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

TECHNICAL REPORT NO. 1005

May 1992

ESTIMATION FOR AUTOREGRESSIVE PROCESSES WITH POSITIVE INNOVATIONS

by

Paul D. Feigin[1]
and Sidney I. Resnick[2]

[1]Faculty of Industrial Engineering and Management, Technion, IIT, Haifa, Israel. This research was partially supported by the United States-Israel Binational Science Foundation (BSF) Grant No. 89-00323/1.

[2]Partially supported by NSF Grant DMS 9100027 at Cornell University.

Estimation for Autoregressive Processes with Positive Innovations

Paul D. Feigin*
Faculty of Industrial Engineering and Management
TECHNION — Israel Institute of Technology, Haifa 32000, ISRAEL

Sidney I. Resnick*,†
CORNELL University, Ithaca, New York 14853-7501

May, 1992

Abstract

We consider stationary autoregressive processes of order p which have positive parameters and *positive* innovations. The main results concern the rate of consistency of parameter estimators for the case $p = 2$. These estimators are defined in terms of estimating equations. Relevant asymptotic theory is developed in the wider context of vector autoregressive processes with positive innovations having a distribution with regularly varying left or right tails. These weak convergence results may be of independent interest.

Key words and phrases: autoregressive processes, estimating equations, consistency, weak convergence of point processes, generalized martingales.

AMS 1980 subject classification: Primary — 62M10,62F12; Secondary — 60G55, 60G48, 60F17.

This paper is dedicated to the memory of Micha Yadin, our colleague and dear friend.

---

*This research was partially supported by the United States-Israel Binational Science Foundation (BSF) Grant No. 89-00323/1

†Partially supported by NSF Grant DMS 9100027 at Cornell University

1 Introduction

We consider the *autoregressive process of order p*, denoted by AR(p), with positive innovations, and with positive autoregressive coefficients. These processes are defined by the following relation:

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + Z_t \ ; \ t = 0, \pm 1, \pm 2, \dots \quad ; \tag{1.1}$$

where we assume that $\{Z_t\}$ is an independent and identically distributed sequence of random variables with left endpoint of their common distribution being 0. We will make more specific assumptions in Section 2 when we present our main results.

Based on observation of $\{X_0, X_1, \dots, X_n\}$ we are interested in estimating the parameters, and in determining the asymptotic properties of these estimators as $n \to \infty$.

The case of $p = 1$ was discussed by Davis and McCormick(1989). They used a Poisson Random Measure (PRM) approach to obtain the asymptotic distribution of the natural estimator in the positive innovation context when the innovations distribution, F, varies regularly at 0 and satisfies a suitable moment condition. This estimator is

$$\hat{\phi}^{(n)} = \bigwedge_{j=1}^{n} \frac{X_j}{X_{j-1}} \quad . \tag{1.2}$$

where $\wedge$ denotes the minimum operator.

For the case of $p > 1$, a straightforward generalization of (1.2) does not seem to perform well–Andel (1989). Of course, one may ignore the special nature of the innovations and use the Yule-Walker estimators (see, for example, Brockwell and Davis(1991)), but we show that one can sometimes do better than the $n^{1/2}$ rate of convergence.

Andel (1989) considered the case $p = 2$ and suggested two estimators of (ϕ_1, ϕ_2). One is based on a maximum likelihood argument and is the one we consider here. This estimator is obtained by solving equations which turn out to be examples of *generalized martingale estimating equations* as described in Feigin(1991). Andel (1989) found by simulation that this estimator converges at a faster rate than the Yule-Walker estimator.

Our main results are presented in Section 2. They establish the *rate of consistency* of the estimators of ϕ_1 and ϕ_2 for the case $p = 2$ and so explain some of Andel's findings. Also, for the case $p = 1$, the asymptotic distribution result of Davis and McCormick(1989) is extended by showing that the suitably normalized estimator also has a limit distribution when the distribution, F, of the innovations is regularly varying at ∞, and satisfies an inverse moment condition.

In Section 3 the relevant weak convergence results for the AR(p) model are developed in somewhat more generality than is immediately required for the $p = 2$ case. The proofs use similar techniques to those of Davis and McCormick(1989), but extend the results in two ways: (a) the vector autoregressive process is considered; and (b) they also allow the roles of the conditions on the left and right hand tail of F to be interchanged. These results form the basis of the proofs of the theorems which are to be found in Section 4.

2 Main Results

After suggesting and motivating the estimators to be considered, we set out the conditions under which the consistency results will hold. The model (1.1) is assumed in all that follows.

2.1 Estimating Equations ($p = 2$)

We define subsets of the time indices:

$$T_n(1) = \{t : 2 \leq t \leq n; X_{t-1} > (1+\delta)X_{t-2}\} \quad (2.1)$$
$$T_n(2) = \{t : 2 \leq t \leq n; X_{t-2} > (1+\delta)X_{t-1}\} \quad (2.2)$$

where δ is a sufficiently small, but otherwise arbitrary, positive number.

Now define $\hat{\boldsymbol{\phi}}^{(n)} = (\hat{\phi}_1^{(n)}, \hat{\phi}_2^{(n)})^T$ to be the solution of

$$M_k^{(n)}(\boldsymbol{\phi}) \equiv \bigwedge_{t \in T_n(k)} (X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}) = 0 \ ; \ k = 1, 2 \quad . \quad (2.3)$$

As long as the sets $T_n(k)$ are not empty, there will be a unique solution of (2.3).

One motivation for these two estimating equations comes from the maximum likelihood analysis for the case in which $\Pr(Z_t > x) = \exp(-x)$. In this case, conditionally on $\{X_0 = x_0, X_{-1} = x_{-1}\}$, the likelihood is proportional to (using $\mathrm{I}(\cdot)$ to indicate the indicator function):

$$\mathrm{I}\left(\bigwedge_{t=1}^{n} (X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}) > 0\right) \exp\left[\phi_1 \sum_1^n X_{t-1} + \phi_2 \sum_1^n X_{t-2}\right] \quad (2.4)$$

and the corresponding maximum likelihood estimator will be approximately determined by solving the linear program (LP)

$$\max(\phi_1 + \phi_2) \quad (2.5)$$

subject to

$$\phi_1 X_{t-1} + \phi_2 X_{t-2} \leq X_t \; ; \; t = 1, \ldots, n \quad . \tag{2.6}$$

Note that the fact that $\sum_1^n X_{t-1} / \sum_1^n X_{t-2} \approx 1$ in the stationary case justifies the simplified (approximate) form of the objective function.

