

# Estimation for Nonnegative First-Order Autoregressive Processes with an Unknown Location Parameter

Andrew Bartlett, William McCormick

University of Georgia, Athens, USA  
 Email: wpmcc@uga.edu

Received September 10, 2012; revised October 10, 2012; accepted October 17, 2012

## ABSTRACT

Consider a first-order autoregressive processes  $X_t = \phi X_{t-1} + Z_t$ , where the innovations are nonnegative random variables with regular variation at both the right endpoint infinity and the unknown left endpoint  $\theta$ . We propose estimates for the autocorrelation parameter  $\phi$  and the unknown location parameter  $\theta$  by taking the ratio of two sample values chosen with respect to an extreme value criteria for  $\phi$  and by taking the minimum of  $X_t - \hat{\phi}_n X_{t-1}$  over the observed series, where  $\hat{\phi}_n$  represents our estimate for  $\phi$ . The joint limit distribution of the proposed estimators is derived using point process techniques. A simulation study is provided to examine the small sample size behavior of these estimates.

**Keywords:** Nonnegative Time Series; Autoregressive Processes; Extreme Value Estimator; Regular Variation; Point Processes

## 1. Introduction

In many applications, the desire to model the phenomena under study by non-negative dependent processes has increased. An excellent presentation of the classical theory concerning these models can be found, for example, in Brockwell and Davis [1]. Recently, advancements in such models have shifted focus to some specialized features of the model, e.g. heavy tail innovations or non-negativity of the model. In this paper we examine the behavior of traditional estimates under conditions leading to non-Gaussian limits. For example, the standard approach to parameter estimation within the AR(1) process is through the Yule-Walker estimator;

$$\hat{\phi}_{LS} = \frac{\sum_{t=1}^{n-1} (X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{t=1}^n (X_{t+1} - \bar{X})^2}, \quad (1.1)$$

where  $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$ .

A slightly different approach presented in Mathew and McCormick [2] used linear programming to obtain estimates for  $\phi$  and  $\theta$  under certain optimization constraints. While there are many established methods to estimate the autocorrelation coefficient in an AR(1) model, there are just a few approaches on estimating the unknown location parameter in an AR(1) model. One,

was mentioned in McCormick and Mathew [2] where they considered

$$\hat{\theta}_{\text{range}} = X_{j^*} - \hat{\phi}_{\text{range}} X_{i^*-1}, \quad \text{where } \hat{\phi}_{\text{range}} = \frac{X_{i^*+1} - X_{j^*}}{X_{i^*} - X_{j^*-1}},$$

$i^*$  and  $j^*$  provides the index of the maximal and minimal  $X_i$  respectively for  $1 \leq i \leq n$ .

In this paper we examine estimation questions and asymptotic properties of alternative estimates for  $\phi$  and  $\theta$  respectively, relating to the model

$$X_t = \phi X_{t-1} + Z_t, t \geq 1,$$

where  $0 < \phi < 1, \theta > 0$  and  $\{Z_t\}$  is an i.i.d. sequence of nonnegative random variables whose innovation distribution  $F$  is assumed to be regularly varying at infinity with index  $-\beta$  and regularly varying at  $\theta$  with index  $\alpha$ , where  $\theta$ , denotes the unknown but positive left endpoint. As a result of not restricting the innovations  $\{Z_t\}$  to be bounded on a finite range, we can first estimate the autoregressive parameter  $\phi$  through regular variation at infinity and then estimate the positive but unknown location parameter through regular variation at  $\theta$ , the left endpoint.

While we have mentioned a few established estimation procedures, one notable exception was that of maximum likelihood. Although typically intractable and intricate in the time series setting, when the innovations in the AR(1) model are exponential, the maximum likelihood procedure

ture had a major contribution on the estimation of positive heavy tailed time series. With these considerations in mind, Raftery [3] determined the limiting distribution of the maximum likelihood estimate for the autocorrelation coefficient  $\phi$ . As a result, the estimator

$$\hat{\phi}_n = \bigwedge_{t=1}^n \frac{X_t}{X_{t-1}}, \tag{1.2}$$

was considered. The realization of this estimator was the stepping stone for the work done in this paper along with Davis and McCormick [4] which first considered this alternative estimator and used a point process approach to obtain the asymptotic distribution of the natural estimator  $\hat{\phi}_n$ . This was done in the context that the innovations distribution  $F$  varies regularly at 0, the left endpoint, and satisfy some moment condition.

The work presented in this paper is an extension of the work done in Davis and McCormick [4] including the following contributions to dependent time series with heavy-tail innovations. The first contribution involves the development of estimates for the autocorrelation coefficient and unknown location parameter under regular variation at both endpoints, with a rate of convergence  $n^{1/\beta} \ell(n)$ , where  $\ell$  is slowly varying function. The second contribution involves using an extreme value method, e.g. point processes to establish the asymptotic distribution of the proposed estimators and weak convergence for the asymptotically independent joint distribution. An initial observation is that our estimation procedure is especially easy to implement for both  $\phi$  and  $\theta$ . That is, the autoregressive coefficient  $\phi$  in the causal AR(1) process is estimated by taking the minimum of the ratio of two sample values while estimation for the unknown location parameter  $\theta$  was achieved through minimizing  $X_t - \hat{\phi}_n X_{t-1}$  over the observed series.

This naturally motivates a comparison between the estimation procedure presented in this paper and the standard linear programming estimates mentioned above, since within a nonnegative AR(1) model the linear programming estimate reduces to the estimate proposed, namely,  $\min_{1 \leq t \leq n} (X_t / X_{t-1})$ , where  $\{X_t\}$  denotes the AR(1) process. This comparison along with the comparison between Mathew and McCormick's [2] optimization method and Bartlett and McCormick [5] extreme value method was performed through simulation and is presented in Section 3. The results found appear to demonstrate a favorable performance for our extreme value method over the 3 alternative estimators.

The main proofs in this paper rely heavily on point process methods from extreme value theory. The essential idea is to first establish the convergence of a sequence of point processes based on simple quantities and then apply the continuous mapping theorem to obtain

convergence of the desired statistics. More background information on point processes, regular variation, and weak convergence can be found in Resnick [6]. Also, a nice survey on linear programming estimation procedures and nonnegative time series can be found in Anděl [7], Anděl [8], and Datta and McCormick [9], whereas more applications on modeling the phenomena with heavy tailed distributions and ensuing estimation issues can be found in Resnick [10].

The rest of the paper is organized as follows: asymptotic limit results for the autocorrelation parameter  $\phi$ , unknown location parameter  $\theta$ , and joint distribution of  $(\phi, \theta)$  are presented in Section 2, while Section 3 is concerned with the small sample size behavior of these estimates through simulation.

## 2. Asymptotics

The following point process limit result is fundamental. Since the result makes no use of an ARMA structure, we present it for more general linear models subject to usual summability conditions on the coefficients. In that regard for this result, we assume that  $\{X_n, n \geq 0\}$  is the stationary linear process given by

$$X_n = \sum_{j=0}^{\infty} c_j Z_{n-j}$$

with  $\sum_{j=0}^{\infty} |c_j|^\delta < \infty$  for some  $0 < \delta < \beta, \delta \leq 1$ . Further-

more for this result we may relax our assumptions on the innovation distribution and we require that  $|Z_1|$  has a regularly varying tail distribution, *i.e.*,

$P(|Z_1| > x) = x^{-\beta} \ell(x), x > 0$  for a slowly varying function  $\ell$  and the innovation distribution is tail balanced

$$\frac{P(Z_1 > x)}{P(|Z_1| > x)} \rightarrow p \text{ and } \frac{P(Z_1 \leq -x)}{P(|Z_1| > x)} \rightarrow q \text{ as } x \rightarrow \infty.$$

Define point processes

$$\xi_n = \sum_{k=1}^n \epsilon_{(b_n^{-1} X_{k-1}, Z_k)}, n \geq 1$$

and let  $\sum_{k \geq 1} \epsilon_{j_k}$  denote PRM( $\nu$ ) on  $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$  where  $\nu$  has Radon-Nikodym derivative with respect to Lebesgue measure

$$\frac{d\nu}{d\lambda}(x) = p\beta x^{-\beta-1} 1_{(0,\infty)}(x) + q\beta |x|^{-\beta-1} 1_{(-\infty,0)}(x).$$

Let  $\{Z_{i,k}, i \geq 0, k \geq 1\}$  be an iid array with  $Z_{i,k} \stackrel{d}{=} Z_1$

and independent of  $\sum_{k=1}^{\infty} \epsilon_{j_k}$ . Define

$$\xi = \sum_{k=1}^{\infty} \sum_{i=0}^{\infty} \epsilon_{(c_i j_k, Z_{i,k})}$$

Our basic result is to show that  $M_p(\mathbb{R}_0 \times \mathbb{R})$  equipped with the topology of vague convergence

$$\xi_n \xrightarrow{d} \xi \text{ in}$$

which is close in statement and spirit to Theorem 2.4 in Davis and Resnick [11]. In view of the commonality of the two results, we present only the needed changes to the Davis and Resnick proof to accommodate the current setting. Aside from keeping track of the time when points occur, *i.e.* large jumps, the difference in the point processes considered here with those in Davis and Resnick [11] is the inclusion of marks, *i.e.* the second component of the point  $(b_n^{-1}X_{k-1}, Z_k)$ . This complication induces an additional weak dependence in the points which is addressed in Lemma 2.2 through a straight forward blocking argument. First, we establish weak convergence of marked point processes of a normalized vector of innovations. For a positive integer  $m$  define

$$Z_k^{(m)} = (Z_k, Z_{k-1}, \dots, Z_{k-m+1})$$

and point process

$$\mathcal{I}_n^{(m)} = \sum_{k=1}^n \epsilon_{(b_n^{-1}Z_k^{(m)}, Z_{k+1})}, n \geq 1.$$

Let  $e_i = (0, \dots, 1, \dots, 0), 1 \leq i \leq m$  denote the standard basis vectors for  $\mathbb{R}^m$ . Define an associated marked point process with the first component placed on an axis by

$$\tilde{\mathcal{I}}_n^{(m)} = \sum_{k=1}^n \sum_{i=1}^m \epsilon_{(b_n^{-1}Z_k e_i, Z_{k+i})}$$

In the following Lemma, we show that  $\mathcal{I}_n^{(m)}$  and  $\tilde{\mathcal{I}}_n^{(m)}$  are asymptotically indistinguishable in the following sense. Let  $\mathbb{R}_0^m = \mathbb{R}^m / \{0\}$  and  $E = \mathbb{R}_0^m \times \mathbb{R}$ . Consider the class of rectangles

$$\mathcal{S} = \left\{ R \times (-\infty, x] : R = \times_{i=1}^m (c_i, d_i], x < \infty \text{ and } \bar{R} \times (-\infty, x] \subset E \right\}$$

