The Annals of Applied Probability
1998, Vol. 8, No. 4, 1156–1183

TAIL INDEX ESTIMATION FOR DEPENDENT DATA[1]

BY SIDNEY RESNICK[2] AND CATALIN STĂRICĂ

Cornell University and University of Pennsylvania

A popular estimator of the index of regular variation in heavy-tailed models is Hill's estimator. We discuss the consistency of Hill's estimator when it is applied to certain classes of heavy-tailed stationary processes. One class of processes discussed consists of processes which can be appropriately approximated by sequences of m-dependent random variables and special cases of our results show the consistency of Hill's estimator for (i) infinite moving averages with heavy-tail innovations, (ii) a simple stationary bilinear model driven by heavy-tail noise variables and (iii) solutions of stochastic difference equations of the form

$$Y_t = A_t Y_{t-1} + Z_t, \quad -\infty < t < \infty$$

where $\{(A_n, Z_n), -\infty < n < \infty\}$ are iid and the Z's have regularly varying tail probabilities. Another class of problems where our methods work successfully are solutions of stochastic difference equations such as the ARCH process where the process cannot be successfully approximated by m-dependent random variables. A final class of models where Hill estimator consistency is proven by our tail empirical process methods is the class of hidden semi-Markov models.

1. Introduction. This paper discusses how to estimate the Pareto index or the index of regular variation for stationary dependent sequences. If $\{X_t, -\infty < t < \infty\}$ is a stationary time series with the property that

$$P[X_t > x] \sim x^{-\alpha} L(x), \qquad x \to \infty,$$

L being a slowly varying function, then a key question in tail estimation is how to estimate the index α. A popular estimator which arose in the iid context as a conditional maximum likelihood estimator is Hill's estimator [Hill (1975)], which is defined as follows. For $1 \le i \le n$, write $X_{(i)}$ for the ith largest value of $X_1, X_2, \dots, X_n$. Hill's estimator based on the observations $X_1, \dots, X_n$ is

$$H^X_{k,n} = \frac{1}{k} \sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k+1)}}. \tag{1.1}$$

This estimator has been well studied when $\{X_n\}$ is iid [Hall (1982), Mason (1982, 1988), Mason and Turova (1994), de Haan and Resnick (1998),

Received September 1996; revised December 1997.
[1]Supported in part by NSF Grant DMS-94-00535 at Cornell University.
[2]Support from NSA Grant MDA904-95-H-1036.
AMS 1991 *subject classifications*. Primary 60G70; secondary 60G10.
Key words and phrases. Hill estimation, tail estimator, heavy tails, tail empirical process, ARCH model, bilinear model, moving average, hidden Markov model.

Geluk et al. (1997), Davis and Resnick (1984), Häusler and Teugels (1985) and Resnick and Stărică (1997a, b)] and our goal here is to better understand its behavior when it is applied to stationary dependent sequences. Related papers which study Hill's estimator in the dependent case are Hsing (1991), Rootzen, Leadbetter and de Haan (1990) and Rootzen (1995).

A great deal of time series analysis has been based on the assumption that the structure of the series can be described by linear models. In the traditional setting of a stationary time series with finite variance, every purely nondeterministic process can be expressed as a linear process driven by an uncorrelated input sequence. From a second-order point of view, linear models are sufficient for data analysis. The situation is totally different when the stationary series has heavy tails and perhaps infinite variance. In this case we have no such confidence that heavy-tailed linear models are sufficiently flexible and rich enough for modeling purposes and, in any case, for heavy-tailed infinite-order moving averages it is already known [Resnick and Stărică (1995) and see also Section 3] that Hill's estimator is consistent. Thus in this paper we concentrate on nonlinear models.

Linear models do not seem to describe adequately the underlying random mechanism when heavy tails are present [Davis and Resnick (1996) and Resnick (1998)]. Insistence upon modeling heavy-tailed data with linear time series can be quite misleading [Feigin and Resnick (1996)]. A popular nonlinear alternative to the linear model is the bilinear process introduced by Mohler (1973) and considered by Granger and Andersen (1978). To date, little use has been made of bilinear models in heavy-tailed data analysis, though Davis and Resnick (1996) present some evidence for their relevance. Other worthy nonlinear models which we consider are two classes of random coefficient models, one of which includes the important example of the ARCH process [Engle (1982)] and hidden semi-Markov models or random variables defined on a semi-Markov chain. Such models have recently been used to fit times between packet transmissions at a terminal in the stimulating paper by Meier-Hellstern, Wirth, Yan and Hoeflin (1991).

Section 2 presents two general theorems which can be applied to prove the consistency of Hill's estimator for heavy-tailed stationary sequences. Section 3 applies one of the theorems to the case of processes which can be approximated by m-dependent sequences. Among the examples considered are infinite-order moving averages, simple bilinear processes and solutions of certain random coefficient autoregressions. Section 4 applies the other theorem from Section 2 to a class of random coefficient autoregressions which includes the first-order ARCH process. This result yields not only an estimator for the Pareto index of the ARCH process but also an estimator of one of the scaling parameters. Section 5 deals directly with hidden semi-Markov models using Laplace functional methods.

The tail empirical measure plays a central role in our approach to proving the consistency of Hill's estimator. This method was also used in Resnick and Stărică (1995). For using this method, we need the following notation. Let $\mathbb{E} := (0, \infty]$ be the one point uncompactification of $[0, \infty]$ so that the compact sets of

$\mathbb{E}$ are of the form U^c, where $0 \in U$ and U is an open set in $[0, \infty)$. Suppose $\mathscr{E}$ is the Borel σ-field on $\mathbb{E}$. Let $M_+(\mathbb{E})$ be the space of positive Radon measures on $\mathscr{E}$ endowed with the vague topology [Resnick (1987) and Kallenberg (1983)]. Let $C_K^+(\mathbb{E})$ be the space of continuous, nonnegative functions on $\mathbb{E} = (0, \infty]$ with compact support. The vague topology on $M_+(\mathbb{E})$ can be generated by a countable family of semi-norms

$$H = \left\{ p_f \colon M_+(\mathbb{E}) \to \mathbb{R}_+ \colon p_f(\mu) = \mu(f), |f| \le 1, f \in C_K^+(\mathbb{E}) \right\}$$

[Resnick (1987), Proposition 3.17 and Lemma 3.11], turning $M_+(\mathbb{E})$ into a complete, separable, metric space. Convergence of $\mu_n \in M_+(\mathbb{E})$ to $\mu_0 \in M_+(\mathbb{E})$ in the vague topology is denoted $\mu_n \to_v \mu_0$. For $x \in \mathbb{E}$ and $A \in \mathscr{E}$ define

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

2. General consistency results. We now prove two general Hill estimator consistency results for heavy-tailed stationary sequences. The first, Proposition 2.1, is designed to be easily specialized for processes which can be approximated by m-dependent sequences and this specialization comes in Proposition 2.2. Proposition 2.3 is similar to Proposition 2.1 but is better suited for application to the ARCH model (cf. Section 4). The proofs of Propositions 2.1 and 2.3 use the standard big block–little block technique explained carefully and exploited in Leadbetter, Lindgren and Rootzen (1988). See also Hsing, Hüsler and Leadbetter (1988) as well as Davis and Resnick (1988) where a parallel result for Poisson convergence is given.

PROPOSITION 2.1. *Suppose for each $n = 1, 2, \dots$ that $\{X_{n,i}, i \ge 1\}$ is a stationary sequence of random elements of $\mathbb{E}$. Let $\{k = k(n)\}$ be a sequence such that $k \to \infty$, $n/k \to \infty$. Suppose $\{X_{n,i}\}$ satisfies the following two conditions*:

(i) *For any $f \in C_K^+(\mathbb{E})$,*

$$\lim_{n\to\infty} \frac{n}{k^2} \sum_{j=2}^{k} E\left(f(X_{n,1}) f(X_{n,j})\right) = 0. \tag{2.1}$$

(ii) *For any sequence $\{l_n\}$ such that $l_n \to \infty$ and*

$$\frac{l_n}{k} \to 0 \tag{2.2}$$

and intervals

$$I_1 = [1, k - l_n], I_2 = [k+1, 2k - l_n], \dots, I_{[n/k]} = [([n/k]-1)k, [n/k]k - l_n], \tag{2.3}$$

we have, for $f \in C_K^+(\mathbb{E})$,

$$\lim_{n\to\infty} E\left(\prod_{j=1}^{[n/k]} \exp\left\{-\frac{1}{k}\sum_{i\in I_j} f(X_{n,i})\right\}\right) - \prod_{j=1}^{[n/k]} E\left(\exp\left\{-\frac{1}{k}\sum_{i\in I_j} f(X_{n,i})\right\}\right) = 0. \tag{2.4}$$

Assume also that

$$\frac{n}{k} P(X_{n,1} \in \cdot) \to_v \nu, \tag{2.5}$$

where $\nu(\{x\}) = 0$ *for any* $x \in (0, \infty]$. *Then*

$$\nu_n := \frac{1}{k}\sum_{i=1}^{n} \varepsilon_{X_{n,i}} \Rightarrow \nu \tag{2.6}$$

in $M_+(\mathbb{E})$. *Moreover, if* $X_{n,i} = X_i/b_n$, $i = 1, \dots, n$, *where* $\{X_n, n \geq 1\}$ *is a sequence of stationary random variables and* $b_n \to \infty$ *and if* ν *satisfies* $\int_1^\infty \log(u)\nu(du) < \infty$, *it also follows that*

$$H_{k,n}^X := \frac{1}{k}\sum_{i=1}^{k} \log\frac{X_{(i)}}{X_{(k+1)}} \to_P \int_1^\infty \log(u)\nu(du). \tag{2.7}$$

REMARK. Condition (2.1) is implied by the condition that, for any $x > 0$,

$$\lim_{n\to\infty} \frac{n}{k^2}\sum_{j=2}^{k} P[X_{n,1} > x, X_{n,j} > x] = 0. \tag{2.8}$$