Now any bounded solution of (2.5, 2.6) must be at the intersection of two lines of the form

$$\phi_1 X_{t-1} + \phi_2 X_{t-2} = X_t \; ; \tag{2.7}$$

one with $X_{t-1} > X_{t-2}$ and one with $X_{t-1} < X_{t-2}$. (The solution is unbounded if one of these inequalities never holds.) From these considerations we arrive at the equations (2.3), except that we only consider a subset of the lines defined by the value of δ in (2.1, 2.2). In our proofs we require that $\delta > 0$. Since δ is arbitrarily small this restriction is not of great practical importance. We are currently investigating whether letting $\delta = 0$ could change the rate of consistency of the resultant estimators.

Another motivation for considering the estimating equations (2.3) comes from regarding them as *generalized* martingale estimating equations. The development for general p is as follows.

We denote by $Z_t(\boldsymbol{\phi})$ the following:

$$Z_t(\boldsymbol{\phi}) = X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} \tag{2.8}$$

Note that in (1.1) $Z_t = Z_t(\boldsymbol{\phi}^{(0)})$ if $\boldsymbol{\phi}^{(0)}$ is the true value of $\boldsymbol{\phi}$.

We now employ some of the notation and definitions of Feigin(1991) concerning *generalized* martingales. We consider the process $\{Z_t(\boldsymbol{\phi})\}$ as taking values in the semigroup $(\mathbf{R}, \wedge)$. In this space, we define the *generalized* expectation ($\mathcal{E}$) as the *left endpoint* of the distribution of the argument. In this context:

$$\mathcal{E}^{\boldsymbol{\phi}} Z_t(\boldsymbol{\phi}) = 0 \tag{2.9}$$

since we have assumed that the left endpoint of the distribution of Z_t is 0. The notation $\mathcal{E}^{\boldsymbol{\phi}}$ designates the generalized expectation when the true value of the parameter vector is $\boldsymbol{\phi}$.

In the semigroup $(\mathbf{R}, \wedge)$ the martingales are defined by the successive minima of *constant* $\mathcal{E}$-expectation "increments". In other words,

$$M_0^{(n)}(\boldsymbol{\phi}) \equiv \bigwedge_{t=1}^{n} Z_t(\boldsymbol{\phi}) \tag{2.10}$$

is a martingale with increment expectation (according to $\mathcal{E}^{\phi}$) equal to zero. A martingale equation for estimating ϕ is obtained by setting the martingale to its "expected" value:

$$M_0^{(n)}(\phi) = 0 \quad . \tag{2.11}$$

Using martingale transforms we seek alternative martingales that will generate more estimating equations. They can be obtained by adjusting the "t-th increment" of $M_0^{(n)}(\phi)$ by past ($\mathcal{F}_{t-1}$) information. The equations (2.3) for $M_k^{(n)}$ are obtained by defining the "increments" $U_t^{(k)}$ in place of $Z_t(\phi)$ as follows:

$$U_t^{(k)} = \begin{cases} Z_t & \text{if } X_{t-k} > (1+\delta)X_{t-3+k} \\ \infty & \text{otherwise} \end{cases} \quad \text{for } k = 1, 2 \quad . \tag{2.12}$$

In the analogy for the case of the AR(p) process with (ordinary) zero mean innovations the least squares estimating equations are:

$$Q_k^{(n)}(\phi) \equiv \sum_{t=1}^{n} X_{t-k} Z_t(\phi) = 0 \; ; \; k = 1, \ldots, p \quad . \tag{2.13}$$

We note that each of the $\{Q_k^{(n)}(\phi)\}$ are (ordinary) martingale sequences under $\mathcal{P}^{\phi}$, the measure induced by the process $\{X_t\}$ when ϕ is the value of the vector of parameters in (1.1). The equation (2.11), based on $M_0^{(n)}$, is therefore the analogue of the single equation

$$Q_0^{(n)}(\phi) = \sum_{t=1}^{(n)} Z_t(\phi) = 0 \quad , \tag{2.14}$$

in the ordinary setting. To obtain the martingales in (2.13) martingale transforms have been used. The increments of $Q_0^{(n)}$ have also been adjusted (multiplied by X_{t-k}) by past ($\mathcal{F}_{t-1}$) information.

2.2 Conditions

We now set out the formal conditions under which the consistency results hold. Although our main theorems only deal with the case $p = 2$, since much of the asymptotic theory is developed for general p we also state the conditions for general p.

Condition M (model) The process $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ satisfies the equations (1.1) where $\{Z_t\}$ is an independent and identically distributed sequence of random variables with essential infimum (left endpoint) equal to 0.

Condition S (stationarity) The coefficients $\phi_1, \ldots, \phi_p$ are non-negative and satisfy the stationarity condition: $\Phi(z) \equiv 1 - \sum_1^p \phi_i z^i$ has no roots in the unit disk.

Condition L (left tail) The distribution F of the innovations Z_t satisfies, for some $\alpha > 0$:

1. $\lim_{s \to 0} \frac{F(sx)}{F(s)} = x^\alpha$ for all $x > 0$;
2. $\mathrm{E}(Z_t^\beta) = \int_0^\infty u^\beta F(du) < \infty$ for some $\beta > \alpha$.

Condition R (right tail) The right tail of the distribution F of the innovations Z_t satisfies, for some $\alpha > 0$:

1. $\lim_{s \to \infty} \frac{1-F(sx)}{1-F(s)} = x^{-\alpha}$ for all $x > 0$;
2. $\mathrm{E}(Z_t^{-\beta}) = \int_0^\infty u^{-\beta} F(du) < \infty$ for some $\beta > \alpha$.

Some preliminary remarks concerning these conditions may help clarify their importance.

Remark 1 We consider only the case of non-negative autoregressive coefficients, corresponding to the first case discussed in Davis and McCormick(1989). Another approach, possibly restricting Z_t to having bounded support, is required for the more general case.

The stationarity condition together with the non-negativity of the autoregressive coefficients leads to the condition that

$$\phi_1 + \cdots + \phi_p < 1 \quad . \tag{2.15}$$

This fact can quite easily be verified by contradiction by considering the process

$$m_t = \bigwedge_{j=1}^{p} X_{t-j} \; ; \; t = 1, 2, \ldots . \tag{2.16}$$

It is monotonically increasing if $\phi_i > 0$; $i = 1, \ldots, p$ and $\phi_1 + \cdots + \phi_p \geq 1$.

Remark 2 In Section 3, we will need the left or right tail conditions to show that for an appropriate sequence $q_n \sim L(n)n^{1/\alpha}$ where L is slowly varying, we have weak convergence of $q_n \bigwedge_{t \in T_n(k)} Z_t / X_{t-k}$. The moment conditions are required in order to ensure that the limit distribution is determined by the regularly varying tail, and not by the other one. If one were not seeking to prove weak convergence results, but solely bounds on rates of convergence, then one could possibly do away with these extra conditions. At the moment our results are based on weak convergence of PRM's, and so we need these restrictions. Moreover, our hope is to eventually derive asymptotic distributions for the estimates, and so we present our results with this goal in mind.