**Lemma 2.1.** *As  $n$  tends to infinity*

$$\mathcal{I}_n^{(m)}(B) - \tilde{\mathcal{I}}_n^{(m)}(B) \xrightarrow{p} 0 \text{ for all } B \in \mathcal{S}.$$

*Proof.* Following the proof presented in Proposition 2.1 of Davis and Resnick [11], suppose that  $B \in \mathcal{S}$  is such that for some  $1 \leq i' \leq m, B \cap (\mathbb{R}e_{i'} \times \mathbb{R}) \neq \emptyset$ . As noted in Davis and Resnick [11] for all  $i \neq i'$ , one has  $c_i < 0 < d_i$ . Observe that

$$\begin{aligned} \mathcal{I}_n^{(m)}(B) &\leq \mathcal{I}_n^{(m)}(\mathbb{R} \times \dots \times [c_{i'}, d_{i'}] \times \dots \times \mathbb{R} \times (-\infty, x]) \\ &= \sum_{k=-i'+2}^{n-i'+1} \epsilon_{(b_n^{-1}Z_k e_{i'}, Z_{k+i'})}(\mathbb{R} \times \dots \times [c_{i'}, d_{i'}] \times \dots \times \mathbb{R} \times (-\infty, x]) \end{aligned} \tag{2.1}$$

Similarly let  $c = \vee_{i \neq i'} c_i < 0$  and  $d = \wedge_{i \neq i'} d_i > 0$ . Then

$$\begin{aligned} &\mathcal{I}_n^{(m)}(B) \\ &\geq \mathcal{I}_n^{(m)}([c, d] \times \dots \times [c_{i'}, d_{i'}] \times \dots \times [c, d] \times (-\infty, x]) \\ &\geq \mathcal{I}_n^{(m)}(\mathbb{R} \times \dots \times [c_{i'}, d_{i'}] \times \dots \times \mathbb{R} \times (-\infty, x]) \\ &\quad - \sum_{i \neq i'} \mathcal{I}_n^{(m)}(E_i \times (-\infty, x]) \end{aligned} \tag{2.2}$$

where  $y = (y_1, \dots, y_m) \in E_i$  according to  $y_i \notin [c, d]$  and  $y_{i'} \in [c_{i'}, d_{i'}]$ . Note that

$$\begin{aligned} &E[\mathcal{I}_n^{(m)}(E_i \times (-\infty, x])] \\ &\leq nP\left(\frac{1}{b_n}|Z_1| > |c| \wedge d, \frac{1}{b_n}|Z_1| > |c_{i'}| \wedge |d_{i'}|\right) \\ &= o(1). \end{aligned} \tag{2.3}$$

Thus from (2.1)-(2.3) we obtain

$$\begin{aligned} &E[\mathcal{I}_n^{(m)}(B) - \tilde{\mathcal{I}}_n^{(m)}(B)] \\ &\leq 2(i' - 1)P\left(\frac{1}{b_n}|Z_1| > |c_{i'}| \wedge |d_{i'}|\right) + o(1) = o(1) \end{aligned} \tag{2.4}$$

as  $n \rightarrow \infty$ .

Then the result follows as in Davis and Resnick [11], Proposition 2.1, completing the proof.  $\square$

**Lemma 2.2.** *Let  $\mathcal{V}_n^{(m)}$  and  $\mathcal{V}^{(m)}$  be the point processes on the space  $E = \mathbb{R}_0 \times \mathbb{R}^m$  defined by*

$$\mathcal{V}_n^{(m)} = \sum_{k=1}^n \epsilon_{(b_n^{-1}Z_k, Z_{k+m})} \text{ and } \mathcal{V}^{(m)} = \sum_{k=1}^{\infty} \epsilon_{(Y_k, Y_k^{(m)})}$$

where  $\{Y_k^{(m)}, k \geq 1\}$  is an iid sequence with  $Y_1^{(m)} = (Z_1, \dots, Z_m)$  and is independent of  $\sum_{k=1}^{\infty} \epsilon_{j_k}$ . Then in  $M_p(E)$ ,

$$\mathcal{V}_n^{(m)} \xrightarrow{d} \mathcal{V}^{(m)}.$$

*Proof.* We employ a blocking argument to establish this result. Let  $r_n$  be a sequence of integers such that  $r = r_n \rightarrow \infty$  as  $n \rightarrow \infty$  and  $r = o(n)$ . Let  $h = \lfloor n/r \rfloor$  and  $l \geq m$ . Define blocks

$$\begin{aligned} I_{n,s} &= [r(s-1) + 1, \dots, rs - l], \\ J_{n,s} &= [rs - l + 1, \dots, rs] \\ &\text{for } 1 \leq s \leq h \text{ and } J_{n,h+1} = [rh + 1, \dots, n]. \end{aligned}$$

Then it is clear that for  $s \neq t$

$$\begin{aligned} &\sigma(Z_{k+i}, k \in I_{n,s}, 0 \leq i \leq m) \\ &\perp\!\!\!\perp \sigma(Z_{k+i}, k \in I_{n,t}, 0 \leq i \leq m). \end{aligned}$$

Write

$$\mathcal{V}_{n,s}^{(m)} = \sum_{k \in I_{n,s}} \epsilon_{(b_n^{-1}Z_k, Z_{k+m})} \text{ and } \tilde{\mathcal{V}}_{n,t}^{(m)} = \sum_{k \in J_{n,t}} \epsilon_{(b_n^{-1}Z_k, Z_{k+m})}$$

Then

$$\mathcal{V}_n^{(m)} = \sum_{s=1}^h \mathcal{V}_{n,s}^{(m)} + \sum_{t=1}^{h+1} \tilde{\mathcal{V}}_{n,t}^{(m)}. \tag{2.5}$$

Let  $B \subset E$  be a disjoint union of rectangles

$$B = \bigcup_{i=1}^j B_i \tag{2.6}$$

where  $B_i = [c_i, d_i] \times R_i$  with  $R_i = \times_{l=1}^m [x_{il}, y_{il}]$ . Let  $\mu = \nu \times F \times \dots \times F$  denote the mean measure of  $\mathcal{V}^{(m)}$  which is PRM( $\mu$ ) on  $E$ . To complete the proof we first show that for all sets  $B$  of the form given in (2.6) that

$$\lim_{n \rightarrow \infty} P(\mathcal{V}_n^{(m)}(B) = 0) = \exp(-\mu(B)). \tag{2.7}$$

The above limit result follows from the easily verifiable relations:

$$\begin{aligned} P(\mathcal{V}_{n,1}^{(m)}(B) = \dots = \mathcal{V}_{n,h}^{(m)}(B) = 0) \\ = P^h(\mathcal{V}_{n,1}^{(m)}(B) = 0); \end{aligned} \tag{2.8}$$

$$P(\mathcal{V}_{n,s}^{(m)}(B_k) \wedge \mathcal{V}_{n,s}^{(m)}(B_l) \geq 1) = O\left(\frac{r^2}{n^2}\right) \tag{2.9}$$

for  $1 \leq k \neq l \leq j$ ;

$$\begin{aligned} P\left(\mathcal{V}_{n,s}^{(m)}\left(\bigcup_{i=1}^j B_i\right) \geq 1\right) \\ = \sum_{i=1}^j P(\mathcal{V}_{n,s}^{(m)}(B_i) \geq 1) + O\left(\frac{r^2}{n^2}\right); \end{aligned} \tag{2.10}$$

$$P(\mathcal{V}_{n,1}^{(m)}(B_i) \geq 1) = \frac{r}{n} \mu(B_i)(1 + o(1)); \tag{2.11}$$

and

$$\sum_{t=1}^{h+1} \tilde{\mathcal{V}}_{n,t}^{(m)}(B) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty. \tag{2.12}$$

Indeed, in view of (2.5) and (2.12), (2.7) is equivalent to showing

$$\lim_{n \rightarrow \infty} P\left(\sum_{s=1}^h \mathcal{V}_{n,s}^{(m)}(B) = 0\right) = \exp(-\mu(B)) \tag{2.13}$$

and the above relation holds by (2.8), (2.10), and (2.11), viz.

$$\begin{aligned} P\left(\sum_{s=1}^h \mathcal{V}_{n,s}^{(m)}(B) = 0\right) &= \left(1 - P(\mathcal{V}_{n,1}^{(m)}(B) \geq 1)\right)^h \\ &= \left(1 - \sum_{i=1}^j \frac{r}{n} \mu(B_i) + o(r/n)\right)^h \\ &\rightarrow \exp\left(-\sum_{i=1}^j \mu(B_i)\right). \end{aligned}$$

It is immediate that for a rectangle

$B = [c, d] \times \times_{i=1}^m [x_i, y_i] \subset E$  we have

$$\lim_{n \rightarrow \infty} E[\mathcal{V}_n^{(m)}(B)] = \mu(B). \tag{2.14}$$

Therefore the result is seen to hold by (2.7) and (2.14) by application of Theorem 4.7 in Kallenberg [12].  $\square$

**Lemma 2.3.** Let  $\mathcal{I}_n^{(m)}$  and  $\mathcal{I}^{(m)}$  be point processes on the space  $E = \mathbb{R}_0^{(m)} \times \mathbb{R}$

$$\mathcal{I}_n^{(m)} = \sum_{k=1}^n \epsilon_{(b_n^{-1} Z_k^{(m)}, Z_{k+1})}, n \geq 1 \text{ and } \mathcal{I}^{(m)} = \sum_{k=1}^{\infty} \sum_{i=1}^m \epsilon_{(j_k e_i, Z_{i,k})}$$

Then in  $M_p(E)$ ,

$$\mathcal{I}_n^{(m)} \xrightarrow{d} \mathcal{I}^{(m)}.$$

*Proof.* We begin by applying the argument used in Theorem 2.2 of Davis and Resnick [11] with the modification that the relevant composition of maps of point processes is given by

$$\begin{aligned} \sum_{k \geq 1} \epsilon_{(u_k, v_{mk}, \dots, v_{1k})} &\mapsto \left( \sum_{k \geq 1} \epsilon_{(u_k, v_{mk})}, \dots, \sum_{k \geq 1} \epsilon_{(u_k, v_{1k})} \right) \\ &\mapsto \left( \sum_{k \geq 1} \epsilon_{(u_k e_1, v_{1k})}, \dots, \sum_{k \geq 1} \epsilon_{(u_k e_m, v_{mk})} \right) \\ &\mapsto \sum_{k \geq 1} \sum_{i=1}^m \epsilon_{(u_k e_i, v_{ik})}. \end{aligned}$$

Each map being continuous, the composition is a continuous map from  $M_p(\mathbb{R}_0 \times \mathbb{R}^m)$  to  $M_p(\mathbb{R}_0^{(m)} \times \mathbb{R})$  with each space being equipped with the topology of vague convergence. Therefore by the continuous mapping theorem and Lemma 2.2 we obtain

$$\begin{aligned} \tilde{\mathcal{I}}_n^{(m)} = \sum_{k=1}^{\infty} \sum_{i=1}^m \epsilon_{(b_n^{-1} Z_k e_i, Z_{k+i})} &\xrightarrow{d} \sum_{k=1}^{\infty} \sum_{i=1}^m \epsilon_{(j_k e_i, Z_{i,k})} \\ = \mathcal{I}^{(m)}. \end{aligned} \tag{2.15}$$

Finally we complete the proof by Lemma 2.1 and (2.15) arguing as in Davis and Resnick [11].

