/\alpha}, \infty] - y\right) \Rightarrow W(y), \tag{2.1}$$

in $D[0, \infty)$, *where* $\{W(t), t \geq 0\}$ *is a standard Brownian motion.*

Proof: We first prove an invariance principle by centering to zero expectations and show

$$X_n(y) := \sqrt{k}(E_{k,n}(y) - E(E_{k,n}(y))) \Rightarrow W(y) \tag{2.2}$$

in $D[0, \infty)$. Proving convergence of the finite dimensional distributions is straightforward (see de Haan and Resnick (1994)). From Theorem 15.6 of Billingsley (1968), in order to show (2.2) in $D[0, \infty)$, it suffices to show for any $0 \leq y_1 < y < y_2$ that

$$\limsup_{n\to\infty} E\left(|\mathcal{E}_n(y) - \mathcal{E}_n(y_1)|^2 |\mathcal{E}_n(y_2) - \mathcal{E}_n(y)|^2\right) < (y_2 - y_1)^2. \tag{2.3}$$

Set

$$\begin{aligned} \alpha_i &= \epsilon_{X_i/b(n/k)}[y^{-1/\alpha}, y_1^{-1/\alpha}) - P(X_1/b(n/k) \in [y^{-1/\alpha}, y_1^{-1/\alpha})) = \epsilon_{X_i/b(n/k)}[y^{-1/\alpha}, y_1^{-1/\alpha}) - p, \\ \beta_i &= \epsilon_{X_i/b(n/k)}[y_2^{-1/\alpha}, y^{-1/\alpha}) - P(X_1/b(n/k) \in [y_2^{-1/\alpha}, y^{-1/\alpha})) = \epsilon_{X_i/b(n/k)}[y_2^{-1/\alpha}, y^{-1/\alpha}) - q. \end{aligned}$$

Note as $n \to \infty$

$$\frac{n}{k}p \to y - y_1, \quad \frac{n}{k}q \to y_2 - y.$$

Then the expectation in (2.3) is the same as

$$\frac{1}{k^2}E(\sum_{i=1}^n \alpha_i)^2(\sum_{i=1}^n \beta_i)^2 = \frac{n}{k^2}E(\alpha_1^2\beta_1^2) + \frac{n(n-1)}{k^2}E(\alpha_1^2)E(\beta_1^2) + \frac{2n(n-1)}{k^2}(E(\alpha_1\beta_1))^2$$
$$= \frac{n}{k^2}(p^2q + pq^2 + p^2q^2) + \frac{n(n-1)}{k^2}pq(1-p)(1-q) + \frac{2n(n-1)}{k^2}p^2q^2$$
$$= \frac{n(n-1)}{k^2}pq + o(\frac{k}{n}) \sim (y-y_1)(y_2-y) < (y_2-y_1)^2$$

as $n \to \infty$ as required for tightness.

Now it remains to remove the centering by expectations and replace by centering by the identity function. This is accomplished by means of the second order condition since (1.11) and (1.12) imply:

$$\sqrt{k}\left(\frac{n}{k}P(\frac{X_1}{b(n/k)} \geq t^{-1/\alpha}) - t\right) \to 0$$

when $n \to \infty$. □

3 Asymptotic behaviour of the Hill process

The derivation of the asymptotic behaviour of the Hill process $H_{k,n}(\cdot)$ has as starting point Proposition 2.1. The proof is based on the observation that Hill process can almost be expressed as a functional of the tail empirical process and its inverse process. The main ingredients are continuity arguments and a convergence together technique.

Proposition 3.1 *Under* **Condition 1** *and* **Condition 2**, *as $n \to \infty$, $k \to \infty$ and $k/n \to 0$,*

$$\mathcal{H}_{k,n}(y) = \sqrt{k}\left(H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right) \Rightarrow \frac{1}{\alpha y}W(y). \tag{3.1}$$

in $D(0,\infty)$, where W is a standard Brownian motion.

Proof: We first sketch the outline of the proof. The Hill process is a functional of a process akin to the tail empirical process:

$$H_{k,n}(y) = -\frac{1}{\lceil ky\rceil}\int_0^1 \sum_{i=1}^n \epsilon_{X_i/X_{(\lceil ky\rceil)}}[x^{-1/\alpha},\infty]\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}. \tag{3.2}$$

The weak convergence proved in Proposition 2.1 will allow us to take **STEP 1**

$$\sqrt{k}\left(\frac{1}{ky}\int_{1/a^\alpha}^1 \sum_{i=1}^n \epsilon_{X_i/X_{(\lceil ky\rceil)}}[x^{-1/\alpha},\infty]\frac{dx^{-1/\alpha}}{x^{-1/\alpha}} - \frac{1-a^{-\alpha}}{\alpha}\right)$$
$$\Rightarrow \frac{1}{y}\left(\int_{1/a^\alpha}^1 W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}} - \frac{1-a^{-\alpha}}{\alpha}W(y)\right) \tag{3.3}$$

in $D(0,\infty)$, for any $a > 1$. **Step 2** will be to prove that the same convergence holds even when $a = \infty$. The natural way to attempt this is a converging together argument. Together with (3.2) this will imply

$$\sqrt{k}\left(\frac{\lceil ky\rceil}{ky}H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right) \Rightarrow \frac{1}{\alpha y}W(y) \tag{3.4}$$

in $D(0,\infty)$. Finally **Step 3** will cover the distance between this last convergence and the conclusion of the theorem. Now that we have a broad image of the path to follow let us detail the steps:

STEP 1

Let us first note that the inverse of the tail empirical process is given by

$$E_{k,n}^{\leftarrow}(y) = \left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}$$

and it is an element in $D[0,\infty)$. By Vervaat's lemma (Vervaat (1972)) and Proposition 2.1 it follows that

$$\sqrt{k}\left(\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha} - y\right) \Rightarrow -W(y) \tag{3.5}$$

in $D[0,\infty)$. Define $D_1[0,\infty)$ to be the set of all positive, increasing functions in $D[0,\infty)$ with the topology which it inherits from $D[0,\infty)$. By a continuity argument the following joint convergence holds:

$$\left(\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[x^{-1/\alpha},\infty] - x\right), \left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}, \sqrt{k}\left(\left(\frac{X_{(\lceil kz\rceil)}}{b(n/k)}\right)^{-\alpha} - z\right)\right) \tag{3.6}$$
$$\Rightarrow \quad (W(x), y, -W(z))$$

in $D[0,\infty)\times D_1[0,\infty)\times D[0,\infty)$. For $a > 0$, define the operator

$$T : D[0,\infty)\times D_1[0,\infty)\times D[0,\infty) \mapsto D(0,\infty)$$

by

$$T(f,g,h)(y) = \frac{1}{y}\int_{1/a^\alpha}^{1} f(xg(y))\frac{dx^{-1/\alpha}}{x^{-1/\alpha}} + \frac{1-a^{-\alpha}}{\alpha y}h(y).$$

Note that this operator is continuous at continuous functions, a fact which together with (3.6) implies

$$\sqrt{k}\frac{1}{y}\int_{1/a^\alpha}^{1}\left[\frac{1}{k}\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[\left(x(\frac{X_{(\lceil ky\rceil)}}{b(n/k)})^{-\alpha}\right)^{-1/\alpha},\infty] - x\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}\right]\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}$$
$$+\sqrt{k}\frac{1-a^{-\alpha}}{\alpha}\left(\frac{1}{y}\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha} - 1\right)$$
$$\Rightarrow \frac{1}{y}\int_{1/a^\alpha}^{1} W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}} - \frac{1-a^{-\alpha}}{\alpha y}W(y),$$

in $D(0,\infty)$ which is equivalent to:

$$\sqrt{k}\left(\frac{1}{ky}\int_{1/a^\alpha}^{1}\left(\sum_{i=1}^{n}\epsilon_{X_i/X_{(\lceil ky\rceil)}}[x^{-1/\alpha},\infty]-x\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}\right)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}\right)$$
$$+\sqrt{k}\frac{1-a^{-\alpha}}{\alpha}\left(\frac{1}{y}\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}-1\right)$$
$$\Rightarrow \frac{1}{y}\int_{1/a^\alpha}^{1}W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}-\frac{1-a^{-\alpha}}{\alpha y}W(y)$$

in $D(0,\infty)$. This completes **Step 1**.

Step 2

We now prove (3.4) using Theorem 4.2 of Billingsley (1968). First notice that as $a\to\infty$

$$\int_{1/a^\alpha}^{1}W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}-\frac{1-a^{-\alpha}}{\alpha}W(y)\Rightarrow\int_0^1 W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}-\frac{1}{\alpha}W(y),$$

in $D(0,\infty)$, if, for example

$$P\left(\sup_{y\in[1,2]}\left|\int_0^{1/a^\alpha}W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}+\frac{a^{-\alpha}}{\alpha}W(y)\right|>\delta\right)\to 0 \tag{3.7}$$

for $a\to\infty$. But

$$\sup_{y\in[1,2]}\left|\int_0^{1/a^\alpha}W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}+\frac{a^{-\alpha}}{\alpha}W(y)\right|$$
$$\le\sup_{y\in[1,2]}\int_0^{1/a^\alpha}\left|\frac{W(xy)}{xy}\right|xy\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}+\frac{a^{-\alpha}}{\alpha}\bigvee_{s=1}^{2}|W(s)|$$
$$\le\frac{2a^{-\alpha}}{\alpha}\bigvee_{s=0}^{2/a^\alpha}\left|\frac{W(s)}{s}\right|+\frac{a^{-\alpha}}{\alpha}\bigvee_{s=1}^{2}|W(s)|.$$

Now (3.7) follows easily from the last inequality. Also one easily can check that

$$\int_0^1 W(xy)\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}-\frac{1}{\alpha}W(y)\stackrel{d}{=}W(y).$$

To finish the converging together argument we should verify that

$$\lim_{a\to\infty}\limsup_{n}P\left(\frac{1}{\sqrt{k}}\sup_{y\in[1,2]}\int_0^{1/a^\alpha}\left|\sum_{i=1}^{n}\epsilon_{X_i/X_{(\lceil ky\rceil)}}[x^{-1/\alpha},\infty]-kx\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}\right|\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}>\frac{\delta}{2}\right)$$
$$+P\left(\sqrt{k}\sup_{y\in[1,2]}\left|\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}-y\right|>a^\alpha\alpha\frac{\delta}{2}\right)=0.$$