Remark 3 Condition L is satisfied if a density f of F exists which is continuous at 0 and with $f(0) > 0$. In this case $\alpha = 1$. Another common case where Condition L holds is for the Weibull distributions of the form $F(x) = 1 - \exp\{-x^{\alpha}\}$.

Remark 4 What does condition R mean when the tail of F near 0 is also regularly varying? We have

$$EZ_1^{-\beta} = \int_0^{\infty} \Pr[Z_1^{-\beta} > x]dx \tag{2.17}$$

$$= \int_0^{\infty} F(u)u^{-\beta-1}du \tag{2.18}$$

so if F is regularly varying near 0, the index being greater than β is sufficient to guarantee Condition R. In other words, if both tails are regularly varying, then the rate of convergence is determined by the tail with the smaller value index! If both tails have the same index, determining the weak convergence theory becomes more delicate.

2.3 Theorems

We now state the two main theorems. We denote by $\hat{\boldsymbol{\phi}}^{(n)}$ a solution of (2.3) and by $\boldsymbol{\phi}^0$ the true value of the parameter vector.

The first theorem deals with the rate of consistency of the estimators under the left tail or right tail regular variation.

Theorem 1 *Under conditions M,S, together with L or R, for the case $p = 2$, the solution $\hat{\boldsymbol{\phi}}^{(n)}$ of (2.3),for some choice of δ small enough, satisfies:*

$$L(n)n^{1/\alpha}(\hat{\boldsymbol{\phi}}^{(n)} - \boldsymbol{\phi}^{(0)}) = O_p(1) \quad ; \tag{2.19}$$

where $L(n)$ is a slowly varying function at ∞.

For $\alpha < 2$ this theorem shows that we can achieve a rate of consistency better than $n^{1/2}$ which characterizes the Yule-Walker estimates. The quantity δ that appears in the statement of the theorem determines the estimating equations (2.3) through the definitions (2.1,2.2).

The second theorem deals with the case for $p = 1$ and shows that an exact limit theorem is available when a right tail regular variation condition is available. The corresponding left tail version is the original theorem proved by Davis and McCormick(1989).

Theorem 2 *Under conditions M,S and R the estimator* $\hat{\phi}^{(n)}$ *of (1.2) satisfies*

$$\Pr\left(b_n(\hat{\phi}^{(n)} - \phi^{(0)}) > x\right) \to \exp\{-cx^\alpha\} \tag{2.20}$$

where

$$b_n = \left(\frac{1}{1-F}\right)^{\leftarrow}(n) \tag{2.21}$$

and

$$c = \int_0^\infty \left(1 - \prod_{l=0}^{\infty}(1 - F(\phi^l s))\right) \alpha s^{-\alpha-1} ds. \tag{2.22}$$

3 Asymptotic Theory

In sections 3.1 and 3.2 we discuss limit theory necessary for the understanding of the asymptotic behavior of our estimators of the autoregressive parameters. In Section 3.1 we assume the innovation variables of the autoregression have distributions concentrating on $[0, \infty)$ with regularly varying left tail and right tail controlled by a moment condition (Condition L). In Section 3.2 we consider the reverse situation, namely that the right tail is regularly varying and the left tail is controlled by a moment condition (Condition R). Joint limit distributions are obtained for vectors built from minima of ratios of the X's.

3.1 Left tail analysis in the multivariate case

Consider a first order d-dimensional vector autoregression of the form

$$\boldsymbol{X}_t = \Phi \boldsymbol{X}_{t-1} + \boldsymbol{Z}_t. \tag{3.1}$$

Here, for each t, $\boldsymbol{X}_t$ is a d-dimensional column vector, $\boldsymbol{Z}_t$ is d-dimensional with positive components and $\{\boldsymbol{Z}_t\}$ is iid and Φ is a $d \times d$ matrix of non-negative coefficients. We assume the distribution of $\boldsymbol{Z}_t$ is regularly varying near 0 so that there exists a Radon measure ν on $[0, \infty)^d$ such that for $x \in [0, \infty)^d$

$$\lim_{t \to 0} \frac{\Pr[\boldsymbol{Z}_1 \le tx]}{\Pr[\boldsymbol{Z}_1 \le t\mathbf{1}]} = \nu([0, x]) \tag{3.2}$$

(cf. Resnick, 1987, chapter 5). Here $\mathbf{1}$ is a column of ones of length d. The measure ν has the following homogeneity property: for $t > 0$

$$\nu(t\cdot) = t^\alpha \nu(\cdot) \tag{3.3}$$

for some $\alpha > 0$. Equivalently there exists a regularly varying sequence a_n with index $-1/\alpha$ such that as $n \to \infty$

$$n \Pr[a_n^{-1}\boldsymbol{Z}_1 \in \cdot] \xrightarrow{v} \nu(\cdot) \tag{3.4}$$

in the sense of vague convergence of measures on $[0,\infty)^d$.

We also need a condition controlling the right tails of the $\boldsymbol{Z}$'s and we assume there exists $\beta > \alpha$ such that for $j = 1, \ldots, d$

$$\mathrm{E}Z_{1,j}^{\beta} < \infty \quad . \tag{3.5}$$

If the Perron-Frobenius eigenvalue of Φ is less than 1, there is no trouble showing that the unique solution of (3.1) is given by the convergent series

$$\boldsymbol{X}_t = \sum_{j=0}^{\infty} \Phi^j \boldsymbol{Z}_{t-j}. \tag{3.6}$$

We now explore some limit theory useful for estimation of autoregressive coefficients. From (3.4) we get an equivalent statement about weak convergence of point processes, namely

$$\sum_{t=1}^{n} \epsilon_{\boldsymbol{Z}_t/a_n} \Rightarrow \sum_k \epsilon_{j_k} \tag{3.7}$$

in $[0,\infty)^d$ and where the limit is Poisson with mean measure ν (Resnick, 1987, page 154). The following sequence of steps parallels the procedure given in Davis and McCormick (1989). Define the approximations

$$\boldsymbol{X}_t^{(q)} = \sum_{j=0}^{q} \Phi^j \boldsymbol{Z}_{t-j} \tag{3.8}$$

and then an argument involving m-dependence shows that

$$\sum_{t=1}^{n} \epsilon_{(a_n^{-1}\boldsymbol{Z}_t, \boldsymbol{X}_{t-1}^{(q)})} \Rightarrow \sum_k \epsilon_{(j_k, \boldsymbol{Y}_k^{(q)})} \tag{3.9}$$

in $M_p([0,\infty)^{2d})$, the space of point measures on $[0,\infty)^{2d}$, where $\{\boldsymbol{Y}_k^{(q)}\}$ are iid and independent of $\{j_k\}$ and $\boldsymbol{Y}_1^{(q)} \stackrel{d}{=} \boldsymbol{X}_1^{(q)}$. We wish to show we can replace q by ∞. As $q \to \infty$ we get on the right side of (3.9)

$$\sum_k \epsilon_{(j_k, \boldsymbol{Y}_k^{(q)})} \Rightarrow \sum_k \epsilon_{(j_k, \boldsymbol{Y}_k)} \tag{3.10}$$

where $\{\boldsymbol{Y}_k\}$ is ¡iid and independent of $\{j_k\}$ and $\boldsymbol{Y}_1 = \boldsymbol{X}_1$. It remains to show, according to Billingsley, 1968, Theorem 4.2, that for any $f \in C_K([0,\infty)^{2d})$, the space of continuous functions with compact support on $[0,\infty)^{2d}$, that for any $\eta > 0$

$$\lim_{q\to\infty} \limsup_{n\to\infty} \Pr[|\sum_{t=1}^{n} f(a_n^{-1}\boldsymbol{Z}_t, \boldsymbol{X}_{t-1}^{(q)}) - \sum_{t=1}^{n} f(a_n^{-1}\boldsymbol{Z}_t, \boldsymbol{X}_{t-1})| > \eta] = 0. \tag{3.11}$$