We are now ready to present our fundamental result.

**Theorem 2.1.** Let  $\xi_n$  and  $\xi$  be the point processes on the space  $E = \mathbb{R}_0 \times \mathbb{R}$  defined by

$$\xi_n = \sum_{k=1}^n \epsilon_{(b_n^{-1} X_{k-1}, Z_k)}, n \geq 1 \text{ and } \xi = \sum_{k=1}^{\infty} \sum_{i=0}^{\infty} \epsilon_{(c_i j_k, Z_{i,k})}$$

where  $\sum_{k=1}^{\infty} \epsilon_{j_k}$  is PRM( $\nu$ ) and  $\{Z_{i,k}, i \geq 0, k \geq 1\}$  is

an iid array with  $Z_{i,k} \stackrel{d}{=} Z_1$  and independent of  $\sum_{k=1}^{\infty} \epsilon_{j_k}$ .

Then in  $M_p(E)$ ,

$$\xi_n \xrightarrow{d} \xi.$$

**Remark.** Apart from considering a time coordinate and restricting the process to an AR(1) process, the above Theorem 2.1 and Theorem 3.1 in Mathew and Mc-

Cormick [2] consider essentially the same point process limit result. However, their result gave a wrong limit point process. This error is corrected in the current paper.

*Proof.* Observe that the map

$$(z_k, z_{k-1}, \dots, z_{k-m+1}) \mapsto \sum_{i=0}^{m-1} c_i z_{k-i}$$

induces a continuous map on point processes given by

$$\sum_{k=1}^{\infty} \epsilon_{(z_k, z_{k-1}, \dots, z_{k-m+1}, z_{k+1})} \mapsto \sum_{k=1}^{\infty} \epsilon_{\left(\sum_{j=0}^{m-1} c_j z_{k-j}, z_{k+1}\right)}$$

Thus we obtain from Lemma 2.3 that

$$\sum_{k=1}^n \epsilon_{\left(b_n^{-1} \sum_{i=0}^{m-1} c_i z_{k-i}, z_{k+1}\right)} \xrightarrow{d} \sum_{k=1}^{\infty} \sum_{i=0}^{m-1} \epsilon_{(c_i j_k, Z_{i,k})}. \quad (2.16)$$

The result now follows from (2.16) by the same argument in Davis and Resnick [11] to finish their Theorem 2.4.

Returning to the AR(1) model under discussion in this paper and the estimate  $\hat{\phi}_n$  given in (1.2), we obtain the following asymptotic limit result.

**Theorem 2.2.** *Let  $\{X_t, t \geq 0\}$  be the stationary solution to the AR(1) recursion  $X_t = \phi X_{t-1} + Z_t$  and*

$$\hat{\phi}_n = \bigwedge_{t=1}^n \frac{X_t}{X_{t-1}}. \text{ Under the assumptions that } 0 < \phi < 1$$

*and the innovation distribution  $F$  has regularly varying right tail with index  $-\beta$  and finite positive left endpoint  $\theta$ ,*

$$\lim_{n \rightarrow \infty} P\left(b_n (\hat{\phi}_n - \phi) > x\right) = e^{-x^\beta EW^{-\beta}} \text{ for all } x > 0,$$

where  $W = \bigwedge_{i=0}^{\infty} \frac{Z_i}{\phi^i}$  and  $b_n = F^{\leftarrow}(1-1/n)$ .

*Proof.* For  $x > 0$  define a subset

$$Q_x = \left\{ (x_1, x_2) : \frac{x_2}{x_1} \leq x, x_1 > 0, x_2 \geq \theta \right\}.$$

Then note that for the point processes

$$\xi_n = \sum_{k=1}^n \epsilon_{(b_n^{-1} X_{k-1}, Z_k)}, \text{ we have}$$

$$\{\xi_n(Q_x) = 0\} = \left\{ \bigwedge_{k=1}^n \frac{Z_k}{b_n^{-1} X_{k-1}} > x \right\} = \left\{ b_n (\hat{\phi}_n - \phi) > x \right\}.$$

Applying Theorem 2.1 in the case of an AR(1) process so that  $c_i = \phi^i, i \geq 0$ , we have

$$\xi_n \xrightarrow{d} \xi = \sum_{k=1}^{\infty} \sum_{i=0}^{\infty} \epsilon_{(\phi^i j_k, Z_{i,k})}.$$

Note that as a subset of  $E = (0, \infty) \times [\theta, \infty)$ , the set  $Q_x$  is a bounded continuity set with respect to the limit point process  $\xi$  so that

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(b_n (\hat{\phi}_n - \phi) > x\right) &= P(\xi(Q_x) = 0) \\ &= P\left(\bigwedge_{k=1}^{\infty} \bigwedge_{i=0}^{\infty} \frac{Z_{i,k}}{\phi^i j_k} > x\right) = P\left(\bigwedge_{k=1}^{\infty} \frac{W_k}{j_k} > x\right) \end{aligned}$$

where  $(W_k = \bigwedge_{i=0}^{\infty} \phi^{-i} Z_{i,k}, k \geq 1)$  is an i.i.d. sequence independent of  $(j_k, k \geq 1)$ . Let

$$\tilde{\xi} = \sum_{k \geq 1} \epsilon_{(j_k, W_k)}. \quad (2.17)$$

Then by Proposition 5.6 in Resnick [10], we have that if  $G$  denotes the distribution of  $W_1$ , then  $\tilde{\xi}$  is a Poisson random measure on  $E$  with mean measure  $\mu = \nu \times G$ , where  $\nu(dx) = \beta x^{-\beta-1} 1_{[0, \infty)}(x) dx$ . Using (2.17) we can write

$$P\left(\bigwedge_{k=1}^{\infty} \frac{W_k}{j_k} > x\right) = P(\tilde{\xi}(Q_x) = 0) = \exp(-\mu(Q_x)).$$

Since  $\mu(Q_x) = x^\beta EW^{-\beta}$ , the result follows.

**Corollary 2.1.** *Under conditions given in Theorem 2.2,*

$$\hat{\phi}_n \xrightarrow{a.s.} \phi.$$

*Proof.* Since  $b_n = F^{\leftarrow}(1-1/n) \rightarrow \infty$  we have  $\hat{\phi}_n \xrightarrow{p} \phi$ . But this implies  $\hat{\phi}_n \xrightarrow{a.s.} \phi$ . since  $\hat{\phi}_n \geq \phi$  and is non-increasing.

Let us now define our estimate of  $\theta$ :

$$\hat{\theta}_n = \bigwedge_{t \in I_n} (X_t - \hat{\phi}_n X_{t-1}),$$

where we define the index set

$I_n = \left\{ t : 1 \leq t \leq n \text{ and } X_{t-1} \leq (a_n b_n)^\rho \right\}$  where  $0 < \rho < 1$  is a fixed value.

**Lemma 2.4.** *Under the assumptions that  $F$  is regularly varying with index  $\alpha$  at its positive left endpoint  $\theta$  and  $\bar{F}$  is regularly varying with index  $-\beta$  at infinity, its right endpoint, and  $\alpha > \beta$ , then*

$$a_n^{-1} \left( \hat{\theta}_n - \bigwedge_{t \in I_n} Z_t \right) \xrightarrow{p} 0, \text{ as } n \rightarrow \infty,$$

where  $a_n = F^{\leftarrow}(1/n) - \theta$ . Furthermore, for any  $y \geq 0$

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(a_n^{-1} (\hat{\theta}_n - \theta) > y\right) \\ = \lim_{n \rightarrow \infty} P\left(a_n^{-1} \left(\bigwedge_{t=1}^n Z_t - \theta\right) > y\right) = e^{-y^\alpha}. \end{aligned}$$

*Proof.* Since  $\alpha > \beta$ , we have  $\lim_{n \rightarrow \infty} a_n b_n = \infty$ . Therefore since  $(b_n (\hat{\phi}_n - \phi), n \geq 1)$  is a tight sequence by Theorem 2.2 and since  $\max_{t \in I_n} X_{t-1} / (a_n b_n)^\rho \leq 1$  with  $0 < \rho < 1$ , we have

$$a_n^{-1} \left( \hat{\phi}_n - \phi \right) \bigvee_{t \in I_n} X_{t-1} \xrightarrow{p} 0.$$

The first statement now follows since

$$a_n^{-1} \left| \hat{\theta}_n - \bigwedge_{t \in I_n} Z_t \right| \leq a_n^{-1} \left| \hat{\phi}_n - \phi \right| \bigvee_{t \in I_n} X_{t-1}.$$

For the second statement observe

$$\begin{aligned} 0 &\leq P \left( a_n^{-1} \left( \bigwedge_{t \in I_n} Z_t - \theta \right) > y \right) - P \left( a_n^{-1} \left( \bigwedge_{t=1}^n Z_t - \theta \right) > y \right) \\ &= P \left( a_n^{-1} \left( \bigwedge_{t=1}^n Z_t - \theta \right) \leq y < a_n^{-1} \left( \bigwedge_{t \in I_n} Z_t - \theta \right) \right) \\ &\leq P \left( \bigcup_{1 \leq t \leq n} \left( X_{t-1} > (a_n b_n)^\rho \text{ and } a_n^{-1} (Z_t - \theta) \leq y \right) \right) \\ &\leq n P \left( X_0 > (a_n b_n)^\rho \right) P(Z_1 \leq a_n y + \theta) = o(1). \end{aligned} \tag{2.18}$$

The result for the second statement now follows from (2.18) and the first part of the lemma. Finally, the identification of the limit distribution is well known. □

A useful observation follows from this lemma which we state as a corollary.