The fact that the second term goes to zero follows from

$$\sqrt{k}\sup_{y\in[1,2]}\left|\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}-y\right|\Rightarrow\sup_{y\in[1,2]}|W(y)|,$$

which is a consequence of (3.5). We concentrate now on the remaining term. Let $\varepsilon>0$. First, perform a change of variables $u=x(X_{(\lceil ky\rceil)}/b(n/k))^{-\alpha}$ and notice that

$$\begin{aligned}
&P\left(\frac{1}{\sqrt{k}}\sup_{y\in[1,2]}\int_0^{1/a^{\alpha}}\left|\sum_{i=1}^{n}\epsilon_{X_i/X_{(\lceil ky\rceil)}}[x^{-1/\alpha},\infty]-kx\left(\frac{X_{(\lceil ky\rceil)}}{b(n/k)}\right)^{-\alpha}\right|\frac{dx^{-1/\alpha}}{x^{-1/\alpha}}>\frac{\delta}{2}\right)\\
&\le P\left(\frac{1}{\sqrt{k}}\sup_{y\in[1,2]}\int_0^{(aX_{(\lceil ky\rceil)}/b(n/k))^{-\alpha}}\left|\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}>\frac{\delta}{2}\right)\\
&\le P\left(\frac{1}{\sqrt{k}}\int_0^{(aX_{(k)}/b(n/k))^{-\alpha}}\left|\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}>\frac{\delta}{2}\right)\\
&\le P\left(\frac{1}{\sqrt{k}}\int_0^{(1+\varepsilon)a^{-\alpha}}\left|\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}>\frac{\delta}{2}\right)\\
&+P\left(\left|\left(\frac{X_{(k)}}{b(n/k)}\right)^{-\alpha}-1\right|>\varepsilon\right),
\end{aligned}$$

due to $(X_{(\lceil ky\rceil)}/b(n/k))^{-\alpha}\le(X_{(k)}/b(n/k))^{-\alpha}$ for $y\in[1,2]$ and $(X_{(k)}/b(n/k))^{-\alpha}\xrightarrow{P}1$ (by (3.5)). We therefore concentrate on

$$\begin{aligned}
&P\left(\frac{1}{\sqrt{k}}\int_0^{(1+\varepsilon)a^{-\alpha}}\left|\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}>\frac{\delta}{2}\right)\\
&\le P\Big(\frac{1}{\sqrt{k}}\int_0^{(1+\varepsilon)a^{-\alpha}}\left|\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]-E\Big(\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]\Big)\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}\\
&>\frac{\delta}{2}-\frac{1}{\sqrt{k}}\int_0^{(1+\varepsilon)a^{-\alpha}}\left|E\Big(\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]\Big)-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}\Big).
\end{aligned}\tag{3.8}$$

We now claim that

$$\begin{aligned}
&\frac{1}{\sqrt{k}}\int_0^{(1+\varepsilon)a^{-\alpha}}\left|E\Big(\sum_{i=1}^{n}\epsilon_{X_i/b(n/k)}[u^{-1/\alpha},\infty]\Big)-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}\\
&=\frac{1}{\sqrt{k}}\int_0^{(1+\varepsilon)a^{-\alpha}}\left|nP\left(\frac{X_i}{b(n/k)}\ge u^{-1/\alpha}\right)-ku\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}\\
&=\int_0^{(1+\varepsilon)a^{-\alpha}}\sqrt{k}\left|\frac{n}{k}P\left(\frac{X_i}{b(n/k)}\ge u^{-1/\alpha}\right)-u\right|\frac{du^{-1/\alpha}}{u^{-1/\alpha}}
\end{aligned}\tag{3.9}$$

tends to 0 for $n \to \infty$. Once proved, this claim will allow us to substitute δ' for

$$\frac{\delta}{2} - \frac{1}{\sqrt{k}} \int_0^{(1+\varepsilon)a^{-\alpha}} \left| E(\sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty]) - ku \right| \frac{du^{-1/\alpha}}{u^{-1/\alpha}}$$

in (3.8). Notice that due to the uniform convergence in **Condition 1.(i)** the expression

$$\frac{1}{k} E(\sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty]) - u = \frac{n}{k} P\left(\frac{X_i}{b(n/k)} \geq u^{-1/\alpha}\right) - u$$

keeps constant sign for large n and for any $u \geq 1$. After performing the change of variable $y = u^{-1/\alpha}$ the integral becomes

$$\int_0^{(1+\varepsilon)a^{-\alpha}} \sqrt{k} \left| \frac{n}{k} P\left(\frac{X_i}{b(n/k)} \geq u^{-1/\alpha}\right) - u \right| \frac{du^{-1/\alpha}}{u^{-1/\alpha}}$$

$$\left| \int_0^{(1+\varepsilon)a^{-\alpha}} \sqrt{k} \left(\frac{n}{k} P(\frac{X_i}{b(n/k)} \geq u^{-1/\alpha}) - u\right) \frac{du^{-1/\alpha}}{u^{-1/\alpha}} \right|$$

$$\left| \int_{a(1-\varepsilon)^{-1/\alpha}}^{\infty} \sqrt{k} \left(\frac{\overline{F}(b(n/k)u)}{\overline{F}(b(n/k))} - u^{-\alpha}\right) \frac{du}{u} \right|$$

$$= \left| \int_{a(1-\varepsilon)^{-1/\alpha}}^{\infty} \sqrt{k} \left(\frac{L(b(n/k)u)}{L(b(n/k))} - 1\right) u^{-\alpha-1} du \right| \sim |M| \sqrt{k} g(b(n/k)).$$