This follows since if $f \in C_K^+(\mathbb{E})$ and we set $[c, \infty]$ for the support of f and set $\|f\| = \sup_{\mathbb{E}} f(x)$, then

$$f \leq \|f\| 1_{[c,\infty]}$$

and

$$E\big(f(X_{n,1}) f(X_{n,j})\big) \leq \|f\|^2 P[X_{n,1} > c, X_{n,j} > c].$$

PROOF. Suppose $f \in C_K^+(\mathbb{E})$. To show (2.6), it suffices to show [Kallenberg (1983) and Resnick (1987)]

$$\lim_{n\to\infty} E\exp\left\{-\frac{1}{k}\sum_{i=1}^{n} f(X_{n,i})\right\} = \exp\{-\nu(f)\}. \tag{2.9}$$

For typographical ease, we write $f_i = f(X_{n,i})$ and $p = [n/k]$. Then

$$
\begin{aligned}
I_j &= \{(j-1)k+1, \dots, jk-l_n\}, \\
I_j^* &= \{jk-l_n+1, \dots, jk\}, \qquad j = 1, \dots, p-1,
\end{aligned} \tag{2.10}
$$

and

$$
I_p = \{(p-1)k+1, \dots, pk-l_n\}, \qquad I_p^* = \{pk-l_n+1, \dots, n\}. \tag{2.11}
$$

We have

$$
\begin{aligned}
&\left| E\exp\left\{-\frac{1}{k}\sum_{i=1}^{n} f_i\right\} - \exp\{-\nu(f)\}\right| \\
&\quad \le \left| E\exp\left\{-\frac{1}{k}\sum_{i=1}^{n} f_i\right\} - E\exp\left\{-\frac{1}{k}\sum_{j=1}^{p}\sum_{i\in I_j} f_i\right\}\right| \\
&\qquad + \left| E\exp\left\{-\frac{1}{k}\sum_{j=1}^{p}\sum_{i\in I_j} f_i\right\} - \left(E\exp\left\{-\frac{1}{k}\sum_{i\in I_j} f_i\right\}\right)^p\right| \\
&\qquad + \left| \left(E\exp\left\{-\frac{1}{k}\sum_{i\in I_j} f_i\right\}\right)^p - \left(E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\}\right)^p\right| \\
&\qquad + \left| \left(E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\}\right)^p - \exp\{-\nu(f)\}\right| \\
&\quad = I + II + III + IV.
\end{aligned}
$$

Let us look at the individual terms in turn.

We have

$$
\begin{aligned}
I &\le \left| E\exp\left\{-\frac{1}{k}\sum_{j=1}^{p}\left(\sum_{i\in I_j} f_i + \sum_{i\in I_j^*} f_i\right)\right\}\right| \\
&\le E\left| 1 - \exp\left\{-\frac{1}{k}\sum_{j=1}^{p}\sum_{i\in I_j^*} f_i\right\}\right| \\
&\le \sum_{j=1}^{p}\sum_{i\in I_j^*}\frac{1}{k}Ef_i \\
&\le pl_n\frac{1}{k}Ef_1 \sim \frac{l_n}{k}\frac{n}{k}Ef_1 \\
&\sim \frac{l_n}{k}\nu(f) \to 0
\end{aligned}
$$

as $n \to \infty$ from (2.2) and (2.5). Term III is handled very similarly:

$$\begin{aligned} III &\le p\left|E\exp\left\{-\frac{1}{k}\sum_{i\in I_1} f_i\right\} - E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\}\right| \\ &\le pE\left|1-\exp\left\{-\frac{1}{k}\sum_{i\in I_1^*} f_i\right\}\right| \\ &\le p\frac{l_n}{k}Ef_1 \to 0. \end{aligned}$$

Term II goes to 0 because of condition (2.4).

For term IV, we set $y_i = 1-\exp\{-(1/k)f_i\}$ and observe

$$\begin{aligned} E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\} &= E\prod_{i=1}^{k}(1-y_i) \\ &\le 1 - E\sum_{i=1}^{k} y_i + E\sum_{1\le i<j\le k} y_i y_j \\ &\le 1 - kEy_1 + k\sum_{l=2}^{k} Ey_1 y_l \end{aligned}$$

and thus

$$\left(E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\}\right)^p \le \left(1-\frac{kp(Ey_1-\sum_{l=2}^{k} Ey_1y_l)}{p}\right)^p.$$

Now

$$\begin{aligned} kp\sum_{l=2}^{k} Ey_1y_l &\sim n\sum_{l=2}^{k} E\left(1-\exp\left\{-\frac{1}{k}f_1\right\}\right)\left(1-\exp\left\{-\frac{1}{k}f_l\right\}\right) \\ &\le \frac{n}{k^2}\sum_{l=2}^{k} Ef_1f_l \to 0 \end{aligned}$$

by (2.1). Also

$$kpEy_1 \sim nE\left(1-\exp\left\{-\frac{1}{k}f_1\right\}\right) \le \frac{n}{k}Ef_1 \to \nu(f)$$

and

$$\begin{aligned} nE\left(1-\exp\left\{-\frac{1}{k}f_1\right\}\right) &\ge nE\left(\frac{f_1}{k}-\frac{f_1^2}{2k^2}\right) \\ &\sim \nu(f) - \frac{1}{2k}\nu(f^2) \to \nu(f). \end{aligned}$$

Thus we conclude

$$\limsup_{n\to\infty}\left(E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\}\right)^p \le \exp\{-\nu(f)\}.$$

A slightly simpler argument gives

$$\liminf_{n\to\infty}\left(E\exp\left\{-\frac{1}{k}\sum_{i=1}^{k} f_i\right\}\right)^p \ge \exp\{-\nu(f)\}$$

and this completes the proof of (2.9).

To prove (2.7), we make use of Proposition 2.4 of Resnick and Stărică (1995) which shows that the convergence of the tail measure implies the consistency of Hill's estimator. □

Proposition 2.1 will be applied primarily to proving the consistency of Hill's estimator for stationary processes which can be approximated by truncated versions which are m-dependent. In order for this approximation strategy to be successful, the truncated m-dependent approximation must carry enough information about the tail behavior of the marginal distribution of the original process $\{X_t\}$. This is true of the processes considered in the examples in Section 3 and false for certain random coefficient models such as the ARCH process considered in Section 4. The adaptation of Proposition 2.1 to processes which can be successfully approximated by m-dependent processes is given next.

PROPOSITION 2.2. *Suppose, for each $n \ge 1$, $m \ge 1$, $\{X_{n,i}^{(m)}, i \ge 1\}$ is a stationary sequence of m-dependent random elements of $\mathbb{E}$ and, for each $n \ge 1$, $\{X_{n,i}, i \ge 1\}$ is a stationary sequence of random elements of $\mathbb{E}$. Suppose there exist Radon measures $\nu^{(m)}$ on $\mathbb{E}$ and a sequence $k = k(n)$, $k \to \infty$ and $n/k \to \infty$, such that, for any fixed $m \ge 1$,*

$$\frac{n}{k}P\big(X_{n,i}^{(m)} \in \cdot\big) \to_v \nu^{(m)} \tag{2.12}$$

as $n \to \infty$. Suppose further that

$$\nu^{(m)} \to_v \nu \tag{2.13}$$

as $m \to \infty$. Finally, assume that

$$\lim_{m\to\infty}\limsup_{n\to\infty}\frac{n}{k}P\big(\big|X_{n,1}^{(m)} - X_{n,1}\big| > \varepsilon\big) = 0 \tag{2.14}$$

for all $\varepsilon > 0$. Then

$$\frac{1}{k}\sum_{i=1}^{n}\varepsilon_{X_{n,i}} \Rightarrow \nu \tag{2.15}$$

in $M_+(\mathbb{E})$.

Moreover, if $X_{n,i} = X_i/b_n$, $i = 1, \ldots, n$, *where* $\{X_n, n \geq 1\}$ *is a stationary sequence*, $b_n \to \infty$ *and* ν *satisfies* $\int_1^\infty \log(u)\nu(du) < \infty$, *then*

$$H^X_{k,n} := \frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{(i)}}{X_{(k+1)}} \to_P \int_1^\infty \log(u)\nu(du). \tag{2.16}$$

PROOF. We first show that, for any fixed m,

$$\frac{1}{k}\sum_{i=1}^{n}\varepsilon_{X^{(m)}_{i,n}} \Rightarrow \nu^{(m)} \tag{2.17}$$

by checking that the hypotheses of Proposition 2.1 hold for $\{X^{(m)}_{n,i}, i \geq 1\}$. Since (2.12) holds, we need only check conditions (2.1) and (2.4). Condition (2.4) holds trivially since, for $l(n) > m$,

$$\left\{\sum_{i\in I_j} f\left(X^{(m)}_{i,n}\right),\ j = 1, \ldots, p\right\}$$

are independent random variables. To check condition (2.1), note that

$$\begin{aligned}
&\frac{n}{k^2}\sum_{j=2}^{k} P\left(X^{(m)}_{n,1} > x, X^{(m)}_{n,j} > y\right) \\
&\quad\leq \frac{n}{k^2}\left(\sum_{j=2}^{m} P\left(X^{(m)}_{n,1} > x, X^{(m)}_{n,j} > y\right) + \sum_{j=m+1}^{k} P\left(X^{(m)}_{n,1} > x, X^{(m)}_{n,j} > y\right)\right) \\
&\quad\leq \frac{n}{k^2}\sum_{j=2}^{m} P\left(X^{(m)}_{n,j} > y\right) + \frac{n}{k}P\left(X^{(m)}_{n,1} > x\right)P\left(X^{(m)}_{n,1} > y\right) \\
&\quad= \frac{(m-1)n}{k^2}P\left(X^{(m)}_{n,1} > y\right) + \frac{k}{n}\frac{n^2}{k^2}P\left(X^{(m)}_{n,1} > x\right)P\left(X^{(m)}_{n,1} > y\right) \\
&\quad= \frac{m-1}{k}\left(\nu^{(m)}\left((y,\infty]\right) + o(1)\right) + \frac{k}{n}\left(\nu^{(m)}\left((x,\infty]\right)\nu^{(m)}\left((y,\infty]\right) + o(1)\right).
\end{aligned}$$

Therefore

$$\lim_{n\to\infty}\frac{n}{k^2}\sum_{j=2}^{k} P\left(X^{(m)}_{n,1} > x, X^{(m)}_{n,j} > y\right) \to 0,$$

which completes the proof of (2.17).