**Corollary 2.2.** *Under the assumptions of Lemma 2.4 for any  $x, y > 0$*

$$\begin{aligned} &\lim_{n \rightarrow \infty} \left\{ P b_n \left( \left( \hat{\phi}_n - \phi \right) > x, a_n^{-1} \left( \hat{\theta}_n - \theta \right) > y \right) \right. \\ &\quad \left. - P \left( b_n \left( \hat{\phi}_n - \phi \right) > x, a_n^{-1} \left( \bigwedge_{t=1}^n (Z_t - \theta) > y \right) \right) \right\} = 0 \end{aligned}$$

*Proof.* By Lemma 2.4 we have

$$\begin{aligned} &\lim_{n \rightarrow \infty} \left\{ P \left( b_n \left( \hat{\phi}_n - \phi \right) > x, a_n^{-1} \left( \hat{\theta}_n - \theta \right) > y \right) \right. \\ &\quad \left. - P \left( b_n \left( \hat{\phi}_n - \phi \right) > x, a_n^{-1} \left( \bigwedge_{t \in I_n} Z_t - \theta \right) > y \right) \right\} = 0 \end{aligned}$$

The result then follows from

$$\begin{aligned} 0 &\leq E 1 \left\{ b_n \left( \hat{\phi}_n - \phi \right) > x \right\} \\ &\cdot \left[ 1 \left\{ a_n^{-1} \left( \bigwedge_{t \in I_n} Z_t - \theta \right) > y \right\} - 1 \left\{ a_n^{-1} \left( \bigwedge_{t=1}^n Z_t - \theta \right) > y \right\} \right] \\ &\leq P \left( a_n^{-1} \left( \bigwedge_{t \in I_n} Z_t - \theta \right) > y \right) - P \left( a_n^{-1} \left( \bigwedge_{t=1}^n Z_t - \theta \right) > y \right). \end{aligned}$$

□

Corollary 2.2 allows a simplification in determining the joint asymptotic behavior of  $(\hat{\phi}_n, \hat{\theta}_n)$  by allowing us to replace  $\hat{\theta}_n$  with  $\min_{1 \leq t \leq n} Z_t$ . The next lemma will provide another useful simplification—this time on  $\hat{\phi}_n$ .

For a positive integer  $m$  define

$$X_t^{(m)} = \sum_{i=0}^{m-1} \phi^i Z_{t-i}.$$

**Lemma 2.5.** *Let  $U_n^{(m)}$  and  $U_n$  be defined as*

$$U_n^{(m)} = \bigwedge_{t=1}^n \frac{Z_t}{b_n^{-1} X_{t-1}^{(m)}} \text{ and } U_n = \bigwedge_{t=1}^n \frac{Z_t}{b_n^{-1} X_{t-1}}.$$

Then for any  $\epsilon > 0$

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} P \left( \left| U_n^{(m)} - U_n \right| > \epsilon \right) = 0.$$

*Proof.* We first note for any positive  $M$  that

$$\begin{aligned} P \left( \left| U_n^{(m)} - U_n \right| > \epsilon \right) &= P \left( U_n^{(m)} U_n \left| \frac{1}{U_n^{(m)}} - \frac{1}{U_n} \right| > \epsilon \right) \\ &\leq P \left( \left| \frac{1}{U_n^{(m)}} - \frac{1}{U_n} \right| > \epsilon / M^2 \right) + P \left( U_n^{(m)} > M \right). \end{aligned}$$

In order to calculate  $P \left( \left| \frac{1}{U_n^{(m)}} - \frac{1}{U_n} \right| > \epsilon \right)$  we partition

$X_t$ . That is, we write

$$X_t = X_t^{(m)} + X_t^{r(m)}$$

where  $X_t^{r(m)} = \sum_{j=m}^{\infty} \phi^j Z_{t-j}$ , so that

$$0 \leq \bigvee_{t=1}^n \frac{X_{t-1}}{Z_t} - \bigvee_{t=1}^n \frac{X_{t-1}^{(m)}}{Z_t} \leq \bigvee_{t=1}^n \frac{X_{t-1}^{r(m)}}{Z_t}.$$

Define point processes

$$\xi_n^{(m)} = \sum_{t=1}^n \mathcal{E}_{(b_n^{-1} X_{t-1}^{(m)}, Z_t)} \text{ and } \zeta_n^{(m)} = \sum_{k=1}^{\infty} \sum_{i=m}^{\infty} \epsilon_{(\phi^i j_k, Z_{i,k})}$$

where  $\{j_k, k \geq 1\}$  and  $\{Z_{i,k}, k \geq 1, i \geq 0\}$  have the distributional properties given in Theorem 2.2. Applying Theorem 2.2 with  $c_i = \phi^i$  for  $i \geq m$  and  $c_i$  equal to 0 otherwise, we obtain

$$\xi_n^{(m)} \xrightarrow{d} \zeta_n^{(m)}.$$

Then letting for  $x > 0$

$$\mathcal{R}_x = \left\{ (u, v) : u > 0, v > \theta, \text{ and } \frac{u}{v} > x \right\},$$

we have

$$\begin{aligned} &\lim_{n \rightarrow \infty} P \left( \bigvee_{t=1}^n b_n^{-1} X_{t-1}^{(m)} / Z_t \leq x \right) \\ &= \lim_{n \rightarrow \infty} P \left( \zeta_n^{(m)}(\mathcal{R}_x) = 0 \right) = P \left( \zeta_n^{(m)}(\mathcal{R}_x) = 0 \right). \end{aligned}$$

Setting

$$V_k^{(m)} = \bigvee_{i=m}^{\infty} \frac{\phi^i}{Z_{i,k}} \text{ and } \zeta^{(m)} = \sum_{k=1}^{\infty} \epsilon_{(j_k, V_k^{(m)})}$$

we have

$$P(\xi^{(m)}(\mathcal{R}_x) = 0) = P\left(\bigvee_{k=1}^{\infty} j_k V_k^{(m)} \leq x\right) = P(\zeta^{(m)}(\mathcal{R}_x) = 0),$$

where

$$\mathcal{S}_x = \{(u, v) : u > 0, v > \theta, \text{ and } uv > x\}.$$

Since  $\zeta^{(m)}$  is Poisson random measure with mean measure  $\mu_m = \nu \times H_m$  where  $V_1^{(m)} \sim H_m$  and

$$\mu_m(\mathcal{S}_x) = x^{-\beta} E(V_1^{(m)})^\beta,$$

we obtain

$$\lim_{n \rightarrow \infty} P\left(\bigvee_{t=1}^n b_n^{-1} X_{t-1}^{(m)} / Z_t \leq x\right) = \exp\left(-x^{-\beta} E(V_1^{(m)})^\beta\right).$$

Next since

$$\left|\frac{1}{U_n^{(m)}} - \frac{1}{U_n}\right| \leq \bigvee_{t=1}^n \frac{b_n^{-1} X_{t-1}^{(m)}}{Z_t},$$

we have for large  $n$  that

$$\begin{aligned} P\left(\left|\frac{1}{U_n^{(m)}} - \frac{1}{U_n}\right| > \epsilon/M^2\right) &\leq P\left(\bigvee_{t=1}^n \frac{b_n^{-1} X_{t-1}^{(m)}}{Z_t} > \epsilon/M^2\right) \\ &\leq 2\left(1 - \exp\left(-\epsilon^{-\beta} M^{2\beta} E(V_1^{(m)})^\beta\right)\right) \\ &\leq 2\epsilon^{-\beta} M^{2\beta} E(V_1^{(m)})^\beta. \end{aligned}$$

Therefore, since  $V_1^{(m)} \leq \phi^m / \theta$ ,

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} P\left(\left|\frac{1}{U_n^{(m)}} - \frac{1}{U_n}\right| > \epsilon/M^2\right) = 0.$$

Next, note that from the limit law for the maximum obtained above, by replacing  $X_{t-1}^{(m)}$  with  $X_{t-1}^{(m)}$  and by taking reciprocals, we derive the limit law for minimum,

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(U_n^{(m)} = \bigwedge_{t=1}^n \frac{Z_t}{b_n^{-1} X_{t-1}^{(m)}} \leq x\right) \\ = 1 - \exp\left(-x^\beta E\tilde{W}_m^{-\beta}\right) \end{aligned} \tag{2.19}$$

where  $\tilde{W}_m$  has the distribution of  $\bigwedge_{i=0}^{m-1} Z_{i,k} / \phi^i$ . Thus, for any integer  $m \geq 1$ ,

$$\lim_{n \rightarrow \infty} P(U_n^{(m)} \leq x) \geq 1 - \exp\left(-(x/\theta)^\beta\right).$$

Thus for any  $\epsilon > 0$ , we have for  $M$  large enough that

$$\limsup_{n \rightarrow \infty} \sup_{m \geq 1} P(U_n^{(m)} > M) < \epsilon$$

completing the proof.  $\square$

With Lemma 2.5 in hand we can now focus our attention to the limiting joint distribution of  $(U_n^{(m)}, \bigwedge_{t=1}^n Z_t)$ . This will be accomplished by a blocking argument. To that end for a fixed positive integer  $k$ , let  $r_n = \lfloor n/k \rfloor$  and define blocks for  $i = 1, \dots, \lfloor n/r_n \rfloor$  by

$$J_i = \left[(i-1)r_n + 1, \dots, ir_n - q\right],$$

and  $J'_i = [ir_n - q + 1, \dots, ir_n]$

where  $q$  is a positive integer greater than  $m$ . Furthermore, let

$$J'_0 = [r_n \lfloor n/r_n \rfloor + 1, \dots, n].$$

Now we define the events

$$\chi_i = \left\{ \exists l \in J_i : \left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}}\right) \leq x \text{ or } a_n^{-1}(Z_l - \theta) \leq y \right\},$$

$i = 1, \dots, l_n$

and

$$\chi'_i = \left\{ \exists l \in J'_i : \left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}}\right) \leq x \text{ or } a_n^{-1}(Z_l - \theta) \leq y \right\},$$

$i = 0, \dots, l_n$

where  $l_n = \lfloor n/r_n \rfloor$ . We begin by showing that the events  $\chi'_i$  are negligible.

**Lemma 2.6.** For any  $x, y > 0$

$$\lim_{n \rightarrow \infty} P\left(\bigcup_{i=0}^{l_n} \chi'_i\right) = 0.$$

*Proof.* Observe that

$$\begin{aligned} nP\left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} \leq x\right) &\leq nP\left(\frac{\theta}{b_n^{-1} X_{l-1}^{(m)}} \leq x\right) \\ &\sim x^\beta \theta^{-\beta} \sum_{i=0}^{m-1} \phi^{i\beta} \leq \frac{x^\beta}{\theta^\beta (1 - \phi^\beta)} \end{aligned} \tag{2.20}$$

and

$$nP(a_n^{-1}(Z_l - \theta) \leq y) \sim y^\alpha. \tag{2.21}$$

Thus for some constant  $c$  and any  $n \geq 1$ ,

$$P\left(\bigcup_{i=0}^{l_n} \chi'_i\right) \leq ck/n$$

establishing the lemma.

Define events  $A_i$  and  $B_i$  by

$$A_i = \left\{ \exists l \in J_i : \left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}}\right) \leq x \right\}$$

and  $B_i = \left\{ \exists l \in J_i : a_n^{-1}(Z_l - \theta) \leq y \right\}.$

The following result provides the asymptotic behavior of the probability of these events.