This is accomplished as on page 242 of Davis and McCormick, 1989. So we have the following.

Proposition 3.1 *Suppose (3.1)–(3.2) hold and that the series in (3.6) is convergent. Then in $M_p([0,\infty)^{2d})$ we have as $n \to \infty$*

$$\sum_{t=1}^{n} \epsilon_{(a_n^{-1}\boldsymbol{Z}_t, \boldsymbol{X}_{t-1})} \Rightarrow \sum_k \epsilon_{(j_k, \boldsymbol{Y}_k)} \quad . \tag{3.12}$$

The particular case we are interested in is the AR(p) process of (1.1). We may rewrite this model in vector form by defining

$$\boldsymbol{X}_t = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix}, \quad \boldsymbol{Z}_t = \begin{pmatrix} Z_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.13}$$

and

$$\Phi = \begin{pmatrix} \phi_1 & & \cdots & & \phi_p \\ 1 & 0 & & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{3.14}$$

Set $d = p$ and suppose that conditions M,S, and L hold. Then (3.2–3.4) hold with

$$\nu(dx_1, \ldots, dx_p) = \epsilon_0(dx_2) \cdots \epsilon_0(dx_p) \alpha x_1^{\alpha-1} dx_1. \tag{3.15}$$

In this setting, Proposition 3.1 informs us that in $M_p([0,\infty]^{2p})$

$$\sum_{t=1}^{n} \epsilon_{(\boldsymbol{Z}_t/a_n, \boldsymbol{X}_{t-1})} \Rightarrow \sum_k \epsilon_{(\boldsymbol{j}_k, \boldsymbol{Y}_k)} \tag{3.16}$$

where $a_n = F^{\leftarrow}(1/n)$ and where the limit is Poisson with mean measure $\nu \times \Pr[\boldsymbol{X}_1 \in \cdot]$ where now $\boldsymbol{X}_1 = (X_1, X_0, \ldots, X_{2-p})$. In terms of the original one-dimensional variables, we extract from this statement that

$$\sum_{t=1}^{n} \epsilon_{(Z_t/a_n, X_{t-1}, \ldots, X_{t-p})} \Rightarrow \sum_k \epsilon_{(j_k, \boldsymbol{Y}_k)} \tag{3.17}$$

in $M_p([0,\infty)^{p+1})$, where the vectors $\{\boldsymbol{Y}_k\}$ are iid with $\boldsymbol{Y}_1 \stackrel{d}{=} (X_{t-1}, \ldots, X_{t-p})$ and $\{\boldsymbol{Y}_k\}$ is independent of $\{j_k\}$. The limit is Poisson with mean measure $\alpha x^{\alpha-1} dx \Pr[\boldsymbol{Y}_1 \in \cdot]$.

We may elaborate the argument in Davis and McCormick (1989), page 242 to obtain

$$\sum_{t=1}^{n} \epsilon_{(Z_t/(a_n X_{t-i}), 1 \le i \le p; X_{t-1}, \ldots, X_{t-p})} \Rightarrow \sum_k \epsilon_{(j_k/Y_{ki}, 1 \le i \le p; \boldsymbol{Y}_k)} \tag{3.18}$$

in $M_p([0,\infty)^{2p})$.

Now define for any $\delta > 0$

$$T_n(k) = \{1 \le t \le n : X_{t-k} > (1+\delta) \vee_{l=1, l\neq k}^{p} X_{t-l}\} \quad ; \tag{3.19}$$

as was done for the case $p = 2$ in (2.1,2.2). Then since

$$\Pr[X_t = (1+\delta) \vee_{l=1, l\neq i}^{p} X_{t-l}] = 0, \quad i = 1, \ldots, p \tag{3.20}$$

(which follows from the fact that the distribution of X_t is continuous; see Proposition 2.1 in Davis and Rosenblatt, 1991) we get by restricting the state space to

$$\{(z_1, \ldots, z_p, x_{-1}, \ldots, x_{-p}) : x_{-i} > (1+\delta) \vee_{l=1, l\neq i}^{p} x_{-l}, z_j \in [0,\infty), j \neq i\} \tag{3.21}$$

that in $(M_p([0,\infty)^{p+1}))^p$

$$\left(\sum_{t \in T_n(i)} \epsilon_{(Z_t/(a_n X_{t-i}); X_{t-1}, \ldots, X_{t-p})}, 1 \le i \le p\right) \Rightarrow \left(\sum_k 1_{[Y_{ki} > (1+\delta)\vee_{l=1,l\neq i}^{p} Y_{kl}]} \epsilon_{(j_k/Y_{ki}, \boldsymbol{Y}_k)}, 1 \le i \le p\right). \tag{3.22}$$

The p point processes on the right side of (3.22) are independent because they are restrictions of a Poisson process to p disjoint regions of the state space. Define

$$\{i\}^> = \{(y_{-1}, \ldots, y_{-p}) : y_{-i} > (1+\delta) \vee_{l=1, l\neq i}^{p} y_{-l}\}. \tag{3.23}$$

The following proposition now follows by applying the functional which extracts the minimum of the first components of the points from each point process.