The proof of (2.15) follows from a converging together argument similar to the proof of Proposition 3.3 in Resnick and Stărică (1995). The conclusion (2.16) follows from Proposition 2.4 of Resnick and Stărică (1995) which shows that the convergence of the tail measure implies the consistency of Hill's estimator. □

For dealing with the ARCH process in Section 4, it is better to have a version of Proposition 2.1 adapted for use with sets rather than $C_K^+(\mathbb{E})$ functions. This is given next.

PROPOSITION 2.3. *Suppose all the assumptions of Proposition* 2.1 *hold except that in place of condition* (2.4) *we assume*

$$\lim_{n\to\infty}\left|E\prod_{j=1}^{[n/k]}\left(1-\frac{1}{k}\sum_{i\in I_j}f(X_{n,i})\right)-\prod_{j=1}^{[n/k]}E\left(1-\frac{1}{k}\sum_{i\in I_j}f(X_{n,i})\right)\right|=0 \tag{2.18}$$

for any function f of the form $f=\sum_{h=1}^{s}\beta_h 1_{(x_h,\infty]}$ where $\beta_h>0$, $h=1,\dots,s$, and $x_h>0$, $h=1,\dots,s$. Then the conclusions of Proposition 2.1 *hold.*

PROOF. We will use the fact that $\nu_n\Rightarrow\nu$ in $M_+(\mathbb{E})$ provided

$$\big(\nu_n(I_1),\dots,\nu_n(I_s)\big)\Rightarrow\big(\nu(I_1),\dots,\nu(I_s)\big),\qquad n\to\infty, \tag{2.19}$$

in $\mathbb{R}^s$ for any s and any intervals $I_i=(x_i,\infty]$, $i=1,2,\dots,s$ [Kallenberg (1983)]. Using multivariate Laplace transforms, we must show that, for any positive $\beta_1,\dots,\beta_s$ and $f=\sum_{h=1}^{s}\beta_h 1_{(x_h,\infty]}$,

$$E\exp\left(-\frac{1}{k}\sum_{i=1}^{n}f(X_{n,i})\right)\to\exp(-\nu(f)). \tag{2.20}$$

Define blocks I_j, I_j^* as in Proposition 2.1 and decompose

$$\begin{aligned}
&\left|E\exp\left(-\frac{1}{k}\sum_{i=1}^{n}f(X_{n,i})\right)-\exp(-\nu(f))\right|\\
&\quad\le\left|E\exp\left(-\frac{1}{k}\sum_{i=1}^{n}f(X_{n,i})\right)-E\exp\left(-\frac{1}{k}\sum_{j=1}^{p}\sum_{i\in I_j}f(X_{n,i})\right)\right|\\
&\qquad+\left|E\exp\left(-\frac{1}{k}\sum_{j=1}^{p}\sum_{i\in I_j}f(X_{n,i})\right)-E\prod_{j=1}^{p}\left(1-\frac{1}{k}\sum_{i\in I_j}f(X_{n,i})\right)\right|\\
&\qquad+\left|E\prod_{j=1}^{p}\left(1-\frac{1}{k}\sum_{i\in I_j}f(X_{n,i})\right)-\prod_{j=1}^{p}E\left(1-\frac{1}{k}\sum_{i\in I_j}f(X_{n,i})\right)\right|\\
&\qquad+\left|\left(1-\frac{1}{p}\frac{p(k-l_n)}{k}Ef(X_{n,1})\right)^{p}-\exp(-\nu(f))\right|\\
&\quad=I+II+III+IV.
\end{aligned}$$

Term I is controlled as in Proposition 2.1. For II, denote $Q := \sum_h \beta_h = \sup_{\mathbb{E}} f(x)$ and we have

$$
\begin{aligned}
II &\le E \sum_{j=1}^{p} \left| \exp\biggl(-\frac{1}{k} \sum_{i \in I_j} f(X_{n,i})\biggr) - 1 + \frac{1}{k} \sum_{i \in I_j} f(X_{n,i}) \right| \\
&\le \frac{p}{2k^2} E \biggl(\sum_{i=1}^{k-l_n} f(X_{n,i}) \biggr)^2 \\
&\le \frac{n}{2k^2} E f^2(X_{n,1}) + \frac{n}{k^2} \sum_{j=2}^{k} E f(X_{n,1}) f(X_{n,j}).
\end{aligned}
$$

By condition (2.1) and (2.5), it follows that $\lim_{n\to\infty} II = 0$. Condition (2.18) is equivalent to $\lim_{n\to\infty} III = 0$. By (2.5) and (2.2),

$$
\biggl(1 - \frac{1}{p} \frac{p(k-l_n)}{k} E f(X_{n,1})\biggr)^p \to \exp\bigl(-\nu(f)\bigr),
$$

and the conclusion (2.6) of the proposition follows. The rest is the same as in Proposition 2.1. □

3. Examples. We now consider three examples of heavy-tailed, dependent, stationary processes which have m-dependent approximations and in each case we apply Proposition 2.2 to demonstrate the consistency of Hill's estimator. The three classes of processes are:

1. infinite-order moving averages of iid heavy-tailed random variables,
2. bilinear processes driven by heavy-tailed innovations and
3. processes satisfying a simple stochastic difference equation with random coefficients.

The first two processes are constructed using a sequence $\{Z_t, -\infty < t < \infty\}$ of iid random variables which, for simplicity, we take to be positive. These random variables have regularly varying tail probabilities; that is, for $x > 0$,

$$
P[Z_1 > x] =: 1 - F(x) =: \bar{F}(x) = x^{-\alpha} L(x), \qquad \alpha > 0, \tag{3.1}
$$

where L is a slowly varying function at ∞.

3.1. *Infinite order moving averages.* Suppose that the sequence $\{c_i, i \ge 0\} \in \mathbb{R}^\infty$ contains at least one positive number and satisfies

$$
0 < \sum_{j=0}^{\infty} |c_j|^\delta < \infty \tag{3.2}
$$

for some $0 < \delta < \alpha \wedge 1$. Then [cf. Cline (1983)] almost surely $\sum_{j=0}^{\infty} c_j Z_j$ converges absolutely and

$$\lim_{x\to\infty} \frac{P(\sum_{j=0}^{\infty} c_j Z_j > x)}{P(Z_1 > x)} = \sum_{\substack{j=0 \\ c_j>0}}^{\infty} c_j^{\alpha} \tag{3.3}$$

so that $\sum_{j=0}^{\infty} c_j Z_j$ also has regularly varying tail probabilities.

Define the moving average of order infinity processes, denoted MA(∞), by

$$X_t = \sum_{j=0}^{\infty} c_j Z_{t-j}, \qquad -\infty < t < \infty. \tag{3.4}$$

Causal ARMA processes can be represented in the form (3.4) [Brockwell and Davis (1991), Chapter 3]. The consistency of the Hill estimator for MA(∞) processes was considered in detail in Resnick and Stărică (1995). See also Resnick and Stărică (1997b).

3.2. *The simple bilinear model.* Let X_t be the stationary bilinear model

$$X_t = cX_{t-1}Z_{t-1} + Z_t, \qquad -\infty < t < \infty, \tag{3.5}$$

where $c > 0$ is a positive constant satisfying

$$c^{\alpha/2} E Z_1^{\alpha/2} < 1. \tag{3.6}$$

Using the bilinear recursion formula (3.5), X_t can be written as an infinite series whose convergence is guaranteed by (3.6) [see Davis and Resnick (1996)],

$$X_t = \sum_{j=0}^{\infty} c^j X_t^{(j)}, \tag{3.7}$$

where

$$X_t^{(0)} = Z_t, \qquad X_t^{(j)} = \left(\prod_{i=1}^{j-1} Z_{t-i}\right) Z_{t-j}^2, \qquad j \ge 1.$$

Corollaries 2.3 and 2.4 of Davis and Resnick (1996) show that

$$\lim_{x\to\infty} \frac{P(\sum_{j=0}^{m} c^j X_t^{(j)} > x)}{P(Z_1^2 > x)} = \sum_{j=1}^{m} c^{j\alpha/2} \left(E Z_1^{\alpha/2}\right)^{j-1} \tag{3.8}$$

and

$$\begin{aligned} \lim_{x\to\infty} \frac{P(\sum_{j=0}^{\infty} c^j X_t^{(j)} > x)}{P(Z_1^2 > x)} &= \sum_{j=1}^{\infty} c^{j\alpha/2} \left(E Z_1^{\alpha/2}\right)^{j-1} \\ &= \frac{c^{\alpha/2}}{1 - c^{\alpha/2} E Z_1^{\alpha/2}}. \end{aligned} \tag{3.9}$$

3.3. *Solutions of stochastic difference equations.* Let $\{Y_t, -\infty < t < \infty\}$ be a process which satisfies the stochastic difference equation

$$Y_t = A_t Y_{t-1} + Z_t, \qquad -\infty < t < \infty, \tag{3.10}$$

where $\{(A_n, Z_n), -\infty < n < \infty\}$ are iid $\mathbb{R}_+^2$-valued random pairs [cf. Vervaat (1979) and Grincevicius (1975)]. For the case which we consider here, Z_1 will have regularly varying tail probabilities and the tail of Z_1 is heavier than that of A_1. We assume the pair (A_0, Z_0) satisfies

$$EA_0^\alpha < 1, \qquad EA_0^\beta < \infty \tag{3.11}$$

for some $0 < \alpha < \beta$ and, as usual,

$$P(Z_0 > x) = x^{-\alpha} L(x), \tag{3.12}$$

where L is a slowly varying function at ∞. By iterating (3.10) we find, for $t \geq 1$,

$$Y_t = \sum_{j=0}^{\infty} \left(\prod_{i=t-j+1}^{t} A_i \right) Z_{t-j} := \sum_{j=0}^{\infty} Y_t^{(j)} \tag{3.13}$$

(where $\prod_{i=t+1}^{t} A_i = 1$). It is suggestive to write also

$$Y_t = \sum_{t=0}^{\infty} C_{t,j} Z_{t-j}, \qquad t \geq 1,$$

where $C_{t,j} = \prod_{i=t-j+1}^{t} A_i$ so that the process is a random coefficient MA(∞) process. Furthermore [Resnick and Willekens (1990), Theorem 2.1, and Grincevicius (1975)]

$$\lim_{x\to\infty} \frac{P\left(\sum_{j=0}^{m} Y_t^{(j)} > x\right)}{P(Z_0 > x)} = \sum_{j=1}^{m} (EA_0^\alpha)^{j-1} \tag{3.14}$$

and

$$\lim_{x\to\infty} \frac{P\left(\sum_{j=0}^{\infty} Y_t^{(j)} > x\right)}{P(Z_0 > x)} = \sum_{j=1}^{\infty} (EA_0^\alpha)^{j-1} = \frac{1}{1 - EA_0^\alpha}. \tag{3.15}$$

We now state the result which applies Proposition 2.2 and yields weak consistency of Hill's estimator for these three processes.