**Lemma 2.7.** For any  $x > 0, y > 0$ , we have as  $k \rightarrow \infty$

$$\lim_{n \rightarrow \infty} P(A_i) \sim \frac{1}{k} x^\beta E\tilde{W}_m^{-\beta} \text{ and } \lim_{n \rightarrow \infty} P(B_i) \sim \frac{y^\alpha}{k}.$$

*Proof.* Since the events  $A_i$  are independent, we have

$$\begin{aligned} &P\left(\bigcap_{i=1}^k \left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} > x, \forall l \in J_i\right)\right) \\ &= \left(P\left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} > x, \forall l \in J_1\right)\right)^k. \end{aligned}$$

Using Lemma 2.6 we have that

$$\begin{aligned} &P\left(\bigwedge_{l=1}^n \frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} > x\right) \\ &= P\left(\bigcap_{i=1}^k \left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} > x, \forall l \in J_i\right)\right) + O\left(\frac{1}{n}\right) \\ &= \left(P\left(\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} > x, \forall l \in J_1\right)\right)^k + O\left(\frac{1}{n}\right). \end{aligned}$$

From (2.19) we showed that

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(U_n^{(m)} = \bigwedge_{l=1}^n \frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} \leq x\right) \\ = 1 - \exp(-x^\beta E\tilde{W}_m^{-\beta}). \end{aligned}$$

Hence using this limit law on  $\bigwedge_{l=1}^n (Z_l/b_n^{-1} X_{l-1}^{(m)})$ , we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \left(P\left[\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} > x, \forall l \in J_1\right]\right)^k \\ = \exp(-x^\beta E\tilde{W}_m^{-\beta}). \end{aligned}$$

Thus,

$$\lim_{n \rightarrow \infty} P[A_i] \sim \frac{x^\beta}{k} E\tilde{W}_m^{-\beta}, \text{ as } k \rightarrow \infty. \quad (2.22)$$

Similarly using the result of Lemma 2.4, we obtain

$$\lim_{n \rightarrow \infty} P[B_i] \sim \frac{y^\alpha}{k}, \text{ as } k \rightarrow \infty. \quad (2.23)$$

Hence the lemma holds.  $\square$

**Lemma 2.8.** For some constant  $c$

$$P(A_i \cap B_i) \leq cP(A_i)P(B_i).$$

**Remark.** Since the cardinality of  $J_i$  depends on  $r_n$  which depends on  $k$  and the events  $A_i$  and  $B_i$  depend on  $n$ ,  $P(A_i)$  and  $P(B_i)$  depend on  $k$  and  $n$ .

The conclusion of this lemma provides that for all  $k$  and  $n$ , there is a constant dependent on no parameters for which the inequality stated there holds.

*Proof.* To calculate the intersection we define the following sets

$$K_1 = K_{1,i} = \{(l_1, l_2) : l_1, l_2 \in J_i \text{ and } l_2 \notin [l_1 - 1 - m, l_1]\}$$

and

$$K_2 = K_{2,i} = \{(l_1, l_2) : l_1, l_2 \in J_i \text{ and } l_2 \in [l_1 - 1 - m, l_1]\}.$$

Now for  $l_1 - 1 - m \leq l_2 \leq l_1 - 1$  and  $n$  sufficiently large,

$$\begin{aligned} &\left\{\frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x\right\} \cap \{a_n^{-1}(Z_{l_2} - \theta) \leq y\} \\ &\subset \left\{\sum_{i \neq l_1 - l_2 - 1, 0 \leq i \leq m-1} \phi^i Z_{l_1-1-i} \geq b_n \frac{Z_{l_1}}{x} - \theta - 1\right\} \\ &\cap \{Z_{l_2} \leq \theta + a_n y\}. \end{aligned}$$

It then follows from (2.20), (2.21), and independence that

$$\begin{aligned} &P\left(\left\{\sum_{i \neq l_1 - l_2 - 1, 0 \leq i \leq m-1} \phi^i Z_{l_1-1-i} \geq b_n \frac{Z_{l_1}}{x} - \theta - 1\right\} \cap \{Z_{l_2} \leq \theta + a_n y\}\right) \\ &= O\left(\frac{1}{n^2}\right). \end{aligned}$$

If  $l_1 = l_2 = l$ , we have

$$\begin{aligned} &P\left(\left\{\frac{Z_l}{b_n^{-1} X_{l-1}^{(m)}} \leq x\right\} \cap \{a_n^{-1}(Z_l - \theta) \leq y\}\right) \\ &= \int_{\theta}^{\theta + a_n y} P\left(X_{l-1}^{(m)} \geq \frac{z}{b_n^{-1} x}\right) dF(z) = O\left(\frac{1}{n^2}\right). \end{aligned}$$

Therefore, for some constant  $c$

$$\begin{aligned} &P\left(\bigcup_{(l_1, l_2) \in K_2} \left\{\frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x\right\} \cap \{a_n^{-1}(Z_{l_2} - \theta) \leq y\}\right) \\ &= c/nk. \end{aligned}$$

In order to handle set  $K_1$ , observe from construction of the blocks  $J_i$  and set  $K_1$  that if  $(l_1, l_2) \in K_1$  then the events

$$\left\{\frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x\right\} \text{ and } \{a_n^{-1}(Z_{l_2} - \theta) \leq y\}$$

are independent. Thus, if we define  $\{Z'_i, i \in \mathbb{Z}\}$  as an independent copy of  $\{Z_i, i \in \mathbb{Z}\}$ , then



$$\begin{aligned} & P\left(\bigcup_{(l_1, l_2) \in K_1} \left\{ \frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x \right\} \cap \{a_n^{-1}(Z_{l_2} - \theta) \leq y\}\right) \\ &= P\left(\bigcup_{(l_1, l_2) \in K_1} \left\{ \frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x \right\} \cap \{a_n^{-1}(Z'_{l_2} - \theta) \leq y\}\right) \\ &\leq P(A_i \cap B'_i) = P(A_i)P(B_i) \leq c/k^2 \end{aligned}$$

where  $B'_i = \{\exists l \in J_i : a_n^{-1}(Z'_l - \theta) \leq y\}$  and where we used Lemma 2.7 in the last step. Thus, we have that for some constant  $c$

$$\begin{aligned} & P(A_i \cap B_i) \\ &= P\left(\bigcup_{(l_1, l_2) \in K_1} \left\{ \frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x \right\} \cap \{a_n^{-1}(Z_{l_2} - \theta) \leq y\}\right) \\ &\cup \bigcup_{(l_1, l_2) \in K_2} \left\{ \frac{Z_{l_1}}{b_n^{-1} X_{l_1-1}^{(m)}} \leq x \right\} \cap \{a_n^{-1}(Z_{l_2} - \theta) \leq y\} \\ &\leq c/k^2 = O(P(A_i)P(B_i)), \end{aligned}$$

which completes the proof in view of Lemma 2.7.  $\square$

**Lemma 2.9.** For any  $x > 0, y > 0$ ,

$$\begin{aligned} & \lim_{n \rightarrow \infty} P\left(\bigwedge_{t=1}^n \left(\frac{Z_t}{b_n^{-1} X_{t-1}^{(m)}}\right) > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &= \exp(-x^\beta E\tilde{W}_m^{-\beta} - y^\alpha). \end{aligned}$$

*Proof.* First from Lemma 2.6, we have as  $n \rightarrow \infty$  that

$$\begin{aligned} & P\left(\bigwedge_{t=1}^n \left(\frac{Z_t}{b_n^{-1} X_{t-1}^{(m)}}\right) > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &= P\left(\bigcap_{i=1}^k \chi_i^c\right) + o(1). \end{aligned}$$

Next by (2.22), (2.23), and Lemma 2.8 we obtain that as  $k$  tends to infinity

$$\lim_{n \rightarrow \infty} P(\chi_i) \sim \frac{1}{k} (x^\beta E\tilde{W}_m^{-\beta} + y^\alpha).$$

Therefore, we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} P\left(\bigcap_{i=1}^k \chi_i^c\right) &= \lim_{k \rightarrow \infty} \left(1 - \frac{1}{k} (x^\beta E\tilde{W}_m^{-\beta} + y^\alpha)\right)^k \\ &= \exp\left(- (x^\beta E\tilde{W}_m^{-\beta} + y^\alpha)\right). \end{aligned}$$

Hence

$$\begin{aligned} & \lim_{n \rightarrow \infty} P\left(\bigwedge_{t=1}^n \left(\frac{Z_t}{b_n^{-1} X_{t-1}^{(m)}}\right) > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &= \exp\left(-x^\beta E\tilde{W}_m^{-\beta} - y^\alpha\right). \end{aligned}$$

$\square$

**Theorem 2.3.** Let  $\{X_t, t \geq 1\}$  denote the stationary AR(1) process such that the innovation distribution  $F$  satisfies  $\bar{F}$  is  $RV_{-\beta}$  at infinity and  $F$  is  $RV_\alpha$  at  $\theta$ .

If  $\alpha > \beta$  then for any  $x > 0, y > 0$  we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} P\left(b_n(\hat{\phi}_n - \phi) > x, a_n^{-1}(\hat{\theta}_n - \theta) > y\right) \\ &= e^{-x^\beta EW^{-\beta} - y^\alpha}, \end{aligned}$$

where  $W = \bigwedge_{j=0}^\infty Z_j / \phi^j$ .

*Proof.* Let us first observe that for  $\epsilon > 0$

$$\begin{aligned} & P\left(U_n^{(m)} > x + \epsilon, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) - P\left(\left|U_n^{(m)} - U_n\right| > \epsilon\right) \\ &\leq P\left(U_n > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &\leq P\left(U_n^{(m)} > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right). \end{aligned}$$

Thus by Lemma 2.9 we obtain

$$\begin{aligned} & \exp\left\{- (x + \epsilon)^\beta E\tilde{W}_m^{-\beta} - y^\alpha\right\} \\ & - \limsup_{n \rightarrow \infty} P\left(\left|U_n - U_n^{(m)}\right| > \epsilon\right) \\ &\leq \liminf_{n \rightarrow \infty} P\left(U_n > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &\leq \limsup_{n \rightarrow \infty} P\left(U_n > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &\leq \exp\left(-x^\beta E\tilde{W}_m^{-\beta} - y^\alpha\right). \end{aligned}$$

Letting  $m$  tend to infinity in the above and then  $\epsilon$  tend to 0, we obtain from Lemma 2.5 and

$\lim_{m \rightarrow \infty} E\tilde{W}_m^{-\beta} = EW^{-\beta}$  that

$$\begin{aligned} & \lim_{n \rightarrow \infty} P\left(U_n > x, a_n^{-1} \bigwedge_{t=1}^n (Z_t - \theta) > y\right) \\ &= \exp\left(-x^\beta EW^{-\beta} - y^\alpha\right). \end{aligned}$$

The theorem now follows from this and Corollary 2.2.