Proposition 3.2 *Suppose conditions M,S, and L hold. Then in* $[0,\infty)^p$ *we have*

$$\left(\bigwedge_{t \in T_n(i)} \frac{Z_t}{a_n X_{t-i}}, 1 \le i \le p\right) \Rightarrow \left(\bigwedge \{\frac{j_k}{Y_{ki}} : Y_{ki} > (1+\delta) \vee_{l=1, l\neq i}^{p} Y_{kl}\}, 1 \le i \le p\right) \tag{3.24}$$

and the components of the limit are independent. The distribution of the ith component of the limit is of Weibull type

$$\Pr[\bigwedge_k \{\frac{j_k}{Y_{ki}} : Y_{ki} > (1+\delta) \vee_{l=1, l\neq i}^{p} Y_{kl}\} > x] = \exp\{-x^\alpha \int_{\{i\}^>} y_i^\alpha \Pr[\boldsymbol{Y}_1 \in d\boldsymbol{y}]\}. \tag{3.25}$$

Proof: Note

$$\sum_k 1_{[Y_{ki} > (1+\delta)\vee_{l=1,l\neq i}^{p} Y_{kl}]} \epsilon_{(j_k, \boldsymbol{Y}_k)} \tag{3.26}$$

is Poisson with mean measure

$$\alpha x^{\alpha-1}dx\,\Pr[\boldsymbol{Y}_1 \in (\cdot) \cap \{(y_{-1},\ldots,y_{-p}) : y_{-i} > (1+\delta)\vee_{l=1,l\neq i}^{p} y_{-l}\}] \quad . \tag{3.27}$$

We are now in a position to compute the limit distribution above:

$$\Pr[\bigwedge_k\{\frac{j_k}{Y_{ki}} : Y_{ki} > (1+\delta)\vee_{l=1,l\neq i}^{p} Y_{kl}\} > x]$$

$$\begin{aligned} &= \exp\{-\int\int_{\{(s,\mathbf{y}):s/y_i\le x\}} \alpha s^{\alpha-1}ds\,\Pr[\boldsymbol{Y}_1 \in d\boldsymbol{y}, \boldsymbol{Y}_1 \in \{i\}^>]\} \\ &= \exp\{-x^{\alpha}\int_{\{i\}^>} y_i^{\alpha}\Pr[\boldsymbol{Y}_1 \in d\boldsymbol{y}]\}. \end{aligned} \tag{3.28}$$

□

3.2 Influence of the right tail

In this section we discuss why the right tail may be important to a decision how to estimate autoregressive coefficients. We present the limit theory which will underlie our approach.

Suppose we have an infinite order moving average process

$$X_t = \sum_{j=0}^{\infty} c_j Z_{t-j}, \quad t = 0, \pm 1, \pm 2, \ldots \tag{3.29}$$

where $\{Z_t\}$ are iid, non-negative with regularly varying tail probabilities as given in Condition R. Define

$$b_n = \left(\frac{1}{1-F}\right)^{\leftarrow}(n) \quad . \tag{3.30}$$

We assume the non-negative sequence $\{c_j\}$ satisfies

$$\sum_{j=0}^{\infty} c_j^{\lambda} < \infty, \quad \text{for some } 0 < \lambda < \alpha \wedge 1. \tag{3.31}$$

This guarantees the infinite series in (3.29) converges and also from Cline (1983)

$$\lim_{t\to\infty}\frac{\Pr[X_1 > t]}{\Pr[Z_1 > t]} = \sum_{j=0}^{\infty} c_j^{\alpha}. \tag{3.32}$$

We seek weak convergence of the sequence of point processes $\sum_{t=1}^{\infty} \epsilon_{(X_t/b_n, Z_{t+1})}$ in the vague topology on $M_p((0,\infty] \times [0,\infty))$, the space of Radon point measures on $(0,\infty] \times [0,\infty)$. We proceed in a series of steps which parallel closely those employed by Davis and Resnick (1985); see also the summary of the Davis and Resnick paper in Resnick (1987), pages 224ff.

First observe that the regular variation Condition R(1) is equivalent to weak convergence of a sequence of point processes

$$\sum_{t=1}^{n} \epsilon_{Z_t/b_n} \Rightarrow \sum_{k} \epsilon_{j_k} \tag{3.33}$$

where the limit point process is Poisson with mean measure $\alpha x^{-\alpha-1}dx,\ x>0$. (Cf. Resnick, 1987, Proposition 3.21.) Fix an integer m and set

$$I_n = \sum_{t=1}^{n} \epsilon_{((Z_{t-1},\dots,Z_{t-m})/b_n, Z_t)} \tag{3.34}$$

so that I_n is a random element of $M_p(([0,\infty]^m \setminus \{\mathbf{o}\}) \times [0,\infty))$. Note that for any $\delta > 0$

$$\begin{aligned} & \mathrm{E} I_n(\cup_{1\le i<j\le m}\{(z_1,\dots,z_m,z_{m+1}) : z_i > \delta, z_j > \delta\}) && (3.35)\\ = & \binom{m}{2} n \Pr[Z_1 > \delta b_n, Z_2 > \delta b_n, Z_3 \ge 0] && (3.36)\\ \sim & \binom{m}{2} \delta^{-\alpha} \Pr[Z_2 > \delta b_n] \to 0 && (3.37) \end{aligned}$$

as $n \to \infty$. Define basis vectors in R^m by $\boldsymbol{e}_l = (0,\dots,1,\dots,0)$, with 1 in the lth spot, $1 \le l \le m$. Define I_n^* by

$$I_n^* = \sum_{t=1}^{n}\sum_{l=1}^{m} \epsilon_{(b_n^{-1}(Z_{t-l}\boldsymbol{e}_l), Z_t)} \tag{3.38}$$

Let ρ be the vague metric on $M_p(([0,\infty]^m \setminus \{\mathbf{o}\}) \times [0,\infty))$. Because of (3.35–3.37) we get, as in the discussion of Resnick, 1987, page 233, that

$$\rho(I_n, I_n^*) \xrightarrow{P} 0 \quad . \tag{3.39}$$

The weak limit behavior of I_n^* is the same as for the sequence

$$I_n^{**} = \sum_{t=1}^{n}\sum_{l=1}^{m} \epsilon_{(b_n^{-1} Z_t \boldsymbol{e}_l, Z_{t+l})} \tag{3.40}$$

(cf. discussion in Resnick, 1987, equation (4.81)). To verify this assertion we can write (set $k = t - l$)

$$\begin{aligned} I_n^* &= \sum_{l=1}^{m}\sum_{k=1}^{n} \epsilon_{(b_n^{-1} Z_k \boldsymbol{e}_l, Z_{k+l})} && (3.41)\\ &\quad + \sum_{l=1}^{m}\sum_{k=1-l}^{0} \epsilon_{(b_n^{-1} Z_k \boldsymbol{e}_l, Z_{k+l})} - \sum_{l=1}^{m}\sum_{k=n-l+1}^{n} \epsilon_{(b_n^{-1} Z_k \boldsymbol{e}_l, Z_{k+l})} && (3.42)\\ &= I_n^{**} + II - III. && (3.43) \end{aligned}$$