COROLLARY 3.1. *Suppose* $\{Z_t\}$ *are iid positive random variables satisfying* (3.1). *The Hill estimator is consistent for* α^{-1} *when applied to either the* MA(∞) *process of Section* 3.1 *or the solution of the random coefficient difference equation described in Section* 3.3. *For the simple bilinear process described in Section* 3.2, *the Hill estimator is consistent for* $2/\alpha$.

PROOF. We apply Proposition 2.2. The key in each case is that each process can be approximated by an m-dependent sequence.

To prove the assertion for the simple bilinear process, let $k \to \infty$, $n/k \to \infty$ and define b_n such that

$$\frac{n}{k} P(X_1 > b_n) \to 1, \qquad n \to \infty. \tag{3.16}$$

For $m \geq 1$, let $X_{n,i}^{(m)} := \sum_{j=0}^{m} c^j X_i^{(j)} / b_n$. Define $X_{n,i} := \sum_{j=0}^{\infty} c^j X_i^{(j)} / b_n$. Since by (3.8) and (3.9) we have, for $x > 0$,

$$\frac{n}{k} P\left(\frac{\sum_{j=0}^{m} c^j X_1^{(j)}}{b_n} > x\right) \to \frac{\sum_{j=1}^{m} c^{j\alpha/2} \left(EZ_1^{\alpha/2}\right)^{j-1}}{\sum_{j=1}^{\infty} c^{j\alpha/2} \left(EZ_1^{\alpha/2}\right)^{j-1}} x^{-\alpha/2},$$

we may define the measures $\nu^{(m)}$ of Proposition 2.2 by

$$\nu^{(m)}\big((x, \infty]\big) := \frac{\sum_{j=1}^{m} c^{j\alpha/2} \left(EZ_1^{\alpha/2}\right)^{j-1}}{\sum_{j=1}^{\infty} c^{j\alpha/2} \left(EZ_1^{\alpha/2}\right)^{j-1}} x^{-\alpha/2}.$$

Note that $\nu^{(m)} \to_v \nu$, where $\nu((x, \infty]) = x^{-\alpha/2}$. Since

$$\begin{aligned} \lim_{m\to\infty} \lim_{n\to\infty} \frac{n}{k} P\big(|X_{n,1}^{(m)} - X_{n,1}| > \varepsilon\big) &= \lim_{m\to\infty} \lim_{n\to\infty} \frac{n}{k} P\left(\frac{\sum_{j=m+1}^{\infty} c^j X_1^{(j)}}{b_n > \varepsilon}\right) \\ &= \lim_{m\to\infty} \frac{\sum_{j=m+1}^{\infty} c^{j\alpha/2} \left(EZ_1^{\alpha/2}\right)^{j-1}}{\sum_{j=1}^{\infty} c^{j\alpha/2} \left(EZ_1^{\alpha/2}\right)^{j-1}} \varepsilon^{-\alpha/2} = 0, \end{aligned}$$

condition (2.14) of Proposition 2.2 is also verified which proves consistency.

The proofs of the results for the MA(∞) process and the solution of the stochastic difference equation are similar. □

We simulated the bilinear process to get a sample of size 5000 using Pareto-distributed Z's satisfying

$$P[Z_1 > x] = x^{-1}, \qquad x > 1.$$

In Figure 1 we give a plot, called the Hill plot, of $\{(k, H_{k,n}^{-1}), 1 \leq k \leq 5000\}$. The graph hovers between 0.5 and 0.6. The correct answer is 0.5.

4. Tail estimation for solutions of stochastic difference equations and ARCH processes. In this section we consider tail estimation for the process $\{Y_t, -\infty < t < \infty\}$ which satisfies the stochastic difference equation

$$Y_t = A_t Y_{t-1} + B_t, \qquad -\infty < t < \infty, \tag{4.1}$$

where $\{(A_n, B_n), -\infty < n < \infty\}$ are iid $\mathbb{R}_+^2$-valued random pairs. In contrast to Section 3, we will now make different assumptions on the tail behavior of the pair (A_n, B_n) which preclude truncating the series solution of (4.1). Solutions to (4.1) include as a particular case the first-order autoregressive

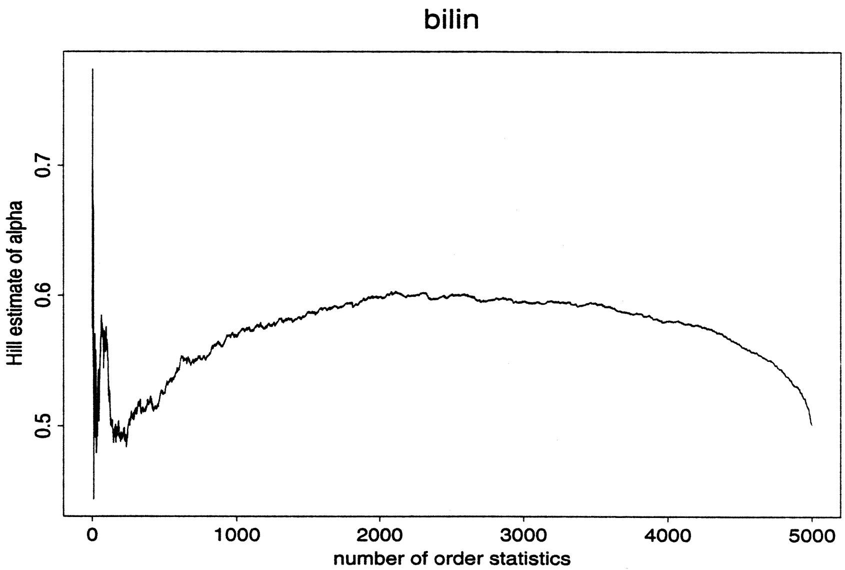

FIG. 1.

conditional heteroscedastic (ARCH) process introduced by Engle (1982). The first-order ARCH process is defined by

$$\xi_t = X_t\big(\beta + \lambda \xi_{t-1}^2\big)^{1/2}, \qquad -\infty < t < \infty, \tag{4.2}$$

where $\{X_t\}$ are iid $N(0,1)$ random variables, $\beta > 0$, $0 < \lambda < 1$. Thus $\{\xi_t^2\}$ satisfies (4.1) with $A_t = \lambda X_t^2$, $B_t = \beta X_t^2$. [Higher-order ARCH processes would satisfy higher-order versions of (4.1) but these are not considered here.]

It is known [Kesten (1973), Vervaat (1979) and Goldie (1991)] that if there exists $\alpha > 0$ with

$$EA_0^\alpha = 1, \qquad EA_0^\alpha \log^+ A_0 < \infty, \qquad 0 < EB_0^\alpha < \infty, \tag{4.3}$$

if $B_0/(1 - A_0)$ is nondegenerate and if the conditional distribution of $\log A_0$ given $A_0 \neq 0$ is nonlattice, then there exists a constant $c > 0$ such that, as $x \to \infty$,

$$P(Y_t > x) \sim cx^{-\alpha}. \tag{4.4}$$

Furthermore [cf. de Haan, Resnick, Rootzen and de Vries (1989), page 220] under the assumptions (4.3) there exist a γ and c_0 such that $0 < \gamma < \alpha$, $0 < c_0 < 1$

$$EA_0^\gamma = c_0 < 1. \tag{4.5}$$

By iterating (4.1) we find, for $t \geq 1$,

$$Y_t = \sum_{j=0}^{\infty}\left(\prod_{i=t-j+1}^{t} A_i\right)B_{t-j} := \sum_{j=0}^{\infty} Y_t^{(j)} \tag{4.6}$$

(where $\prod_{i=t+1}^{t} A_i = 1$). If we iterate (4.1) $t-s$ times for $s<t$, we get

$$Y_t = \sum_{j=0}^{t-s-1} Y_t^{(j)} + \left(\prod_{k=s+1}^{t} A_k\right) Y_s := Y_t^{s,t} + \Pi_{s+1}^t Y_s, \tag{4.7}$$

where

$$\Pi_{s+1}^t = A_t A_{t-1} \cdots A_{s+1} \tag{4.8}$$

and

$$Y_t^{s,t} = B_t + A_t B_{t-1} + A_t A_{t-1} B_{t-2} + \cdots + A_t A_{t-1} \cdots A_{s+2} B_{s+1}. \tag{4.9}$$

Observe that $Y_t^{s,t}$ and Y_s are independent random variables as are Π_{s+1}^t and Y_s.

We begin with a lemma designed to help us check conditions (2.1) and (2.4) for the solution of the stochastic difference equation (4.1).

LEMMA 4.1. *Assume* (4.3) *holds and that* $\varepsilon > 0$ *is given. Suppose* $i_1 < i_2 < \cdots < i_s$ *and* $x_i > 0$ *for* $i = 1, \ldots, s$. *Recall the definition of* γ *and* c_0 *from* (4.5).