Condition 1(ii) together with (1.12) assure us the previous claim is correct. We are in the position now to undertake the last step of our proof, i.e.

$$\lim_{a} \limsup_{n} P\left(\frac{1}{\sqrt{k}} \int_0^{(1+\varepsilon)a^{-\alpha}} \left| \sum_{i=1}^{n} (\epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty] - E\epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty]) \right| \frac{du^{-1/\alpha}}{u^{-1/\alpha}} > \delta'\right)$$

equals 0. By making use of the Chebyshev inequality we get

$$P(\frac{1}{\sqrt{k}} \int_0^{(1+\varepsilon)a^{-\alpha}} \left| \sum_{i=1}^{n} \left(\epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty] - E\epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty]\right) \right| \frac{du^{-1/\alpha}}{u^{-1/\alpha}} > \delta')$$

$$\leq \frac{1}{\delta' k} E\left(\int_0^{(1+\varepsilon)a^{-\alpha}} \left| \sum_{i=1}^{n} \left(\epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty] - E(\epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty])\right) \right| \frac{du^{-1/\alpha}}{u^{-1/\alpha}}\right)^2$$

$$\leq \frac{1}{\delta' k} \int_0^{(1+\varepsilon)a^{-\alpha}} Var\left(\sum_{i=1}^{n} \epsilon_{X_i/b(n/k)}[u^{-1/\alpha}, \infty]\right) \frac{du^{-1/\alpha}}{u^{-2/\alpha}}$$

$$= \frac{n}{\delta' k} \int_{a(1+\varepsilon)^{-1/\alpha}}^{\infty} Var(\epsilon_{X_1/b(n/k)}[u^{-1/\alpha}, \infty]) \frac{du}{u^2} \leq \frac{1}{\delta'} \int_{a(1-\varepsilon)^{-1/\alpha}}^{\infty} \frac{n}{k} P\left(\frac{X_1}{b(n/k)} > u\right) \frac{du}{u^2}.$$

By Potter's inequality (Bingham, Goldie, Teugels (1987), Theorem 1.5.6), for a given δ, there exists an n_0 such that, for $u \geq 1$

$$(1-\delta) u^{-\alpha-\delta} < \frac{n}{k} \overline{F}\left(b\left(\frac{n}{k}\right) u\right) < (1+\delta) u^{-\alpha+\delta}$$

for any $n > n_0$. Therefore it follows that:

$$\limsup_n \int_{a(1-\varepsilon)^{-1/\alpha}}^{\infty} \frac{n}{k}\overline{F}\left(b\left(\frac{n}{k}\right)u\right)\frac{du}{u^2} < \int_{a(1-\varepsilon)^{-1/\alpha}}^{\infty} (1+\delta)u^{-\alpha+\delta-2}du.$$

Making sure that $\delta < 1$ and letting $a \to \infty$ completes the proof of **Step 2**. Thus we have proved (3.4). We now discuss the last step of the proof.

Step 3

We note that

$$\sqrt{k}\left(\frac{\lceil ky\rceil}{ky}H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right) = \sqrt{k}\frac{\lceil ky\rceil}{ky}\left(H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\frac{ky}{\lceil ky\rceil}\right)$$
$$= \frac{\lceil ky\rceil}{ky}\left(\sqrt{k}\left(H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right) + \sqrt{k}\left(\frac{1}{\alpha} - \frac{1}{\alpha}\frac{ky}{\lceil ky\rceil}\right)\right).$$

Due to the fact that

$$\sqrt{k}\left(\frac{\lceil ky\rceil}{ky} - 1\right) \to 0$$

as $k \to \infty$ the conclusion of the theorem follows. □

4 Asymptotics of the smoothed Hill estimator

As we have already mentioned, the smoothing procedure consists in averaging the Hill estimator values $H_{k,n}$ over a broad range of k, the number of order statistics. The order of the number of terms involved in the averaging is k. Therefore, when $n, k \to \infty$ we will be averaging larger and larger numbers of Hill estimator values with a consequent reduction in asymptotic variance. Let us state the result precisely:

Proposition 4.1 *Let $0 < s < t$ be fixed. Then for $n \to \infty$*

$$\sqrt{k}\left(\frac{1}{k(t-s)}\sum_{p=\lceil ks\rceil}^{\lceil kt\rceil} H_{p,n} - \frac{1}{\alpha}\right) \Rightarrow N(0,c)$$

where

$$c = \frac{1}{\alpha^2}\frac{2}{t-s}\left(1 - \frac{s\ln(t/s)}{t-s}\right).$$

Proof: By the continuous mapping theorem integrate both sides of (3.1) to get:

$$\sqrt{k}\int_s^t \left(H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right)dy \Rightarrow \frac{1}{\alpha}\int_s^t \frac{1}{y}W(y)dy. \tag{4.1}$$