It is easy to check that

$$II \xrightarrow{P} 0, \quad III \xrightarrow{P} 0 \tag{3.44}$$

as $n \to \infty$ and this verifies (3.40). Now the sequence of point processes

$$\sum_{t=1}^{n} \epsilon_{(b_n^{-1} Z_t, Z_{t+1}, \ldots, Z_{t+m})}, \quad n \geq 1, \tag{3.45}$$

being based on the stationary $m+1$-dependent sequence $\{(Z_t, \ldots Z_{t+m}), -\infty < t < \infty\}$, is weakly convergent to a Poisson limit

$$\sum_{k} \epsilon_{(j_k, Y_{k1}, \ldots, Y_{km})} \tag{3.46}$$

where $\{Y_{kl}, k \geq 1, 1 \leq l \leq m\}$ are iid random variables with the same distribution as Z_1 and are independent of the $\{j_k\}$ and where the limit Poisson process has mean measure $\alpha x^{-\alpha-1} dx \times \Pr[(Z_1, \ldots, Z_m) \in \cdot]$. By projecting we then get in $(M_p(0,\infty] \times [0,\infty)))^m$ that

$$\left(\sum_{t=1}^{n} \epsilon_{(b_n^{-1} Z_t, Z_{t+1})}, \ldots, \sum_{t=1}^{n} \epsilon_{(b_n^{-1} Z_t, Z_{t+m})}\right) \Rightarrow \left(\sum_{k} \epsilon_{(j_k, Y_{k1})}, \ldots, \sum_{k} \epsilon_{(j_k, Y_{km})}\right) \tag{3.47}$$

and therefore from an easy mapping argument we have

$$\left(\sum_{t=1}^{n} \epsilon_{(b_n^{-1} Z_t \boldsymbol{e}_1, Z_{t+1})}, \ldots, \sum_{t=1}^{n} \epsilon_{(b_n^{-1} Z_t \boldsymbol{e}_m, Z_{t+m})}\right) \Rightarrow \left(\sum_{k} \epsilon_{(j_k \boldsymbol{e}_1, Y_{k1})}, \ldots, \sum_{k} \epsilon_{(j_k \boldsymbol{e}_m, Y_{km})}\right). \tag{3.48}$$

Since addition is vaguely continuous we finally obtain

$$I_n^{**} \Rightarrow \sum_{l=1}^{m} \sum_{k} \epsilon_{(j_k \boldsymbol{e}_l, Y_{kl})} \tag{3.49}$$

whence also

$$I_n \Rightarrow \sum_{l=1}^{m} \sum_{k} \epsilon_{(j_k \boldsymbol{e}_l, Y_{kl})}. \tag{3.50}$$

As on page 235 of Resnick, 1987, apply the mapping

$$(z_0, \ldots, z_{m-1}, z_m) \mapsto (\sum_{j=0}^{m-1} c_j z_j, z_m) \tag{3.51}$$

and Proposition 3.18 of Resnick, 1987 to (3.50) to obtain

$$\sum_{t=0}^{n} \epsilon_{(b_n^{-1} \sum_{j=0}^{m-1} c_j Z_{t-j-1}, Z_t)} \Rightarrow \sum_{k} \sum_{l=0}^{m-1} \epsilon_{(j_k c_l, Y_{k,l+1})}. \tag{3.52}$$

Now one needs an argument justifying replacement of m by ∞. This is almost identical to the one presented on page 237 of Resnick, 1987. We have proved the following proposition.

Proposition 3.3 *Suppose Condition R holds as does (3.31). Then we have, in* $M_p((0,\infty]\times[0,\infty))$, *that*

$$\sum_{t=1}^{n}\epsilon_{(b_n^{-1}X_{t-1},Z_t)}\Rightarrow\sum_k\sum_{l=0}^{\infty}\epsilon_{(j_kc_l,Y_{kl})}\quad; \tag{3.53}$$

and for any positive integer p *we have in* $M_p(([0,\infty]^p\setminus\{\mathbf{o}\})\times[0,\infty))$

$$\sum_{t=1}^{n}\epsilon_{(b_n^{-1}(X_{t-1},X_{t-2},\ldots,X_{t-p}),Z_t)}\Rightarrow\sum_k\sum_{l=0}^{\infty}\epsilon_{(j_k(c_l,c_{l-1},\ldots,c_{l-p+1}),Y_{kl})} \tag{3.54}$$

where for $j<0$ *we set* $c_j=0$, $\{j_k\}$ *are the points of a Poisson process with mean measure* $\alpha x^{-\alpha-1}dx1_{(0,\infty]}(x)$ *and* $\{Y_{kl},k\geq 1,l\geq 0\}$ *are iid random variables with the distribution of* Z_1.

The proof of the second claim in Proposition 3.3 follows as in Theorem 2.4 of Davis and Resnick, 1985. □

We now use the fact that weak convergence holds when restricted to a subset of the state space, provided the limit process has no points on the boundary of the subset with probability 1. We thus conclude that

$$\left\{\sum_{t\in T_n(i)}\epsilon_{b_n^{-1}(X_{t-1},\ldots,X_{t-p}),Z_t)}\;;\;i=1,\ldots,p\right\}\Rightarrow\left\{\sum_k\sum_{l\in C(i)}\epsilon_{(j_k(c_l,\ldots,c_{l-p+1}),Y_{kl})}\;;\;i=1,\ldots,p\right\} \tag{3.55}$$

where

$$C(i)=\{l:c_{l-i+1}>(1+\delta)\vee_{1\leq j\leq p;j\neq i}c_{l-j+1}\}\;;\;\text{for } i=1,\ldots,p\quad. \tag{3.56}$$

For example, if $c_0>c_1>c_2>\cdots$ and $c_i=0$ for $i<0$ and $c_i\geq 0$ for $i>0$, then $C(i)\supseteq\{i-1\}$ for δ small enough. This case is the appropriate one for the AR(1) process of the type satisfying Condition M. For the result (3.55) to follow from (3.54) we require that the set $\{l:c_{l-i+1}=(1+\delta)\vee_{1\leq j\leq p;j\neq i}c_{l-j+1}\}=\emptyset$.

For the purpose of the limit theory for estimation we will need to divide Z_t by $b_n^{-1}X_{t-i}$ in (3.55) in order to prove:

$$\sum_{t\in T_n(i)}\epsilon_{b_nZ_t/X_{t-i}}\Rightarrow\sum_k\sum_{l\in C(i)}\epsilon_{Y_{kl}/(j_kc_{l-i+1})}\quad. \tag{3.57}$$

We now consider how to justify this operation. Define the map

$$T:(0,\infty]\times[0,\infty)\mapsto[0,\infty) \tag{3.58}$$

by

$$T(x,y) = y/x \quad . \tag{3.59}$$

Since $T^{-1}(\{1\})$ is not compact, we note from Proposition 3.18 of Resnick, 1987 that a truncation of the state space will be necessary. The following argument parallels one presented in Davis and McCormick, 1989. Choose M large and restrict the state space to the compact set $[M^{-1},\infty] \times [0,M]$. Then from (3.55) we get that the further restrictions converge:

$$\sum_{t\in T_n(i)} 1_{[b_n^{-1}X_{t-1}>M^{-1},Z_t\leq M]}\epsilon_{(b_n^{-1}X_{t-1},Z_t)} \Rightarrow \sum_k \sum_{l\in C(i)} 1_{[j_k c_{l-i+1}>M^{-1},Y_{kl}\leq M]}\epsilon_{(j_k c_{l-i+1},Y_{kl})} \quad . \tag{3.60}$$