(a) *We have that*

$$\begin{aligned} &\left| P(Y_{i_1} > x_1, \ldots, Y_{i_s} > x_s) - \prod_{l=1}^{s} P(Y_{i_l} > x_l) \right| \\ &\quad \le \sum_{q=1}^{s-1} \left(\prod_{j=1}^{s-q} P(Y_0 > x_j) P(Y_0 \in (x_{s-q+1} - \varepsilon, x_{s-q+1} + \varepsilon]) \right. \\ &\qquad\qquad \left. \times \prod_{j=s-q+2}^{s} P(Y_0 > x_j - \varepsilon) \right) \\ &\qquad + \sum_{j=2}^{s} P(\Pi_1^{i_j - i_{j-1}} Y_0 > \varepsilon). \end{aligned} \tag{4.10}$$

(b) *There exists* $M = M(x_1, x_2, \ldots, x_s)$ *and* $K = K(x_1, \ldots, x_s)$ *such that, for* n *large enough*,

$$\begin{aligned} &\left| P\left(Y_{i_1} > \left(\frac{n}{k}\right)^{1/\alpha} x_1, \ldots, Y_{i_s} > \left(\frac{n}{k}\right)^{1/\alpha} x_s \right) \right. \\ &\quad \left. - P\left(Y_{i_1} > \left(\frac{n}{k}\right)^{1/\alpha} x_1 \right) \cdots P\left(Y_{i_s} > \left(\frac{n}{k}\right)^{1/\alpha} x_s \right) \right| \\ &\qquad \le K\varepsilon(s-1) M^{s-1} \left(\frac{k}{n}\right)^{s} + \varepsilon^{-\gamma} E Y_0^{\gamma} \left(\frac{k}{n}\right)^{\gamma/\alpha} \sum_{j=2}^{s} c_0^{i_j - i_{j-1}}. \end{aligned} \tag{4.11}$$

(c) *There exists $C < \infty$ such that*

$$P(Y_1 > (n/k)^{1/\alpha}x, Y_t > (n/k)^{1/\alpha}y) \tag{4.12}$$
$$\leq P(Y_0 > (n/k)^{1/\alpha}x)P(Y_0 > (n/k)^{1/\alpha}(y-\varepsilon)) + C(k/n)c_0^{t-1}.$$

PROOF. The conclusion of (a) follows from an induction argument. To keep the notation simple, we prove (a) for $s = 2$ and then derive the result for $s = 3$. The basic ingredient of the proof is the observation in (4.7). We have, for $s < t$, $x > 0,\ y > 0$,

$$\begin{aligned}
P(Y_s > x, Y_t > y) &\\
&= P(Y_s > x, Y_t^{s,t} + \Pi_{s+1}^t Y_s > y)\\
&\leq P(Y_s > x, Y_t^{s,t} + \Pi_{s+1}^t Y_s > y, \Pi_{s+1}^t Y_s \leq \varepsilon)\\
&\quad + P(Y_s > x, \Pi_{s+1}^t Y_s > \varepsilon)\\
&\leq P(Y_s > x)P(Y_t^{s,t} > y - \varepsilon) + P(Y_s > x, \Pi_{s+1}^t Y_s > \varepsilon) \qquad (4.13)\\
&\leq P(Y_s > x)P(Y_t > y - \varepsilon) + P(Y_s > x, \Pi_{s+1}^t Y_s > \varepsilon) \qquad (4.14)\\
&\leq P(Y_s > x)P(Y_t > y) + P(Y_0 > x)P(y - \varepsilon < Y_0 \leq y)\\
&\quad + P(\Pi_1^{t-s} Y_0 > \varepsilon).
\end{aligned}$$

This shows that

$$\begin{aligned}
&P(Y_s > x, Y_t > y) - P(Y_s > x)P(Y_t > y)\\
&\qquad \leq P(Y_0 > x)P(Y_0 \in (y-\varepsilon, y]) + P(\Pi_1^{t-s} Y_0 > \varepsilon).
\end{aligned}$$

From (4.14) we also get

$$P(Y_s > x, Y_t > y) \leq P(Y_0 > x)P(Y_0 > y - \varepsilon) + P(Y_0 > x, \Pi_1^{t-s} Y_0 > \varepsilon). \tag{4.15}$$

We will use (4.15) in the proof of (c).

The other half of the inequality in (a) is derived as follows:

$$\begin{aligned}
P(Y_s > x)P(Y_t > y + \varepsilon) &\\
&\leq P(Y_s > x)P\big(Y_t^{s,t} + \Pi_{s+1}^t Y_s > y + \varepsilon, \Pi_{s+1}^t Y_s \leq \varepsilon\big)\\
&\quad + P\big(\Pi_{s+1}^t Y_s > \varepsilon\big) \qquad (4.16)\\
&\leq P\big(Y_s > x, Y_t^{s,t} > y\big) + P\big(\Pi_{s+1}^t Y_s > \varepsilon\big)\\
&\leq P\big(Y_s > x, Y_t > y\big) + P\big(\Pi_1^{t-s} Y_0 > \varepsilon\big).
\end{aligned}$$

Hence

$$\begin{aligned}
&-P(Y_0 > x)P(y < Y_0 < y + \varepsilon) - P(\Pi_1^{t-s} Y_0 > \varepsilon)\\
&\qquad \leq P(Y_s > x, Y_t > y) - P(Y_s > x)P(Y_t > y)\\
&\qquad \leq P(Y_0 > x)P(y - \varepsilon < Y_0 \leq y) + P\big(\Pi_1^{t-s} Y_0 > \varepsilon\big).
\end{aligned}$$

The conclusion of (a) for $s = 2$ follows. Based on the case $s = 2$, we will now prove the inequality for $s = 3$. For $s < t < u$ and $x > 0$, $y > 0$ and $z > 0$, we have

$$\begin{aligned}
&P(Y_s > x, Y_t > y, Y_u > z)\\
&\quad \le P(Y_s > x, Y_t > y)P\big(Y_u^{t,u} > z - \varepsilon\big) + P\big(\Pi_{t+1}^{u} Y_t > \varepsilon\big)\\
&\quad \le P(Y_s > x)P(Y_t > y)P\big(Y_u^{t,u} > z - \varepsilon\big)\\
&\qquad + P(Y_0 > x)P(y - \varepsilon < Y_0 \le y)P(Y_u^{t,u} > z - \varepsilon)\\
&\qquad + P(\Pi_1^{t-s} Y_0 > \varepsilon) + P(\Pi_1^{u-t} Y_0 > \varepsilon)\\
&\quad \le P(Y_s > x)P(Y_t > y)P(Y_u > z)\\
&\qquad + P(Y_0 > x)P(Y_0 > y)P\big(Y_0 \in (z - \varepsilon, z]\big)\\
&\qquad + P(Y_0 > x)P(y - \varepsilon < Y_0 \le y)P(Y_0 > z - \varepsilon)\\
&\qquad + P(\Pi_1^{t-s} Y_0 > \varepsilon) + P(\Pi_1^{u-t} Y_0 > \varepsilon).
\end{aligned}$$

For the other half of the inequality, use (4.16) and independence to get

$$\begin{aligned}
&P(Y_s > x)P(Y_t > y)P(Y_u > z + \varepsilon)\\
&\quad \le P(Y_s > x)P(Y_t > y)P(Y_u^{t,u} > z) + P\big(\Pi_{t+1}^{u} Y_t > \varepsilon\big)\\
&\quad \le P(Y_s > x, Y_t > y)P(Y_u^{t,u} > z)\\
&\qquad + P(Y_s > x)P\big(Y_t \in (y, y + \varepsilon]\big)P(Y_u^{t,u} > z)\\
&\qquad + P\big(\Pi_1^{t-s} Y_0 > \varepsilon\big) + P\big(\Pi_1^{u-t} Y_0 > \varepsilon\big)\\
&\quad \le P(Y_s > x, Y_t > y, Y_u > z)\\
&\qquad + P(Y_0 > x)P\big(Y_0 \in (y - \varepsilon, y]\big)P(Y_0 > z - \varepsilon\big)\\
&\qquad + P(\Pi_1^{t-s} Y_0 > \varepsilon) + P(\Pi_1^{u-t} Y_t > \varepsilon).
\end{aligned}$$

Therefore

$$\begin{aligned}
&-P(Y_0 > x)P(Y_0 > y)P(Y_0 \in (z, z + \varepsilon])\\
&\quad - P(Y_0 > x)P(Y_0 \in (y - \varepsilon, y])P(Y_0 > z - \varepsilon)\\
&\quad - P(\Pi_1^{t-s} Y_0 > \varepsilon) - P(\Pi_1^{u-t} Y_t > \varepsilon)\\
&\qquad \le P(Y_s > x, Y_t > y, Y_u > z) - P(Y_s > x)P(Y_t > y)P(Y_u > z)\\
&\qquad \le P(Y_0 > x)P(Y_0 > y)P(Y_0 \in (z - \varepsilon, z])\\
&\qquad\quad + P(Y_0 > x)P(Y_0 \in (y - \varepsilon, y])P(Y_0 > z - \varepsilon)\\
&\qquad\quad + P(\Pi_1^{t-s} Y_0 > \varepsilon) + P(\Pi_1^{u-t} Y_0 > \varepsilon).
\end{aligned}$$

To prove (b), we note that since the inequality in (a) holds whenever $x_i > 0$, $i = 1, \dots, s$, and $\varepsilon > 0$ (provided $x_i - \varepsilon > 0$, $i = 1, \dots, s$), we may replace x_i by $(n/k)^{1/\alpha} x_i$ and ε by $(n/k)^{1/\alpha}\varepsilon$ to get a valid inequality. The inequality in (b) then results from the one in (a) by using $c_0 = EA_0^\gamma < 1$, $P[Y_0 > x] \sim cx^{-\alpha}$, $x \to \infty$ and Markov's inequality. To see this, note that the upper bound becomes

$$\sum_{q=1}^{s-1}\Biggl(\prod_{j=1}^{s-q} P\biggl[\frac{Y_0}{(n/k)^{1/\alpha}} > x_j\biggr] P\biggl[\frac{Y_0}{(n/k)^{1/\alpha}} \in \bigl(x_{s-q+1}-\varepsilon, x_{s-q+1}+\varepsilon\bigr]\biggr] \times \prod_{j=s-q+2}^{s} P\biggl[\frac{Y_0}{(n/k)^{1/\alpha}} > x_j - \varepsilon\biggr]\Biggr) + \sum_{j=2}^{s} P\biggl[\Pi_1^{i_j - i_{j-1}} Y_0 > \Bigl(\frac{n}{k}\Bigr)^{1/\alpha}\varepsilon\biggr]. \tag{4.17}$$