After some elementary computations we get the left hand side to look like

$$\sqrt{k}\int_s^t \left(H_{\lceil ky\rceil,n} - \frac{1}{\alpha}\right) dy$$

$$= \sqrt{k}\left(\frac{\lceil ks\rceil - ks}{k} H_{\lceil ks\rceil,n} + \frac{1}{k}\sum_{p=\lceil ks\rceil+1}^{\lfloor kt\rfloor} H_{p,n} + \frac{kt - \lfloor kt\rfloor}{k} H_{\lceil kt\rceil,n} - (t-s)\frac{1}{\alpha}\right)$$

$$(4.2) \quad = \sqrt{k}\left(\frac{1}{k}\sum_{p=\lceil ks\rceil}^{\lceil kt\rceil} H_{p,n} - (t-s)\frac{1}{\alpha}\right) + \frac{\lceil ks\rceil - ks - 1}{\sqrt{k}} H_{\lceil ks\rceil,n} + \frac{tk - \lfloor kt\rfloor - 1}{\sqrt{k}} H_{\lceil kt\rceil,n}$$

$$= (t-s)\sqrt{k}\left(\frac{1}{k(t-s)}\sum_{p=\lceil ks\rceil}^{\lceil kt\rceil} H_{p,n} - \frac{1}{\alpha}\right) + O_P\left(\frac{1}{\sqrt{k}}\right)(H_{\lceil ks\rceil,n} + H_{\lceil kt\rceil},n).$$

An easy computation shows that

$$Var \int_s^t \frac{1}{y} W(y)dy = 2[(t-s) - s\ln(t/s)].$$

This computation together with (4.2) ends the proof. □

It is useful to write the previous theorem in terms of a different pair of variables $(s, u := t/s)$. Let $1 \le s, 2 \le u$. Then for $n \to \infty$

$$\sqrt{k}\left(\frac{1}{ks(u-1)}\sum_{p=\lceil ks\rceil}^{\lceil ksu\rceil} H_{p,n} - \frac{1}{\alpha}\right) \Rightarrow N(0,c)$$

where

$$c = \frac{1}{\alpha^2}\frac{2}{s(u-1)}\left(1 - \frac{\ln u}{u-1}\right).$$

How does the choice of s and u influence the size of the variance? Note that

$$u \to \frac{2}{u-1}\left(1 - \frac{\ln u}{u-1}\right)$$

and

$$s \to \frac{1}{s}$$

are decreasing in u, respectively s. Therefore the bigger u and s the better. The next figure shows the dependence of the variance on (s,u). The vertical axis represents the quotient between the asymptotic variance of the smoothed (s,u) estimator and the asymptotic variance of Hill's estimator. The decrease is rather sharp even for small values of s and u.

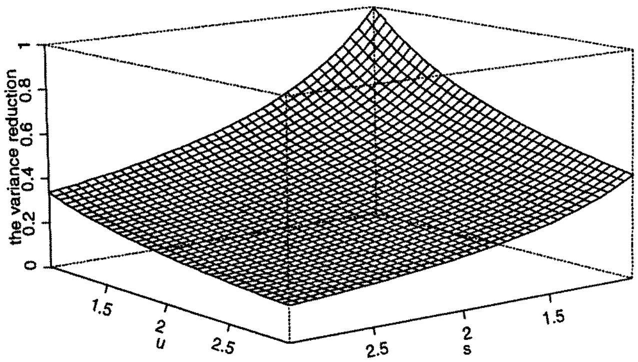

Figure 4

5 Using the smoothed Hill estimator

This section discusses the use of Proposition 4.1 for estimating α. Several remarks and examples about the alternative Hill plot introduced in the first section of the paper end the section.

The first point to be made is that Proposition 4.1 offers considerable assistance in estimating α. Proposition 4.1 lessens the emphasis on choosing the right number of upper order statistics needed to calculate the estimate. Due to the fact that the range of the smoothed estimator is reduced in comparison to the classical Hill estimator, the importance of wisely selecting k diminishes.

Several previous studies advocate choosing k to minimize the asymptotic MSE of Hill's estimator (Hall (1982), Dekkers and de Haan (1991)). For example, adapting the method of Dekkers and de Haan (1991) one finds that for the Cauchy density

$$F'(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$$

the optimal choice $k^* = k^*(n)$ of k is

$$k^* \sim \frac{n^{4/5}}{\pi^{4/5}(\frac{16}{81})^{1/5}}, \quad (n \to \infty).$$

For $n = 1000$ for instance, the right side evaluates to $k^* = 139$. There are two problems with such formulas. First, these formulas require one to know the distribution rather explicitly and thus, although interesting and welcome, are of dubious practical value. Second, the formulas are frequently only asymptotic formulas and the asymptotic equivalence is not helpful for finite samples. If k^* has the asymptotic form given above, then the following is equally acceptable as an asymptotic solution:

$$k_1^* = (1 + \frac{10^{96}}{n})k^*.$$

Evaluating the right side now for $n = 1000$ would yield a rather different value from the one obtained previously, namely $(1 + 10^{93})k^*$. Even if one accepts a value of k^* for finite n from a displayed formula, these seldom work well in practice. For example, in Figure 5 we have displayed the alternative Hill and alternative avHill plots for 1000 iid data from a Cauchy density and evaluating the plot at $k = 139$ gives disappointing results, namely the estimate given by the horizontal line. Therefore, we conclude that attempts to compute an optimal value of k have little applied value and thus we were led to seek alternative methods of inferring the value of α.