### 3. Simulation Study

In this section we assess the reliability of our extreme value estimation method through a simulation study. This included a comparison between our estimation procedure and that of three alternative estimation procedures for both the autocorrelation coefficient  $\phi$  and the unknown location parameter  $\theta$  under two different innovation distributions. Additionally, the degree of approximation for the empirical probabilities of  $\hat{\phi}_{\min}$  and  $\hat{\theta}_{\min}$  to its respective limiting distribution was reported.

To study the performance of the estimators

$$\hat{\phi}_{\min} = \bigwedge_{t=1}^n \frac{X_t}{X_{t-1}} \quad \text{and} \quad \hat{\theta}_{\min} = \bigwedge_{t \in I_n} (X_t - \hat{\phi}_{\min} X_{t-1}) \quad \text{re-}$$

spectively, we generated 5000 replications for the non-negative time series  $(X_0, X_1, \dots, X_n)$  for two different sample sizes (500,1000), where  $\{X_t\}$  is an AR(1) process satisfying the difference equation

$$X_t = \phi X_{t-1} + Z_t, \text{ for } (1 \leq t \leq n) \text{ and } Z_t \geq \theta.$$

The autoregressive parameter  $\phi$  is taken to be in the range from 0 to 1 guaranteeing a nonnegative time series and the unknown location parameter  $\theta$  is positive when the innovations  $Z_t$  are taken to be

$$F(z) = \begin{cases} c(z-\theta)^\alpha, & \text{if } \theta < z < \theta+1, \\ 1-d(z-\theta)^{-\beta}, & \text{if } \theta+1 < z < \infty. \end{cases}$$

For this innovation distribution let  $c$  and  $d$  be non-negative constants such that  $c+d=1$ , then this distribution is regularly varying at both endpoints with index of regular variation  $-\beta$  at infinity and index of regular variation  $\alpha$  at  $\theta$ . For this simulation study two distributions were considered:

- i)  $F_1, c=0, d=1$ , ii)  $F_2, c=0.5, d=0.5$ .

Now observe in case i) the innovation distribution  $F_1$  is a Pareto distribution with a regular varying tail distri-

bution at  $\infty$  with index of regular variation  $-\beta$  and regular varying at  $\theta+1$  with a fixed index  $\alpha=1$ , whereas in case ii) the innovation distribution  $F_2$  is regular varying at  $\infty$  and  $\theta$  with no restriction on  $\alpha$  or  $\beta$ .

First we examine the simulation results for  $\phi=0.9$  under  $F_1$  for each of the six different  $\beta$  values considered by computing 5000 estimates using

$$\hat{\phi}_{\min} = \bigwedge_{t=1}^n \frac{X_t}{X_{t-1}}, \hat{\phi}_{\max} = \frac{X_{t^*+1}}{X_{t^*}}, \hat{\phi}_{\text{range}} = \frac{X_{t^*+1} - X_{j^*}}{X_{t^*} - X_{j^*-1}},$$

where  $t^*$  and  $j^*$  provides the index of the maximal and minimal  $X_i$  respectively for  $1 \leq i \leq n$ , and

$$\hat{\phi}_{LS} = \begin{cases} \frac{\sum_{t=1}^{n-1} X_t X_{t+1}}{\sum_{t=1}^n X_t^2}, & \text{if } 0 < \beta < 1 \\ \frac{\sum_{t=1}^n (X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}, & \text{if } 1 < \beta < 3 \end{cases}$$

where  $\bar{X} = \sum_{t=1}^n X_t/n$ . The means and standard deviations (written below in parentheses), of these estimates are reported in **Table 1** along with the average

**Table 1. Comparison of estimators for  $\phi=0.9$  under  $F_1$ .**

$\beta$	$n$	$\hat{\phi}_{\min}$	$\hat{\phi}_{\max}$	$\hat{\phi}_{\text{range}}$	$\hat{\phi}_{LS}$	95% C.I. Avg. Length			
						Min est.	Max est.	Range est.	LS est.
0.2	500	0.9000 (<0.0001)	0.9002 (0.0015)	0.9002 (0.0016)	0.8988 (0.0219)	<0.0001	<0.0001	<0.0001	0.0288
	1000	0.9000 (<0.0001)	0.9001 (0.0009)	0.9002 (0.0009)	0.8997 (0.0112)	<0.0001	<0.0001	<0.0001	0.0091
0.8	500	0.9004 (0.0016)	0.9026 (0.0083)	0.9032 (0.0083)	0.9231 (0.0251)	0.0046	0.0139	0.0228	0.0657
	1000	0.9002 (0.0008)	0.9020 (0.0064)	0.9014 (0.0064)	0.9158 (0.0186)	0.0023	0.0072	0.0133	0.0529
1.2	500	0.9049 (0.0078)	0.9072 (0.0146)	0.9067 (0.0139)	0.8923 (0.0176)	0.0271	0.0449	0.0548	0.0623
	1000	0.9029 (0.0056)	0.9117 (0.0117)	0.9047 (0.0109)	0.8964 (0.0117)	0.0187	0.0325	0.0389	0.0439
1.8	500	0.9191 (0.0622)	0.9202 (0.0673)	0.9136 (0.0612)	0.8947 (0.0381)	0.0562	0.0763	0.0846	0.0695
	1000	0.9113 (0.0313)	0.9208 (0.0367)	0.9116 (0.0384)	0.8989 (0.0223)	0.0501	0.0690	0.0743	0.0485
2.2	500	0.9236 (0.0909)	0.9313 (0.1050)	0.9176 (0.0244)	0.8917 (0.0281)	0.0670	0.0847	0.0944	0.0709
	1000	0.9207 (0.0412)	0.9293 (0.0456)	0.9151 (0.0217)	0.8961 (0.0228)	0.0623	0.0821	0.0874	0.0494
2.8	500	0.9475 (0.1378)	0.9505 (0.1471)	0.9205 (0.0723)	0.8873 (0.0691)	0.0767	0.0905	0.1050	0.0737
	1000	0.9437 (0.0834)	0.9484 (0.1163)	0.9187 (0.0317)	0.8906 (0.0331)	0.0735	0.0885	0.0956	0.0516

length for a 95 percent empirical confidence intervals with exact coverage. Since the main purpose of this section is to compare our estimator  $\hat{\phi}_{\min}$  to Bartlett and McCormick [5] estimator  $\hat{\phi}_{\max}$ , McCormick and Mathew [2] estimator  $\hat{\phi}_{\text{range}}$ , and Davis and Resnick's [13] estimator  $\hat{\phi}_{LS}$ , the confidence intervals were directly constructed from the empirical distributions of

$$n^{1/\beta} (\hat{\phi}_{\min} - \phi), n^{1/\beta} (\hat{\phi}_{\max} - \phi), n^{1/\beta} (\hat{\phi}_{\text{range}} - \phi),$$

$$\text{and } (n/\log n)^{1/\beta} (\hat{\phi}_{LS} - \phi),$$

respectively.

To evaluate and compare the performance of four location estimators, six different scenarios for  $\alpha$  and  $\beta$  are presented in **Table 2** under  $F_2$ . When  $\theta = 2$ , 5000 estimates for each estimator;

$$\hat{\theta}_{\min} = \bigwedge_{t \in I_n} (X_t - \hat{\phi}_{\min} X_{t-1}), \quad \hat{\theta}_{\text{range}} = X_{j^*} - \hat{\phi}_{\text{range}} X_{j^*-1},$$

$\hat{\theta}_{e1} = (1 - \hat{\phi}_{\text{range}}) \bigwedge_{t=1}^n X_t$ , and  $\hat{\theta}_{e2} = (1 - \hat{\phi}_{LS}) \bigwedge_{t=1}^n X_t$  were obtained. The exponent  $\rho$  inside the index set

$$I_n = \left\{ t : 1 \leq t \leq n \text{ and } X_{t-1} \leq (a_n b_n)^\rho \right\},$$

was set to 0.9.

The means and standard deviations (written below in parentheses), of these estimates are reported in **Table 2** along with the average length for a 95 percent empirical confidence intervals. For convenience, the empirical distributions of  $n^{-1/\alpha} (\hat{\theta}_{\min} - \theta)$ ,  $n^{-1/\alpha} (\hat{\theta}_{\text{range}} - \theta)$ ,

$$q_n (\hat{\theta}_{e1} - \theta - (1 - \hat{\phi}_{\text{range}}) w_n),$$

and

$$q_n (\hat{\theta}_{e2} - \theta - (1 - \hat{\phi}_{LS}) w_n)$$

were respectively used, where

the normalizing constants  $q_n$  and  $w_n$  are obtained through Equations (3.12)-(3.16) of McCormick and Mathew [2].