Note that by Markov's inequality

$$\sum_{j=2}^{s} P\biggl[\Pi_1^{i_j-i_{j-1}} Y_0 > \Bigl(\frac{n}{k}\Bigr)^{1/\alpha}\varepsilon\biggr] \le \Bigl(\frac{k}{n}\Bigr)^{\gamma/\alpha}\varepsilon^{-\gamma}\sum_{j=2}^{s} E\bigl(\Pi_1^{i_j-i_{j-1}} Y_0\bigr)^\gamma = \Bigl(\frac{k}{n}\Bigr)^{\gamma/\alpha}\varepsilon^{-\gamma}\sum_{j=2}^{s} c_0^{i_j-i_{j-1}} EY_0^\gamma.$$

Furthermore, for n sufficiently large and some constant $K = K(x_1, \dots, x_s)$,

$$\frac{n}{k}P\biggl[Y_0 > \Bigl(\frac{n}{k}\Bigr)^{1/\alpha}(x_j - \varepsilon)\biggr] \le M, \qquad j = 1, \dots, s,$$

$$\frac{n}{k}P\biggl[Y_0 \in \Bigl(\frac{n}{k}\Bigr)^{1/\alpha}(x_{s-q+1}-\varepsilon, x_{s-q+1}+\varepsilon]\biggr] \le \varepsilon, \qquad j = 1, \dots, s,$$

and therefore the first summation in (4.17) is bounded by

$$\sum_{q=1}^{s-1}\prod_{j=1}^{s-q}\frac{k}{n}M\varepsilon K\frac{k}{n}\prod_{j=s-q+2}^{s}\frac{k}{n}M = \sum_{q=1}^{s-1}\Bigl(\frac{k}{n}\Bigr)^{s-q}M^{s-q}\varepsilon K\frac{k}{n}\Bigl(\frac{k}{n}\Bigr)^{q-1}M^{q-1} = K\varepsilon(s-1)M^{s-1}\Bigl(\frac{k}{n}\Bigr)^{s},$$

which verifies (4.11).

To prove (c), substitute in (4.15) $s = 1$ and replace ε, x and y by $(n/k)^{1/\alpha}\varepsilon$, $(n/k)^{1/\alpha}x$ and $(n/k)^{1/\alpha}y$. The desired result is shown if we prove

$$P\biggl[Y_0 > \Bigl(\frac{n}{k}\Bigr)^{1/\alpha}x, \Pi_1^{t-1}Y_0 > \Bigl(\frac{n}{k}\Bigr)^{1/\alpha}\varepsilon\biggr] \le c\frac{k}{n}c_0^{t-1}. \tag{4.18}$$

The probability on the left-hand side of 4.18 is

$$\int_x^\infty P\bigl[\Pi_1^{t-1} > \varepsilon u^{-1}\bigr]P\biggl[\frac{Y_0}{(n/k)^{1/\alpha}} \in du\biggr] \le c_0^{t-1}\varepsilon^{-\gamma}\int_x^\infty u^\gamma P\biggl[\frac{Y_0}{(n/k)^{1/\alpha}} \in du\biggr],$$

and since

$$\frac{n}{k}P\left[Y_0 > \left(\frac{n}{k}\right)^{1/\alpha} u\right] \to u^{-\alpha},$$

we get from Karamata's theorem that

$$\int_x^\infty u^\gamma \frac{n}{k} P\left[\left(\frac{n}{k}\right)^{-1/\alpha} Y_0 \in du\right] \to \int_x^\infty u^\gamma \alpha u^{-\alpha-1}\, du < \infty$$

and thus, for all large n, the probability on the left-hand side of 4.18 is bounded by $Cc_0^{t-1}(k/n)$, as was to be proven. □

LEMMA 4.2. *Assume* (4.3) *holds and let* $\{Y_t\}$ *be the solution of* (4.1). *Then condition* (2.1) *or* (2.8) *holds for the array* $\{Y_t/(n/k)^{1/\alpha}\}$; *that is,*

$$\lim_{n\to\infty} \frac{n}{k^2} \sum_{j=2}^{k} P\left(Y_1 > \left(\frac{n}{k}\right)^{1/\alpha} x,\, Y_j > \left(\frac{n}{k}\right)^{1/\alpha} y\right) = 0 \tag{4.19}$$

for any $x > 0$, $y > 0$. *If, in addition, one chooses* l_n *such that* $l_n/k \to 0$ *and*

$$\frac{n}{k} = o(l_n), \tag{4.20}$$

then condition (2.18) *also holds; that is,*

$$\begin{aligned}\lim_{n\to\infty}\Bigg| E \prod_{j=1}^{p} \left(1 - \frac{1}{k}\sum_{i\in I_j} f\left(\frac{Y_i}{(n/k)^{1/\alpha}}\right)\right) \\ - \prod_{j=1}^{p} E\left(1 - \frac{1}{k}\sum_{i\in I_j} f\left(\frac{Y_i}{(n/k)^{1/\alpha}}\right)\right)\Bigg| = 0,\end{aligned} \tag{4.21}$$

where $p = [n/k]$, I_j, $j = 1, \dots, p$, *are defined in* (2.3) *and the function* f *is of the form given in Proposition* 2.3.

PROOF. To check condition (2.1), use (c) of Lemma 4.1:

$$\begin{aligned}&\frac{n}{k^2}\sum_{j=2}^{k} P\left(Y_1 > \left(\frac{n}{k}\right)^{1/\alpha} x,\, Y_{j+1} > \left(\frac{n}{k}\right)^{1/\alpha} y\right) \\ &\quad \le \frac{n(k-1)}{k^2} P\left(Y_0 > \left(\frac{n}{k}\right)^{1/\alpha} x\right) P\left(Y_0 > \left(\frac{n}{k}\right)^{1/\alpha} (y-\varepsilon)\right) \\ &\qquad + C\frac{1}{k}\sum_{j=2}^{k} c_0^{j-1} \to 0\end{aligned}$$

as $n \to \infty$.

To prove condition (4.21) holds, note that

$$\left| E \prod_{j=1}^{p} \left(1 - \frac{1}{k} \sum_{i \in I_j} f\left(\frac{Y_i}{(n/k)^{1/\alpha}}\right)\right) - \prod_{j=1}^{p} E\left(1 - \frac{1}{k} \sum_{i \in I_j} f\left(\frac{Y_i}{(n/k)^{1/\alpha}}\right)\right)\right|$$

$$\le \sum_{u=2}^{p} \frac{1}{k^u} \sum_{1 \le j_1 < j_2 < \cdots < j_u \le p} \sum_{i_1 \in I_{j_1}} \sum_{i_2 \in I_{j_2}} \cdots \sum_{i_u \in I_{j_u}} \left| E\left(\prod_{v=1}^{u} f\left(\frac{Y_{i_v}}{(n/k)^{1/\alpha}}\right)\right) - \prod_{v=1}^{u} Ef\left(\frac{Y_{i_v}}{(n/k)^{1/\alpha}}\right)\right|. \tag{4.22}$$

Also due to the definition of f one has

$$E\left(\prod_{v=1}^{u} f\left(\frac{Y_{i_v}}{(n/k)^{1/\alpha}}\right)\right) - \prod_{v=1}^{u} Ef\left(\frac{Y_{i_v}}{(n/k)^{1/\alpha}}\right)$$

$$= \sum_{h_1=1}^{s} \cdots \sum_{h_u=1}^{s} \beta_{h_1} \cdots \beta_{h_u} \left| P\left(\frac{Y_{i_1}}{(n/k)^{1/\alpha}} > x_{h_1}, \ldots, \frac{Y_{i_u}}{(n/k)^{1/\alpha}} > x_{h_u}\right) - P\left(\frac{Y_{i_1}}{(n/k)^{1/\alpha}} > x_{h_1}\right) \cdots P\left(\frac{Y_{i_u}}{(n/k)^{1/\alpha}} > x_{h_u}\right)\right|. \tag{4.23}$$

From (4.11) it follows that

$$\left| P\left(\frac{Y_{i_1}}{(n/k)^{1/\alpha}} > x_{h_1}, \ldots, \frac{Y_{i_u}}{(n/k)^{1/\alpha}} > x_{h_u}\right) - \prod_{j=1}^{u} P\left(\frac{Y_{i_j}}{(n/k)^{1/\alpha}} > x_{h_j}\right)\right|$$

$$\le K\varepsilon(u-1)M^u \left(\frac{k}{n}\right)^u + \varepsilon^{-\gamma} EY_0^{\gamma} \left(\frac{k}{n}\right)^{\gamma/\alpha} (u-1)c_0^{l_n}.$$

If we denote $Q := \max\{\beta_h \colon h = 1, \ldots, s\}$, then one can bound (4.22) by

$$\sum_{u=2}^{p} \frac{1}{k^u} \sum_{1 \le j_1 < j_2 < \cdots < j_u \le p} \sum_{i_1 \in I_{j_1}} \sum_{i_2 \in I_{j_2}} \cdots \sum_{i_u \in I_{j_u}} (sQ)^u \left(\varepsilon(u-1)\left(M\frac{k}{n}\right)^u + \varepsilon^{-\gamma} EY_0^{\gamma}\left(\frac{k}{n}\right)^{\gamma/\alpha}(u-1)c_0^{l_n}\right)$$

$$\le \sum_{u=2}^{p} \frac{1}{k^u} \binom{p}{u} (k - l_n)^u (sQ)^u \times \left(\varepsilon K(u-1)M^u \left(\frac{k}{n}\right)^u + \varepsilon^{-\gamma} EY_0^{\gamma} \left(\frac{k}{n}\right)^{\gamma/\alpha} (u-1)c_0^{l_n}\right)$$

$$\le K\varepsilon \sum_{u=2}^{p} \binom{p}{u} (u-1)(sQM)^u (k/n)^u + \varepsilon^{-\gamma} EY_0^{\gamma} \left(\frac{k}{n}\right)^{\gamma/\alpha} c_0^{l_n} \sum_{u=2}^{p} \binom{p}{u} (u-1)(sQ)^u$$

$$= K\varepsilon\bigg(\frac{sQMpk}{n}\bigg(1+\frac{sQMk}{n}\bigg)^{p-1} - \bigg(1+\frac{sQMk}{n}\bigg)^{p} + 1\bigg)$$
$$+ \varepsilon^{-\gamma} EY_0^{\gamma}\bigg(\frac{k}{n}\bigg)^{\gamma/\alpha} c_0^{l_n}\big(psQ\{(1+sQ)^{p-1}-1\} - [(1+sQ)^p - 1 - psQ]\big)$$
$$= A + B.$$

When $n \to \infty$, B goes to 0, due to (4.20), and $A \to \varepsilon((sQM-1)\exp(sQM)+1)$. Letting $\varepsilon \to 0$ ends the proof for condition (4.21). □

PROPOSITION 4.1. *Assume* (4.3) *holds and let* $\{Y_t\}$ *be the solution of* (3.10). *Choose* $k(n)$ *such that there exists* $\{l(n)\}$ *satisfying* 2.2 *and* 4.20; *that is*,

$$\frac{n}{k} = o(l(n)) \qquad \textit{and} \qquad l(n) = o(k).$$

Then the Hill estimator applied to the sequence Y_t *is consistent*; *that is*,

$$\frac{1}{k}\sum_{i=1}^{k}\log\frac{Y_{(i)}}{Y_{(k+1)}} \to_P \frac{1}{\alpha}. \tag{4.24}$$

REMARK. Possible choices of $\{k(n)\}$ include $n = o(k^{3/2})$ and $k = n^{\beta}$ for $0.5 < \beta < 1$.