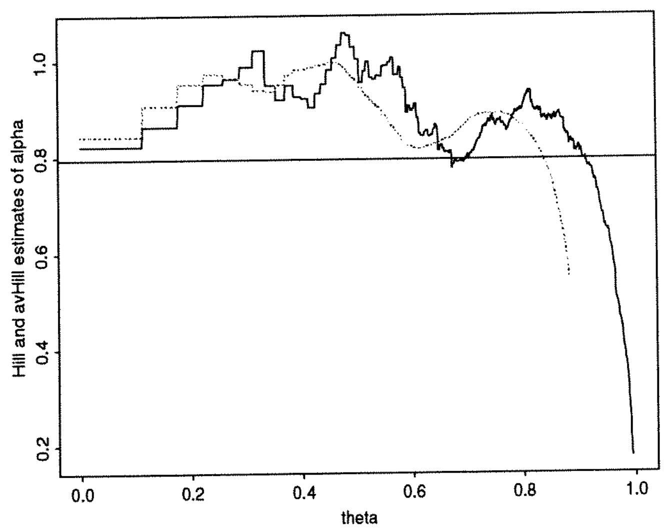

Figure 5

In order to use the traditional Hill plot, there are at least two important decisions to be made. First, a sensible range of k has to be determined. Once this decision is made, the practitioner is usually confronted with a second difficult choice, namely deciding on the specific value of k inside the range which gives the point estimate. This is problematic due to volatility inside the range. It is this second step where our procedure offers assistance: Assuming we believe that $k \in [k_1, k_2]$ is suitable range, we make the avHill plot

$$\left\{ (k, \left(\frac{1}{k(u-1)} \sum_{p=k}^{uk} H_{p,n} \right)^{-1}), k \in [k_1, k_2/u] \right\}.$$

Due to the fact that the range of the smoothed estimator is reduced in comparison to classical Hill, the importance of selecting the optimal k diminishes.

What about the choice of u? Since the asymptotic variance is a decreasing function of u, one would like to choose u as big as possible to ensure the maximum decrease of the variance. However the choice of u is limited by the sample size. Due to the averaging, the bigger the u, the fewer points one gets on the avHill plot. Therefore an equilibrium should be reached between variance reduction and a comfortable number of points on the plot. Usually we use u between $n^{0.1}$ and $n^{0.2}$ where n is the sample size.

We conclude this section with a discussion of the alternative Hill plot introduced in the beginning of the paper.

The alternative Hill plot was defined as the plot displaying the information provided by the estimation as

$$\{(\theta, H^{-1}_{\lceil n^{\theta} \rceil, n}), 0 \le \theta \le 1\}.$$

As we have already mentioned, a significant advantage of the alternative plot over the Hill plot is the way the displayed space is used. In the alternative plot the significant part of the graph, i.e. the part corresponding to a relatively small number of order statistics gets to be shown more clearly, covering a bigger portion of the displayed space. On the other hand, the part of the graph corresponding to a high number of order statistics, which covers a disproportionately large part in the Hill plot gets rescaled. Thus, interpretation of the graph is easier and more accurate.

Here is an example where the difference is striking. The data is drawn from a distribution defined by

$$X = Y \log(Y), \; P(Y > x) = x^{-1}. \tag{5.1}$$

The length of the sample is 60,000. The results are shown in Figure 6.