**Remark.** In the case that  $0 < \beta < 1$ ,  $\hat{\theta}_{e1}$  converge at

**Table 2. Comparison of estimators for  $\phi = 2$  under  $F_2$ .**

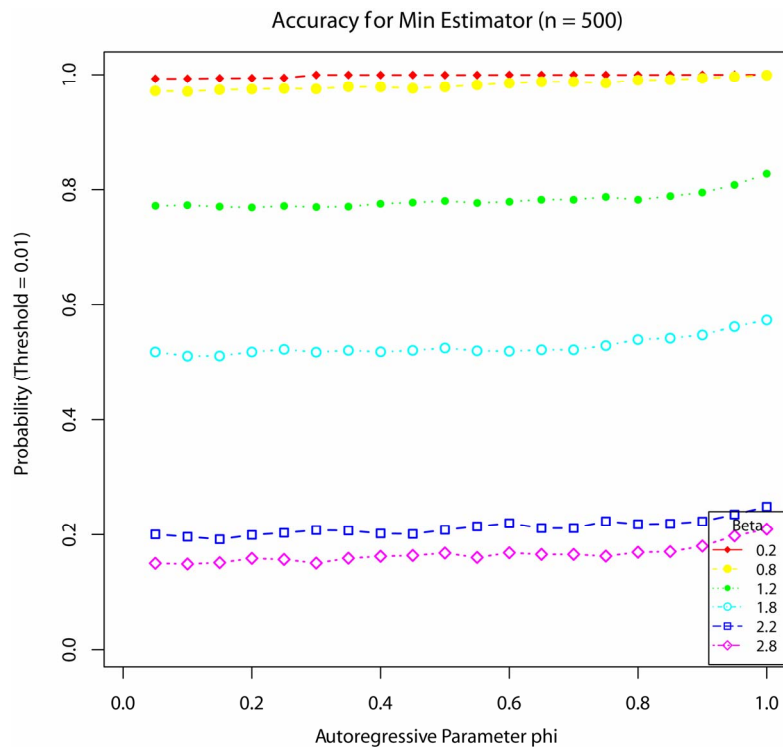
$\beta$	$\alpha$	$n$	$\hat{\theta}_{\min}$	$\hat{\theta}_{\text{range}}$	$\hat{\theta}_{e1}$	$\hat{\theta}_{e2}$	95% C.I. Avg. Length				
							Min est.	Range est.	e1 (range) est.	e2 (LS) est.	
0.5	0.6	500	2.00	2.44	3.16	3.98	<0.0001	2.34	2.09	3.02	
			(0.0087)	(0.4068)	(0.6775)	(0.7657)	-	-	-	-	
		1000	2.00	2.39	3.77	3.67	<0.0001	1.89	1.69	2.60	
			(0.0002)	(0.3605)	(0.4381)	(0.4960)	-	-	-	-	
		1.6	500	1.94	2.68	3.34	4.10	0.0273	2.37	3.18	3.66
				(0.0879)	(0.3566)	(0.6656)	(0.7503)	-	-	-	-
	1000	1.99	2.63	3.94	3.81	0.0152	1.87	2.23	2.79		
		(0.0503)	(0.3199)	(0.4311)	(0.4780)	-	-	-	-		
	2.6	500	1.87	2.77	3.41	4.16	0.1422	2.19	2.85	2.38	
			(0.4697)	(0.7553)	(0.6733)	(0.7632)	-	-	-	-	
	1000	1.89	2.73	3.99	3.87	0.0988	1.74	2.21	2.03		
		(0.4045)	(0.3143)	(0.4276)	(0.4822)	-	-	-	-		
1.5	0.6	500	NA	1.72	2.35	2.15	NA	1.57	1.28	2.21	
			(-)	(0.9410)	(0.5327)	(0.3706)	-	-	-	-	
		1000	NA	1.80	2.19	2.65	NA	1.20	1.25	1.88	
			(-)	(0.9141)	(0.4381)	(0.2565)	-	-	-	-	
		1.6	500	2.80	1.95	2.54	2.26	1.72	2.23	2.18	1.94
				(0.3863)	(0.9569)	(0.5298)	(0.3815)	-	-	-	-
	1000	2.38	2.03	2.37	2.84	1.28	2.10	1.78	1.31		
		(0.2263)	(0.9308)	(0.4461)	(0.2701)	-	-	-	-		
	2.6	500	2.99	2.16	2.79	2.35	1.83	2.47	2.21	2.09	
			(0.3604)	(0.9443)	(0.5271)	(0.3982)	-	-	-	-	
	1000	2.68	2.12	2.27	2.93	1.26	2.10	1.81	1.42		
		(0.2956)	(0.9010)	(0.4398)	(0.2837)	-	-	-	-		

a faster rate than  $\hat{\theta}_{e_2}$  and in the case that  $1 < \beta < 3$ ,  $\hat{\theta}_{e_2}$  converges at a faster rate than  $\hat{\theta}_{e_1}$ . Lastly, since the McCormick and Mathew [2] paper has the restriction that  $Var(Z_1) < \infty$ , only when  $\alpha > 2$  can the estimators  $\hat{\theta}_{e_1}$  and  $\hat{\theta}_{e_2}$  be fairly compared, whereas only when  $\alpha > \beta$  is our estimator applicable.

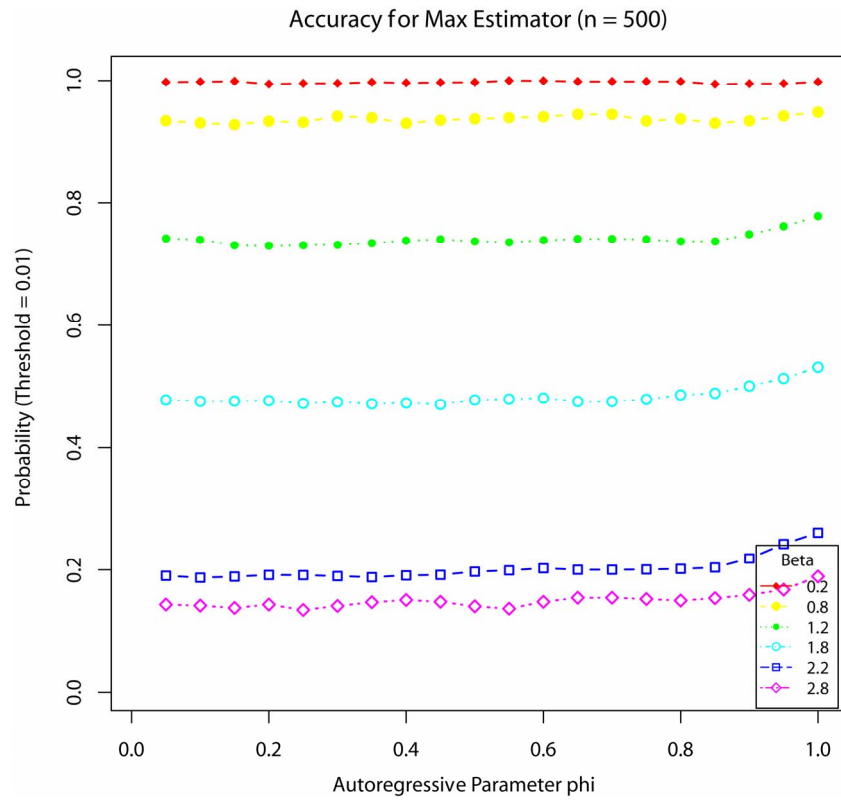
Now observe for the selected  $\beta$  values being considered, **Table 1** shows that our estimator performs at least as well as the three other alternative estimators. This is particularly true under the heavier tail models, *i.e.* when  $0 < \beta < 2$ . In this regime our estimate shows little bias and the average lengths of the confidence intervals are smaller than the other three estimates, sometimes by a wide margin. In particular, when  $\beta = 0.8$  and  $n = 1000$  the 95% confidence interval average length for our method is 3.13, 5.78 and 23 times smaller than the three alternative estimators respectively. This is in part due to the use of one-sided confidence intervals since  $\hat{\phi}_{min} \geq \phi$ , for all  $t \geq 1$ . Naturally, when  $1 < \beta < 3$ , Davis and Resnick least square estimator is more efficient than all three extreme value estimators. While our estimator  $\hat{\phi}_{min}$  will always perform slightly better than the  $\hat{\phi}_{max}$  estimator, Bartlett's and McCormick [5] estimator  $\hat{\phi}_{max}$  main advantage lies with its versatility to perform well for various nonnegative time series, including but not restricted to higher order autoregressive models, along with ARMA models.

**Table 2** reveals that our estimator for  $\theta$  generally performs better than the three alternative estimators for  $\alpha = 0.6, 1.6, 2.6$  when  $0 < \beta < 1$ . This is particularly true when comparing average confidence interval lengths. Although all three estimators  $\hat{\theta}_{min}$ ,  $\hat{\theta}_{range}$ , and  $\hat{\theta}_{e_1}$  converge to the true value of the parameter  $\theta$  as  $n$  tends to infinity respectively, in this setting they may not compete asymptotically with, say, a conditional least square estimator  $\hat{\theta}_{e_2}$  when  $\beta > 1$ . Nonetheless for small sample sizes our simulation study favors  $\hat{\theta}_{range}$  over the other three estimators. The difficulty for a least square estimate is that a small negative bias for the estimate of the autocorrelation parameter  $\phi$  gives rise to a much larger positive bias in the estimate of  $\hat{\theta}_{e_2}$ . While the affect is not as great, the positive bias found in our estimator  $\hat{\phi}_{min}$  and the others for  $\phi$  has a significant effect on the estimate for  $\theta$ .

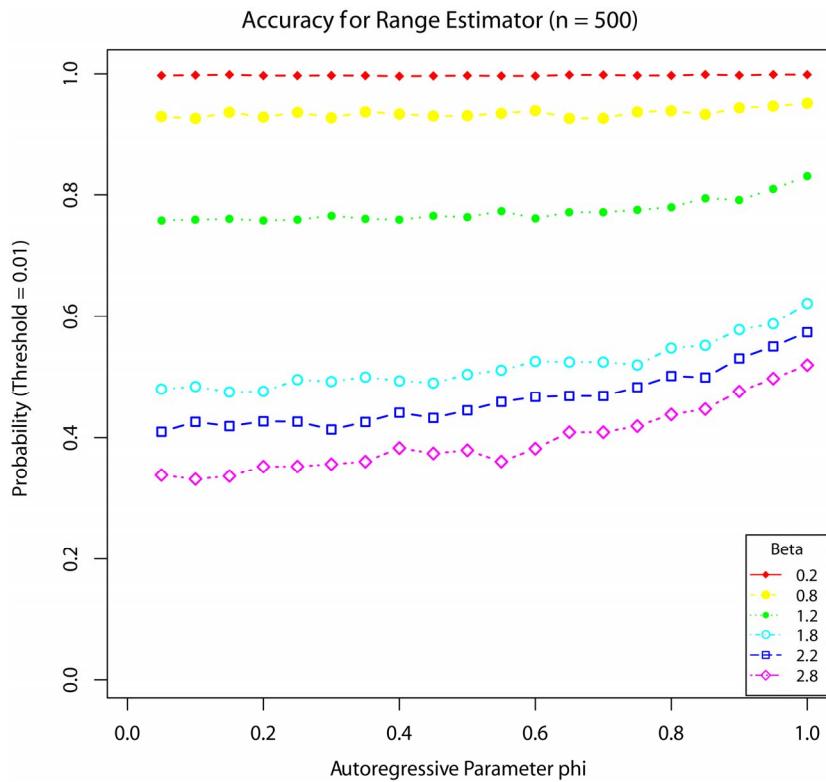
**Figures 1-4** show a comparison between the probability that estimators  $\hat{\phi}_{min}$ ,  $\hat{\phi}_{max}$ ,  $\hat{\phi}_{range}$ , and  $\hat{\phi}_{LS}$  are within 0.01 of the true autocorrelation parameter value, respectively. With a sample size of 500, these figures plotted the sample fraction of estimates which fell within a bound of  $\varepsilon = 0.01$  of the true value. Good performance with respect to this measure is reflected in curves near to 1.0 with diminishing good behavior as curves approach 0.0. When  $0 < \beta < 1$ , the figures seem to show that our estimator compared to the other three