PROOF. The choice of k makes sure that (2.2) and (4.20) hold. The conclusion then follows from Proposition 2.1. □

For the ARCH process $\{\xi_t\}$ given by (4.2), we have

$$P[\xi_t^2 > x] \sim cx^{-\alpha}, \qquad x \to \infty,$$

where α satisfies

$$E(\lambda X_t^2)^{\alpha} = 1,$$

with $\{X_t\}$ being iid $N(0,1)$ random variables. Equivalently, α satisfies

$$\Gamma\big(\alpha + \tfrac{1}{2}\big) = \sqrt{\pi}(2\lambda)^{-\alpha}.$$

Thus the Hill estimator applied to $\{\xi_1^2, \ldots, \xi_t^2\}$ is consistent for α^{-1} and a consistent estimator for λ is obtained from solving

$$\Gamma\big(\hat{\alpha} + \tfrac{1}{2}\big) = \sqrt{\pi}(2\hat{\lambda})^{-\hat{\alpha}}$$

for $\hat{\lambda}$, where $\hat{\alpha}$ is the estimate of α given by the reciprocal of the Hill estimator.

We simulated 7000 data from the ARCH(1) model using $\beta = 1$ and $\lambda = 0.5$. In this case, the true value of α for $\{\xi_t^2\}$ is $\alpha \approx 2.365$. Figure 2 displays the Hill plots which indicate an estimate of α in the neighborhood of 2.1 or 2.2. The AltHill plot in the display is $\{(\theta, H^{-1}_{[n^{\theta}],n}, 0 \le \theta \le 1\}$ and the AltsmooHill plot smooths the AltHill plot. See Resnick and Stărică (1997a) for a discussion of such plots.

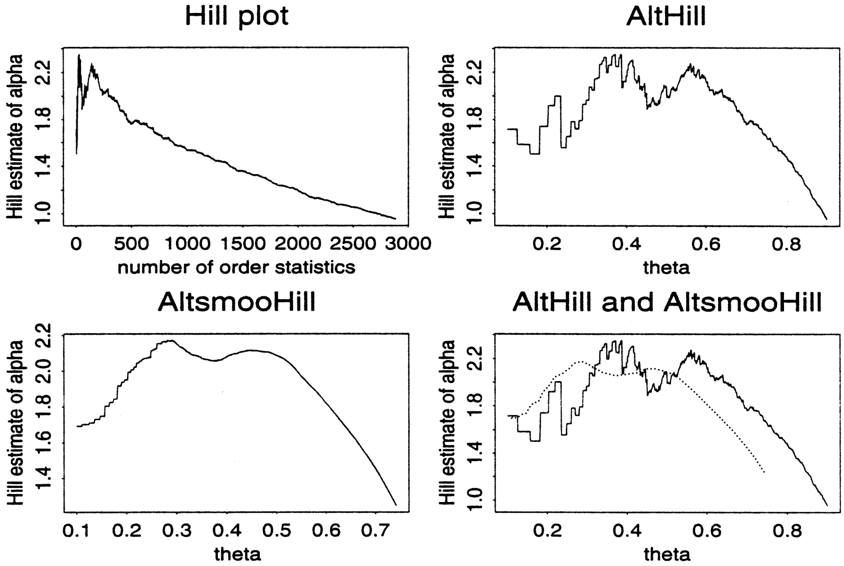

FIG. 2. *Hill plots of* ARCH^2 *when* $\lambda = 0.5$.

5. Hidden Markov models. A heavy-tailed hidden Markov model is proposed in Meier-Hellstern, Wirth, Yan and Hoeflin (1991) to model the times between transmission of packets at a source. We show the Hill estimator is consistent when applied to such models.

The model has the following ingredients. Let $\{J_n, n \geq 0\}$ be an ergodic, m-state Markov chain on the state space $\{1, 2, \ldots, m\}$. Suppose the transition probability matrix of this chain is $P = \{p_{ij}, 1 \leq i, j \leq m\}$ and that the stationary distribution is $\pi' = (\pi_1, \ldots, \pi_m)$. Now suppose, for $i = 1, \ldots, m$, we are given holding time distributions $\{q_n^{(i)}, n \geq 1\}$ concentrating on $\{1, 2, \ldots\}$ and that, for $i = 1, \ldots, m$, $\{D_n^{(i)}, n \geq 0\}$ are iid with common distribution $\{q_n^{(i)}\}$. Define $\{V_n, n \geq 0\}$ by

$$V_j = \begin{cases} J_0, & \text{if } 0 \leq j < D_0^{(J_0)}, \\ J_1, & \text{if } D_0^{(J_0)} \leq j < D_0^{(J_0)} + D_1^{(J_1)}, \\ J_2, & \text{if } D_0^{(J_0)} + D_1^{(J_1)} \leq j < D_0^{(J_0)} + D_1^{(J_1)} + D_2^{(J_2)}, \\ \vdots & \vdots \end{cases}$$

Thus

$$V_j = \sum_{k=0}^{\infty} J_k 1_{\left[\sum_{l=0}^{k-1} D_l^{(J_l)} \leq j < \sum_{l=0}^{k} D_l^{(J_l)}\right]}.$$

The next ingredient we need are distributions $F_1, \dots, F_m$ on $\mathbb{R}_+$ and iid uniform random variables with support $[0, 1]$ which we call $\{U_n, n \geq 0\}$. Define, for $n \geq 0$,

$$X_n = F_{V_n}^{\leftarrow}(U_n) \tag{5.1}$$

and assume $\{U_n\}$, $\{J_n\}$, $\{D_n^{(i)}, n \geq 0, 1 \leq i \leq m\}$ are all independent.

So changes of state follow the Markov chain $\{J_n\}$ and a transition from i to j occurs with probability p_{ij}. Having entered state i, the system stays in state i for k time units with probability $q_k^{(i)}$. While in state i, random variables which we think of as interarrivals are generated from distribution F_i.

PROPOSITION 5.1. *Suppose $\{J_n\}$ is a stationary, ergodic Markov chain and that*

$$ED_n^{(i)} < \infty, \qquad i = 1, \dots, m. \tag{5.2}$$

Suppose, for $\alpha > 0$,

$$\bar{F}_1(x) \sim x^{-\alpha} L(x), \qquad x \to \infty, \tag{5.3}$$

and

$$\lim_{x \to \infty} \frac{\bar{F}_j(x)}{\bar{F}_1(x)} = 0, \qquad j = 2, \dots, m. \tag{5.4}$$

Define the quantile function

$$b(t) = \left(\frac{1}{1 - F_1}\right)^{\leftarrow}(t).$$

If $k \to \infty$, $n/k \to \infty$, then

$$\frac{1}{k} \sum_{i=1}^{n} \varepsilon_{X_i / b(n/k)} \Rightarrow \nu,$$

where

$$\nu((x, \infty]) = \theta_1 x^{-\alpha}$$

and, for $k = 1, \dots, m$,

$$\theta_k = \frac{ED_1^{(k)} \pi_k}{E \sum_{j=1}^{m} D_1^{(j)} \pi_j}.$$

Furthermore, the Hill estimator applied to $\{X_t\}$ is consistent for α^{-1}.

PROOF. The proof uses Laplace functionals. For $f \in C_K^+(\mathbb{E})$, we need to show

$$\Psi_n(f) := E \exp\left\{-\frac{1}{k} \sum_{j=1}^{n} f\left(\frac{X_j}{b(n/k)}\right)\right\} \to \exp(-\nu(f)) = \exp\left\{-\int_{\mathbb{E}} f(x) \nu(dx)\right\}.$$

Define, for $n \geq 0$,

$$N_n^{(j)} = \sum_{l=0}^{n} 1_{[V_l = j]}, \qquad j = 1, \dots, m,$$

$$\mu^{(j)}(n) = \sum_{l=0}^{n} 1_{[J_l = j]}, \qquad j = 1, \dots, m.$$

Because $\{X_n\}$ is conditionally independent given $\{V_n\}$ [see (5.1)], we have

$$\begin{aligned} \Psi_n(f) &= E\left(E\left(\exp\left\{-\frac{1}{k}\sum_{j=1}^{n} f\left(\frac{X_j}{b(n/k)}\right)\right\}\middle| V_0, \dots, V_n\right)\right) \\ &= E\prod_{j=1}^{m}\left(\int_{\mathbb{E}} \exp\left\{\frac{-f(x)}{k}\right\} F_j\left(b\left(\frac{n}{k}\right) dx\right)\right)^{N_n^{(j)}}. \end{aligned} \tag{5.5}$$