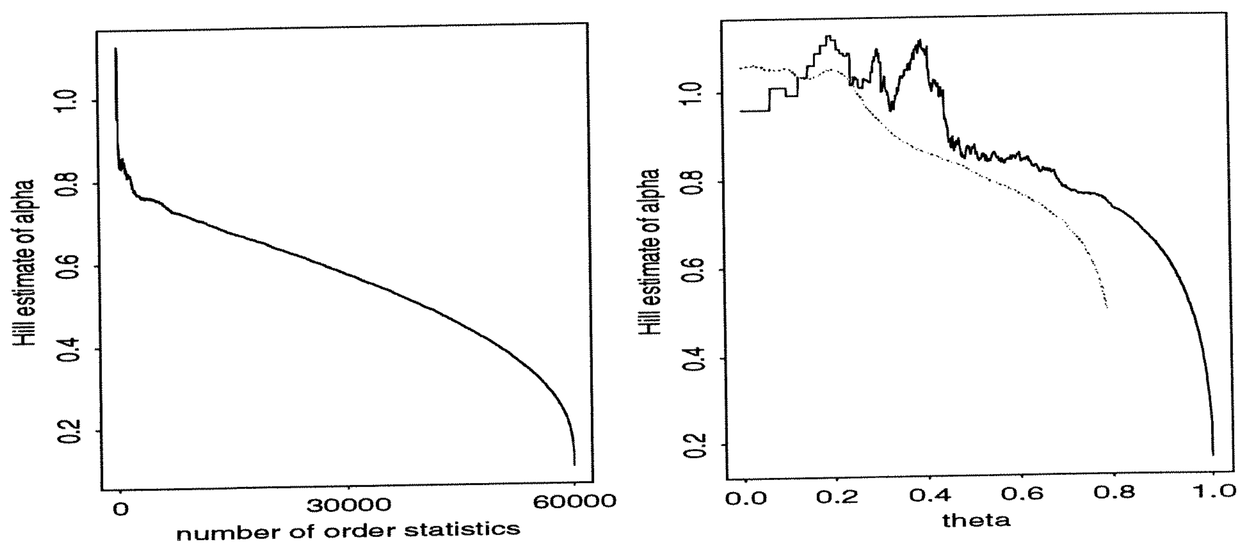

Figure 6

The left hand side of the figure shows the traditional Hill plot and the right hand side shows the altenative plot of the Hill's estimator (the solid graph) together with the smoothed version (the dotted graph).

As in the case of the Hill plot, when dealing with the alternative plot one tries to read the estimated value from a stable portion of the graph. Compare the two plots. The Hill plot does not look too encouraging and can be misleading. A rushed conclusion would be that the estimate is around 0.8. In the case of the alternative Hill plot the graph is put in a better perspective and we can clearly see that for an important number of upper order statistics, i.e. between 1 and $60000^{0.45}$, the estimate is around 1. The alternative plot leaves us with less doubt about the estimate. For sure one could display only the initial portion of the traditional Hill plot on an expanded scale. Even then the significant part of the graph does not get enough of the display space.

We end the section with an example of a set of real data which exhibits large values and seems to be generated by a sequence of independent random variables. The data set represents the interarrival times between packets generated and sent to a host by a terminal during a logged-on session. The terminal hooks to a network through a host and communicates with the network sending and receiving packets. The length of the periods between two consecutive packets received by the host were recorded as our data. The total length of the data is 783. Figure 7 is a time series plot of our data set showing indications of heavy tails.

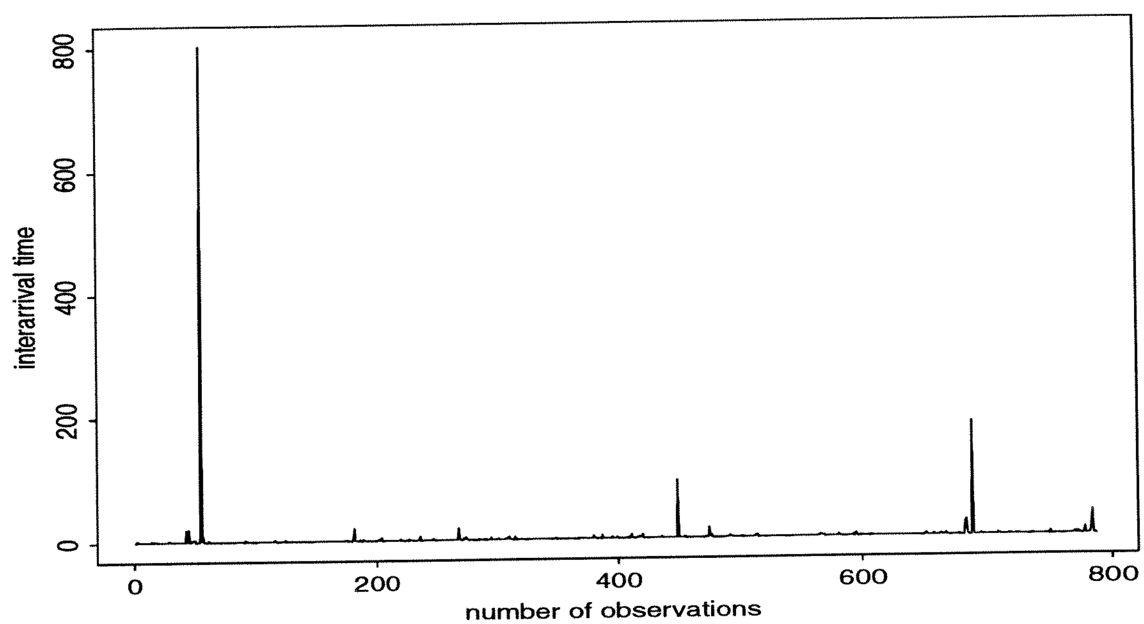

Figure 7

The left hand side of the Figure 8 shows the Hill plot, the right hand side displays the Hill's estimator (the solid graph) together with the smoothed version (the dotted graph) in an alternative Hill plot.