Because the state space is now compact, we may apply T to get

$$\sum_{t\in T_n(i)} 1_{[b_n^{-1}X_{t-1}>M^{-1},Z_t\leq M]}\epsilon_{\frac{Z_t}{X_{t-i}/b_n}} \Rightarrow \sum_k \sum_{l\in C(i)} 1_{[j_k c_{l-i+1}>M^{-1},Y_{kl}\leq M]}\epsilon_{Y_{kl}/(j_k c_{l-i+1})} \quad . \tag{3.61}$$

Letting $M \to \infty$ in the right side of (3.61), yields the right side of (3.57) and by Billingsley, 1968, Theorem 4.2 it suffices to show that for any $f \in C_K([0,\infty))$ and $\eta > 0$

$$\lim_{M\to\infty} \limsup_{n\to\infty} \Pr[|\sum_{t\in T_n(i)} f(\frac{b_n Z_t}{X_{t-i}})1_{[\frac{X_{t-i}}{b_n}>M^{-1},Z_t\leq M]} - \sum_{t\in T_n(i)} f(\frac{b_n Z_t}{X_{t-i}})| > 2\eta] = 0 \quad . \tag{3.62}$$

Suppose the support of f is $[0,c]$ and for simplicity assume $f \leq 1$. The probability on the previous line is bounded by

$$\begin{aligned} \Pr[\sum_{t=1}^n f(\frac{b_n Z_t}{X_{t-i}})1_{[\frac{X_{t-i}}{b_n}\leq M^{-1}]} > \eta] \;&+\; \Pr[\sum_{t=1}^n f(\frac{b_n Z_t}{X_{t-i}})1_{[Z_t>M]} > \eta] && (3.63)\\ &= A + B \quad . && (3.64) \end{aligned}$$

Now B is easily disposed of since

$$\begin{aligned} B &\leq \Pr\{\cup_{t=1}^n [\frac{b_n Z_t}{X_{t-i}} \leq c, Z_t > M]\} && (3.65)\\ &\leq n \Pr[\frac{b_n Z_1}{X_0} \leq c, Z_1 > M] && (3.66)\\ &\leq n \Pr[\frac{b_n M}{X_0} \leq c] && (3.67)\\ &= n \Pr[\frac{X_0}{b_n} \geq \frac{M}{c}] && (3.68) \end{aligned}$$

and from (3.32), as $n \to \infty$ this is (3.69)

$$\sim \text{(const)} \left(\sum_{j=0}^{\infty} c_j^\alpha\right) M^{-\alpha} \to 0 \tag{3.70}$$

as $M \to \infty$.

To kill A we need to assume a bit more, namely that for some $\beta > \alpha$

$$EZ_1^{-\beta} < \infty \quad ; \tag{3.71}$$

in other words, Condition R(2). By an argument similar to the way we bounded B we have

$$\begin{aligned} A &\le n \Pr[\frac{b_n Z_1}{X_0} \le c, \frac{X_0}{b_n} \le M^{-1}] & (3.72)\\ &\le n \Pr[c^{-1} \le \frac{X_0}{b_n Z_1}, Z_1 \le cM^{-1}] & (3.73)\\ &= n \Pr[c^{-1} \le b_n^{-1} X_0 Z_1^{-1} 1_{[Z_1^{-1} > Mc^{-1}]}] \quad . & (3.74) \end{aligned}$$

Since $Z_1^{-1} 1_{[Z_1^{-1} > Mc^{-1}]}$ has a finite β moment, it follows from a result of Breiman, 1965 that as $n \to \infty$ the forgoing is asymptotic to

$$\begin{aligned} &\sim \sum_j c_j^\alpha c^\alpha E Z_1^{-\alpha} 1_{[Z_1^{-1} > Mc^{-1}]} & (3.75)\\ &\to 0 & (3.76) \end{aligned}$$

as $M \to \infty$. This verifies the following statement.

Proposition 3.4 *Suppose Condition R holds and that also (3.31) holds. Then in $M_p([0,\infty))$ we have*

$$\sum_{t \in T_n(i)} \epsilon_{b_n Z_t / X_{t-i}} \Rightarrow \sum_k \sum_{l \in C(i)} \epsilon_{Y_{kl}/(j_k c_{l-i+1})} \quad . \tag{3.77}$$

In fact, for any $p \ge 1$, the following more general statement holds: In $(M_p([0,\infty)))^p$ we have

$$\left\{ \sum_{t \in T_n(i)} \epsilon_{\frac{b_n Z_t}{X_{t-i}}} \; ; \; i = 1, \ldots, p \right\} \Rightarrow \left\{ \sum_k \sum_{l \in C(i)} \epsilon_{Y_{kl}/(j_k c_{l-i+1})} \; ; \; i = 1, \ldots, p \right\} \quad . \tag{3.78}$$

Both (3.77) and (3.78) assume $c_{l-i+1} > 0$ for $l \in C(i)$ — this follows from the definition (3.56).

Proof: Follows from the above with (3.78) being obtained by elaborating slightly the argument leading to (3.77). $\square$

4 Proof of Theorems

In this section we give the proofs of the theorems based on the weak convergence results just established. The following lemma is the required conclusion of those results.

Lemma 1 *Under conditions M,S, and L or R there exists a $\delta > 0$ small enough and a sequence of constants $\{q_n\}$, regularly varying with index $1/\alpha$, for which:*

$$q_n L_k^{(n)} \equiv q_n \bigwedge_{t \in T_n(k)} \frac{Z_t}{X_{t-k}} \Rightarrow W_k \ ; \ k = 1, 2 \tag{4.1}$$

where $\Rightarrow$ denotes weak convergence.

Proof: The quantity δ enters into the definition of $T_n(k)$. The conclusion of the lemma for the case of conditions M,S, and L is Proposition 3.2 with $q_n = 1/a_n$. We need only take care that δ is chosen so that (3.20) is true.

For the case of conditions M,S, and R the result follows from Proposition 3.4, namely (3.77), with $q_n = b_n$. We describe this derivation now.

First we consider properties of the c_j's for the AR(p) case. We set

$$\Phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i \tag{4.2}$$

and then since $\Phi(z)$ has no roots in the unit disk (Condition S)

$$C(z) = \sum_{j=0}^{\infty} c_j z^j = \frac{1}{\Phi(z)} \tag{4.3}$$

determines the sequence $\{c_j\}$ (Brockwell and Davis, 1991). In fact, on writing down the equations for determining the c_j in terms of the ϕ_j, we see that $c_0 = 1$ and $c_j < 1$ for $j \geq 1$ *when Conditions M and S hold.* These results make use of the fact that $\phi_1 + \cdots + \phi_p < 1$, as indicated in Remark 1. Moreover, in the special case that $\{X_t\}$ is an autoregression of order p given by the recursion (1.1) the c_j's will decrease exponentially fast and so (3.31) is assured.