We now study the behavior of $N_n^{(j)}$ and we will prove that, as $n \to \infty$,

$$\frac{N_n^{(n)}}{n} \to_P \theta_j, \qquad j = 1, \dots, m. \tag{5.6}$$

The semi-Markov process $\{V_j\}$ changes states at times $\{S_n\}$ where

$$S_n = \sum_{q=0}^{n} D_q^{(J_q)}$$

and, as $n \to \infty$, we have

$$\begin{aligned} \frac{S_n}{n} &\stackrel{d}{=} \frac{1}{n}\sum_{k=1}^{m}\sum_{i=1}^{\mu^{(k)}(n)} D_i^{(k)} \\ &= \sum_{k=1}^{m} \frac{\sum_{i=1}^{\mu^{(k)}(n)} D_i^{(k)}}{\mu^{(k)}(n)} \frac{\mu^{(k)}(n)}{n} \\ &\to \sum_{k=1}^{m} ED_1^{(k)} \pi_k. \end{aligned}$$

Now we define the process inverse to $\{S_n\}$ as

$$M(t) = \sup\{n\colon S_n \leq t\}$$

so that, as $t \to \infty$,

$$\frac{M(t)}{t} \to \frac{1}{\sum_{k=1}^{m} ED_1^{(k)} \pi_k}.$$

The relevance of $\{S_n\}$ and $\{M(t)\}$ is that

$$\begin{aligned}
\frac{N_n^{(k)}}{n} &\le \frac{1}{n}\sum_{q=1}^{M(n)+1}(S_q - S_{q-1})1_{[J_{q-1}=k]} \\
&= \frac{1}{n}\sum_{q=1}^{M(n)+1} D_q^{(k)} 1_{[J_{q-1}=k]} \\
&\overset{d}{=} \frac{1}{n}\sum_{q=0}^{\mu^{(k)}(M(n)+1)} D_q^{(k)} \\
&= \frac{\sum_{q=0}^{\mu^{(k)}(M(n)+1)} D_q^{(k)}}{\mu^{(k)}(M(n)+1)} \frac{\mu^{(k)}(M(n)+1)}{M(n)+1} \frac{M(n)+1}{n} \\
&\to \frac{ED_1^{(k)}\pi_k}{\sum_{j=1}^m ED_1^{(j)}\pi_j}.
\end{aligned}$$

A lower bound is obtained similarly and this proves (5.6).

Note that because of (5.4) we have, for $x > 0$, that

$$\frac{n}{k}\bar{F}_j\left(b\left(\frac{n}{k}\right)x\right) \to 0, \qquad j = 2, \dots, m.$$

Thus, for $2 \le j \le m$,

$$\begin{aligned}
1 &\ge \left(\int_{\mathbb{E}} \exp\left\{\frac{-f(x)}{k}\right\} F_j\left(b\left(\frac{n}{k}\right) dx\right)\right)^{N_n^{(j)}} \\
&= \left(1 - \frac{\int_{\mathbb{E}}(1 - \exp(-f(x)/k) F_j(b(n/k)\, dx)}{n}\right)^{N_n^{(j)}} \\
&\ge \left(1 - \frac{\int_{\mathbb{E}} f(x)(n/k) F_j(b(n/k)\, dx)}{n}\right)^{n(N_n^{(j)})/n} \\
&\to \exp\{-0\} = 1,
\end{aligned}$$

since

$$\int_{\mathbb{E}} f(x)\frac{n}{k} F_j\left(b\left(\frac{n}{k}\right) dx\right) \to 0 \quad \text{if } f \in C_K^+(\mathbb{E}).$$

For $j = 1$ we claim

$$\int_{\mathbb{E}} \left(1 - \exp\left\{\frac{-f(x)}{k}\right\} n F_1\left(b\left(\frac{n}{k}\right) dx\right) \to \nu(f) \tag{5.7}$$

and, assuming this is true, we get

$$\begin{aligned}&\left(\int_{\mathbb{E}}\exp\left\{\frac{-f(x)}{k}\right\}F_1\left(b\left(\frac{n}{k}\right)dx\right)\right)^{N_n^{(1)}}\\ &\quad=\left(1-\frac{\int_{\mathbb{E}}(1-\exp\{-f(x)/k\})nF_1(b(n/k)\,dx)}{n}\right)^{n(N_n^{(1)}/n)}\\ &\quad\to\exp\{-\nu(f)\}.\end{aligned}\tag{5.8}$$

To verify (5.7), observe that

$$\int_{\mathbb{E}}\left(1-\exp\left\{\frac{-f(x)}{k}\right\}\right)nF_1\left(b\left(\frac{n}{k}\right)dx\right)\le\int_{\mathbb{E}}f(x)\frac{n}{k}F_1\left(b\left(\frac{n}{k}\right)dx\right)\to\nu(f)$$

and

$$\begin{aligned}&\int_{\mathbb{E}}\left(1-\exp\left\{\frac{-f(x)}{k}\right\}\right)nF_1\left(b\left(\frac{n}{k}\right)dx\right)\\ &\quad\ge\int_{\mathbb{E}}\left(\frac{f(x)}{k}-\frac{f^2(x)}{k^2}\right)nF_1\left(b\left(\frac{n}{k}\right)dx\right)\\ &\quad=\nu(f)+o(1)-\frac{1}{k}\int_{\mathbb{E}}f^2(x)\frac{n}{k}F_1\left(b\left(\frac{n}{k}\right)dx\right)\\ &\quad=\nu(f)+o(1)+\frac{1}{k}O(1)\\ &\quad\to\nu(f).\end{aligned}$$

This proves (5.7).

Thus the factors in $\Psi_n(f)$ in (5.5) not corresponding to state 1 converge to 1, while the factor from state 1 converges to the correct limit. The desired result follows from dominated convergence after taking expectations. □

REFERENCES

BROCKWELL, P. and DAVIS, R. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York.

CLINE, D. (1983). Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data. Ph.D. dissertation, Dept. Statist., Colorado State Univ.

DAVIS, R. and RESNICK, S. (1984). Tail estimates motivated by extreme value theory. *Ann. Statist.* **12** 1467–1487.

DAVIS, R. and RESNICK, S. (1988). Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stochastic Process. Appl.* **30** 41–68.

DAVIS, R. and RESNICK, S. (1996). Limit theory for bilinear processes with heavy tailed noise. *Ann. Appl. Prob.* **6** 1191–1210.

ENGLE, R. (1982). Autoregressive conditional heteroscedastic models with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007.

FEIGIN, P. and RESNICK, S. (1996). Pitfalls of fitting autoregressive models for heavy-tailed time series. Unpublished manuscript.

GELUK, J., DE HAAN, L., RESNICK, S. and STĂRICĂ, C. (1997). Second order regular variation, convolution and the central limit theorem. *Stochastic Process. Appl.* **69** 139–159.

GOLDIE, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1** 126–166.

GRANGER, C. W. J. and ANDERSEN, A. (1978). Non-linear time series modelling. In *Applied Time Series Analysis* (D. Findley, ed.) 25–38. Academic Press, New York.

GRINCEVICIUS, A. (1975). One limit distribution for a random walk on the line. *Lithuanian Math. J.* **15** 580–589.

DE HAAN, L. and RESNICK, S. I. (1998). On asymptotic normality of the Hill estimator. *Stochastic Models*. To appear.

DE HAAN, L., RESNICK, S. I., ROOTZEN, H. and DE VRIES, C. (1989). Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.* **32** 213–224.

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

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.

HILL, B. (1975). A simple approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163–1174.

HSING, T. (1991). On tail estimation using dependent data. *Ann. Statist.* **19** 1547–1569.

HSING, T., HÜSLER, J. and LEADBETTER, M. R. (1988). On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields* **78** 97–112.

KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Akademie, Berlin.

KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** 207–248.

LEADBETTER, M., LINDGREN, G. and ROOTZEN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

MASON, D. (1982). Laws of large numbers for sums of extreme values. *Ann. Probab.* **10** 754–764.

MASON, D. (1988). A strong invariance theorem for the tail empirical process. *Ann. Inst. H. Poincaré Probab. Statist.* **24** 491–506.

MASON, D. and TUROVA, T. (1994). Weak convergence of the Hill estimator process. In *Extreme Value Theory and Applications* (J. Galambos, J. Lechner and Simiu, eds.) 419–431. Kluwer, Dordrecht.

MEIER–HELLSTERN, K. S., WIRTH, P. E., YAN, Y. L. and HOEFLIN, D. A. (1991). Traffic models for ISDN data users: office automation application. In *Teletraffic and Datatraffic in a Period of Change. Proceedings of the 13th ITC* (A. Jensen and V. B. Iversen, eds.) 167–192. North-Holland, Amsterdam.

MOHLER, R. R. (1973). Bilinear control processes with applications to engineering, ecology, and medicine. *Math. Sci. Engrg.* **106**.

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

RESNICK, S. (1998). Why non-linearities can ruin the heavy tailed modeler's day. In *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy Tailed Distributions* (R. Adler, R. Feldman, M. S. Taqqu, eds.) Birkhäuser, Boston. In press.

RESNICK, S. and STĂRICĂ, C. (1995). Consistency of Hill's estimator for dependent data. *J. Appl. Probab.* **32**, 139–167.

RESNICK, S. and STĂRICĂ, C. (1997a). Smoothing the Hill estimator. *Adv. in Appl. Probab.* **29** 271–293.

RESNICK, S. and STĂRICĂ, C. (1997b). Asymptotic behavior of Hill's estimator for autoregressive data. *Stochastic Models* **13** 703–723.

RESNICK, S. and WILLEKENS, E. (1990). Moving averages with random coefficients and random coefficient autoregressive models. *Stochastic Models* **7** 511–526.

ROOTZEN, H. (1995). The tail empirical process for stationary sequences. Preprint 1995:9 ISSN 1100-2255, Studies in Statistical Quality Control and Reliability, Chalmers Univ. Technology.

ROOTZEN, H., LEADBETTER, M. and DE HAAN, L. (1990). Tail and quantile estimation for strongly mixing stationary sequences. Technical Report 292, Center for Stochastic Processes, Dept. Statistics, Univ. North Carolina, Chapel Hill, NC.

VERVAAT, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. in Appl. Probab.* **11** 750–783.

SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853
E-MAIL: sid@orie.cornell.edu

DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA 19104-6302
E-MAIL: starica@compstat.wharton.upenn.edu