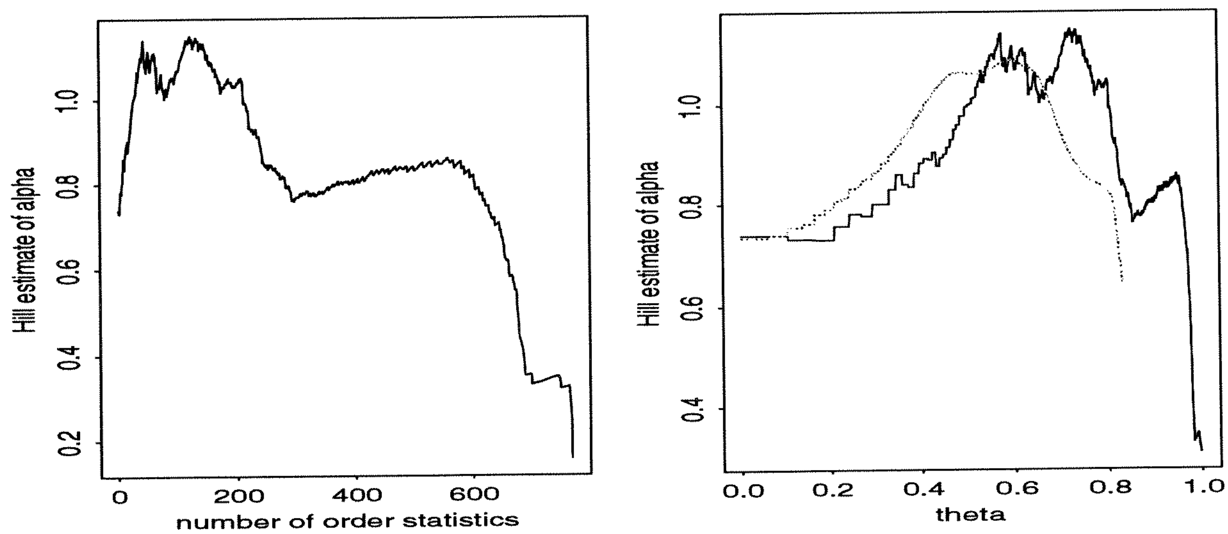

Figure 8

One can see again that the Hill plot can be misleading by visually emphasizing a value close to 0.85. However the number of order statistics used to get an estimate around 0.85 is very large. The alternative plot puts things in better perspective and yields a value around 1.1. One can also see how efficent the smoothing procedure is in this case. The range of the smoothed estimator is seriously reduced in comparison to classical Hill. No matter what two estimates are chosen from the stable part of the graph they are not going to differ by more then 0.02. Therefore the importance of selecting the optimal k diminishes.

Aknowledgement. Paul Feigin suggested averaging the Hill estimator when working on Feigin, Resnick and Stărică (1994). Our endeavour was to provide this procedure with the mathematical backing it needed.

References

[1] Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.

[2] Bingham, N., Goldie, C. and Teugels, J. (1987). Regular Variation. Cambridge University Press, Cambridge, UK.

[3] Csörgő, S., Deheuvels, P. and Mason, D. (1985). Kernel estimates for the tail index of a distribution. Ann. Statist. 13, 1050–1077.

[4] Csörgő, S. and Mason, D. (1985). Central limit theorems for sums of extreme values. Math. Proc. Camb. Philos. Soc. 98, 547–588.

[5] Deheuvels, P. and Mason, D. M. (1991). A tail empirical process approach to some non-standard laws of the iterated logarithm. J. Theor. Probab., 4, 53–85.

[6] Dekkers, A. L. M. and Haan, L. de (1991). Optimal choice of sample fraction in extreme value estimation. Thesis. Erasmus University, Rotterdam.

[7] Feigin, P. D., Resnick, S. I. and Stărică, C. (1994). Testing for independence in heavy tailed and positive innovation time series. Preprint

[8] Geluk, J and Haan, L. de (1987). Regular Variation, Extensions and Tauberian Theorems. CWI tract 40, Centre for Mathematics and Computer Science, Amsterdam, Netherlands.

[9] Goldie, C. and Smith, R. (1987). Slow variation with remainder: Theory and applications. Quarterly Journal of Mathematics 38, 45–71.

[10] Haan, L. de and Stadtmüller, U. (1994). Generalized regular variation of second order. Preprint.

[11] Haan, L. de and Resnick, S. I. (1994). Estimating the limit distribution of multivariate extremes. Stochastic Models 9, 275–309.

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

[13] Häusler, E. and Teugels, J. (1985). On the asymptotic normality of Hill's estimator for the exponent of regular variation. Ann. Statist. 13, 743–756

[14] Hill, B. (1975). A simple approach to inference about the tail of a distribution. Ann. Statist. 3, 1163–1174.

[15] Hsing, T. (1991). On tail estimation using dependent data. Ann. Statist. 19, 1547–1569.

[16] Mason, D. (1982). Laws of large numbers for sums of extreme values. Ann. Probability 10, 754–764.

[17] Mason, D. (1988). A strong invariance theorem for the tail empirical process. Ann. Inst. Henri Poincaré 24, 491-506.

[18] Mason, D. M. and Turova, T. S., (1994). Weak convergence of the Hill estimator process. Preprint.

[19] Pollard, D. (1990). Empirical Processes: Theory and Appliv cations. NSF-CBMS Regional Conference Series in Probability and Statistics, vol. 2.

[20] Resnick, S. (1987). Extreme Values, Regular Variation and Point Processes. Springer-Verlag, Berlin.

[21] Resnick, S. and Stărică, C. (1995) Consistency of Hill's estimator for dependent data. Journal of Applied Probability, to appear.

[22] Rootzen, H. , Leadbetter, M. and De Haan, L. (1990). Tail and quantile estimation for strongly mixing stationary sequences. Preprint, Econometric Institute.

[23] Vervaat, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. Z. Wahrscheinlichkeitstheorie verw. Geb. 23, 249–253