**Figure 1.**  $P\left[|\hat{\phi}_{min} - \phi| < 0.01\right]$  for  $RV_{-\beta}$ .



**Figure 2.**  $P\left[\left|\hat{\phi}_{\max} - \phi\right| < 0.01\right]$  for  $RV_{-\beta}$ .



**Figure 3.**  $P\left[\left|\hat{\phi}_{\text{range}} - \phi\right| < 0.01\right]$  for  $RV_{-\beta}$ .

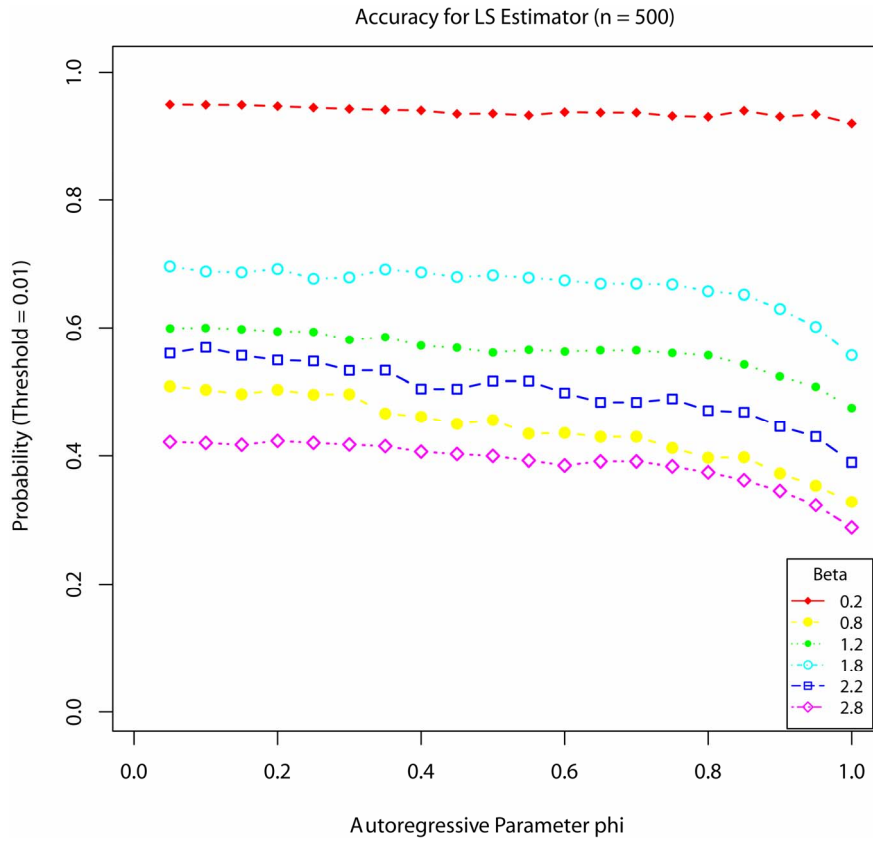


Figure 4.  $P\left[|\hat{\phi}_{LS} - \phi| < 0.01\right]$  for  $RV_{-\beta}$ .

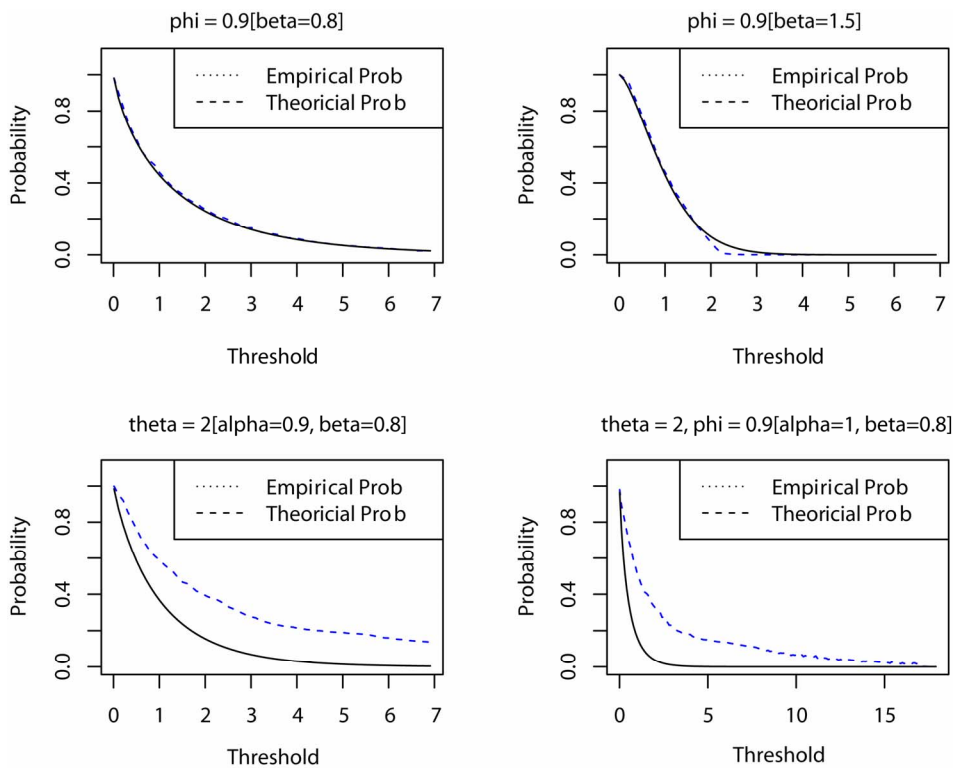


Figure 5. Empirical vs. theoretical probability.

produced a higher fraction of precise estimates, especially compared to Davis and Resnick estimator. When the regular variation index value is closer to 2, we see a higher fraction of the Davis-Resnick estimates showing better accuracy by this measure. The figures also indicate that McCormick and Mathew's range estimator produced a consistent high fraction of precise estimates when  $\beta > 2$ .

Lastly, we performed a Monte Carlo simulation to study the degree of approximation for the empirical probability  $P[b_n(\hat{\phi}_{\min} - \phi) > x]$ ,  $P[a_n^{-1}(\hat{\theta}_{\min} - \theta) > y]$ , and  $P[b_n(\hat{\phi}_{\min} - \phi) > x, a_n^{-1}(\hat{\theta}_{\min} - \theta) > y]$  to its limiting values  $e^{-x^\beta EW^{-\beta}}$ ,  $e^{-y^\alpha}$ , and  $e^{-x^\beta EW^{-\beta} - y^\alpha}$  respectively. The empirical distributions were calculated from 5000 replications of the nonnegative time series  $(X_0, X_1, \dots, X_n)$  for a sample size of 5000, where

$$EW^{-\beta} = 1 - \sum_{i=0}^M \frac{\phi^{\beta(i+3)/2}}{i+2} (1 - \phi^{\beta(i+2)}), \text{ and } M \text{ was set to}$$

500. Additionally, we restricted  $\alpha > \beta$ . The top two plots in **Figure 5** below shows the performance when  $Z_t \sim F_1$  and the autocorrelation coefficient  $\phi$  is 0.9 for  $\alpha=1$  and  $\beta$  equal to 0.8, 1.5 respectively. Observe for  $0 < x < 7$  that the empirical tail probability  $b_n(\hat{\phi}_{\min} - \phi) > x$  mirrors the theoretical probability quite nicely. The lower left plot in **Figure 5** displays the asymptotic performance when  $Z_t \sim F_2$  and the location parameter  $\theta$  is 2 for  $\alpha=0.9$ ,  $\beta=0.8$ . Notice that the convergence rate of the empirical probability to the theoretical probability is extremely slow. This is not surprising since on average our estimate falls more than 0.1 from the true value when  $\beta=0.8$ . The lower right plot in **Figure 5** displays the asymptotic performance when  $Z_t \sim F_1$  for the joint distribution of  $(\hat{\phi}_{\min}, \hat{\theta}_{\min})$ . Observe that this plot solidifies the asymptotic independence between  $b_n(\hat{\phi}_{\min} - \phi)$  and  $a_n^{-1}(\hat{\theta}_{\min} - \theta)$ .

## REFERENCES

- [1] P. J. Brockwell and R. A. Davis, "Time Series: Theory and Methods," 2nd Edition, Springer, New York, 1987.
- [2] W. P. McCormick and G. Mathew, "Estimation for Non-negative Autoregressive Processes with an Unknown Location Parameter," *Journal of Time Series Analysis*, Vol. 14, No. 1, 1993, pp. 71-92. [doi:10.1111/j.1467-9892.1993.tb00130.x](https://doi.org/10.1111/j.1467-9892.1993.tb00130.x)
- [3] A. E. Raftery, "Estimation Efficace Pour un Processus Autoregressif Exponentiel a Densite Discontinue," *Publications de l'Institut de Statistique de l'Universit  de Paris*, Vol. 25, No. 1, 1980, pp. 65-91.
- [4] R. A. Davis and W. P. McCormick, "Estimation for First-Order Autoregressive Processes with Positive or Bounded Innovations," *Stochastic Processes and Their Applications*, Vol. 31, No. 1, 1989, pp. 237-250. [doi:10.1016/0304-4149\(89\)90090-2](https://doi.org/10.1016/0304-4149(89)90090-2)
- [5] A. Bartlett and W. P. McCormick, "Estimation for Non-Negative Time Series with Heavy-Tail Innovations," *Journal of Time Series Analysis*, 2012. <http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291467-9892/earlyview>
- [6] S. I. Resnick, "Point Processes, Regular Variation and Weak Convergence," *Advances in Applied Probability*, Vol. 18, No. 1, 1986, pp. 66-138. [doi:10.2307/1427239](https://doi.org/10.2307/1427239)
- [7] Jiří Anděl, "Non-Negative Autoregressive Processes," *Journal of Time Series Analysis*, Vol. 10, No. 1, 1989, pp. 1-11. [doi:10.1111/j.1467-9892.1989.tb00011.x](https://doi.org/10.1111/j.1467-9892.1989.tb00011.x)
- [8] Jiří Anděl, "Non-Negative Linear Processes," *Applications of Mathematics*, Vol. 36, No. 1, 1991, pp. 277-283.
- [9] S. Datta and W. P. McCormick, "Bootstrap Inference for a First-Order Autoregression with Positive Innovations," *Journal of American Statistical Association*, Vol. 90, No. 1, 1995, pp. 1289-1300. [doi:10.1080/01621459.1995.10476633](https://doi.org/10.1080/01621459.1995.10476633)
- [10] S. I. Resnick, "Heavy-Tail Phenomena: Probabilistic and Statistical Modeling," Springer, New York, 2007.
- [11] R. A. Davis and S. Resnick, "Limit Theory for Moving Averages of Random Variables with Regularly Varying Tail Probabilities," *Annals of Probability*, Vol. 13, No. 1, 1985, pp. 179-195. [doi:10.1214/aop/1176993074](https://doi.org/10.1214/aop/1176993074)
- [12] O. Kallenberg, "Random Measures," Akademie-Verlag, Berlin, 1976.
- [13] R. A. Davis and S. Resnick, "Limit Theory for the Sample Covariance and Correlation Functions of Moving Averages," *Annals of Statistics*, Vol. 14, No. 1, 1986, pp. 533-558. [doi:10.1214/aos/1176349937](https://doi.org/10.1214/aos/1176349937)