Thus far we have shown that the conditions of Proposition 3.4 hold. In order to interpret the conclusion we need to ascertain $C(i)$. Let us concentrate on the case $p = 2$. It is clear from the definition (3.56) that $i - 1 \in C(i)$, for $i = 1, 2$, if δ is small enough. The latter follows from the fact that $c_0 = 1$ is the largest c_j. Therefore, when we take the minimum functional of the points in the state space on both sides of (3.77) we obtain

$$\bigwedge_{t\in T_n(i)} \frac{b_n Z_t}{X_{t-i}} \Rightarrow \bigwedge_k \bigwedge_{l\in C(i)} \frac{Y_{kl}}{j_k c_{l-i+1}}. \tag{4.4}$$

In fact, for the case of the AR(p) of (1.1), we can conclude from (3.78) that

$$b_n \left(\bigwedge_{t\in T_n(1)} \frac{Z_t}{X_{t-1}}, \ldots, \bigwedge_{t\in T_n(p)} \frac{Z_t}{X_{t-p}} \right) \Rightarrow \left(\bigwedge_k \bigwedge_{l\in C(i)} \frac{Y_{kl}}{j_k c_{l-i+1}}, 1 \le i \le p \right). \tag{4.5}$$

$\square$

Proof of Theorem 1: Let $\hat{\boldsymbol{\phi}}^{(n)}$ be a solution of (2.3) and suppose that the two lines that intersect are given by the indices $t_n(k) \in T_n(k)$; $k = 1, 2$. We assume that δ is chosen sufficiently small so that the sets $T_n(k)$ increase indefinitely as $n \to \infty$. The latter follows from the stationarity assumption (Condition S) which assures that there is a positive probability that $X_{t-1} > X_{t-2}$ and a positive probability that $X_{t-1} < X_{t-2}$.

Writing

$$\boldsymbol{\Delta}^{(n)} = \hat{\boldsymbol{\phi}}^{(n)} - \boldsymbol{\phi}^{(0)} \tag{4.6}$$

we have that $\hat{\boldsymbol{\phi}}^{(n)}$ and $\boldsymbol{\Delta}^{(n)}$ satisfy

$$0 = Z_{t_n(k)}(\hat{\boldsymbol{\phi}}^{(n)}) = Z_{t_n(k)} - \Delta_1^{(n)} X_{t_n(k)-1} - \Delta_2^{(n)} X_{t_n(k)-2} \ ; \ k = 1, 2 \tag{4.7}$$

or

$$\Delta_1^{(n)} + U_1^{(n)} \Delta_2^{(n)} = V_1^{(n)} \tag{4.8}$$

$$U_2^{(n)} \Delta_1^{(n)} + \Delta_2^{(n)} = V_2^{(n)} \tag{4.9}$$

where, for $k = 1, 2$,

$$V_k^{(n)} = \frac{Z_{t_n(k)}}{X_{t_n(k)-k}} \tag{4.10}$$

$$U_k^{(n)} = \frac{X_{t_n(k)-3+k}}{X_{t_n(k)-k}} < \frac{1}{1+\delta} \tag{4.11}$$

Solving (4.8,4.9) we obtain

$$\Delta_1^{(n)} = \frac{(V_1^{(n)} - U_1^{(n)} V_2^{(n)})}{(1 - U_1^{(n)} U_2^{(n)})} \tag{4.12}$$

$$\Delta_2^{(n)} = \frac{(V_2^{(n)} - U_2^{(n)} V_1^{(n)})}{(1 - U_1^{(n)} U_2^{(n)})} \quad . \tag{4.13}$$

From (4.11) we can derive that

$$\frac{\delta}{1+\delta}(V_1^{(n)} + V_2^{(n)}) < \Delta_1^{(n)} + \Delta_2^{(n)} < (V_1^{(n)} + V_2^{(n)}) \quad . \tag{4.14}$$

The first inequality comes directly from studying the solutions (4.12,4.13). Indeed

$$\Delta_1^{(n)} + \Delta_2^{(n)} > \left[V_1^{(n)}(1 - \frac{1}{1+\delta}) + V_2^{(n)}(1 - \frac{1}{1+\delta})\right] /1 = \frac{\delta}{1+\delta}\left[V_1^{(n)} + V_2^{(n)}\right] \quad . \tag{4.15}$$

The second inequality in (4.14) comes from the fact that any vector $(x_1, x_2)^T$ that satisfies

$$\begin{aligned} x_1 + u_1 x_2 &\leq b_1 && (4.16)\\ u_2 x_1 + x_2 &\leq b_2 && (4.17)\\ \text{for} \quad u_1, u_2, b_1, b_2 &\geq 0 && (4.18)\\ \text{and} \qquad u_1, u_2 &\leq 1 && (4.19) \end{aligned}$$

is such that

$$x_1 + x_2 \leq b_1 + b_2 \quad . \tag{4.20}$$

This last result comes from considering the LP to maximize $x_1 + x_2$ subject to (4.16,4.17). (For example, looking at the dual LP gives the result easily.)

Moreover, further simple calculations from (4.12,4.13) will reveal that

$$\begin{aligned} |\Delta_k^{(n)}| &\leq (V_1^{(n)} + V_2^{(n)})/\left[1 - \left(\frac{1}{1+\delta}\right)^2\right] && (4.21)\\ &< \frac{1+\delta}{\delta}(V_1^{(n)} + V_2^{(n)}) \ ; \ k = 1,2 \quad . && (4.22) \end{aligned}$$

We proceed to show that $V_1^{(n)} + V_2^{(n)}$ cannot be too much larger than $L_1^{(n)} + L_2^{(n)}$, where $L_k^{(n)}$ is the minimum of Z_t/X_{t-k} for $t \in T_n(k)$ — see (4.1). Indeed, let $t_n^*(k)$ be the index in $T_n(k)$ for which $L_k^{(n)}$ is achieved, and let, correspondingly,

$$U_k^{*n} = \frac{X_{t_n^*(k)-3+k}}{X_{t_n^*(k)-k}} \ ; \ k = 1,2 \quad . \tag{4.23}$$

For $\hat{\phi}^{(n)}$ to solve (2.3) we require that

$$\begin{aligned} \Delta_1^{(n)} + U_1^{*n}\Delta_2^{(n)} &\leq L_1^{(n)} && (4.24)\\ U_2^{*n}\Delta_1^{(n)} + \Delta_2^{(n)} &\leq L_2^{(n)} \quad . && (4.25) \end{aligned}$$

(Otherwise the value of one or both of the $M_k^{(n)}(\hat{\phi}^{(n)})$ would be negative!) But for $\Delta_1^{(n)}, \Delta_2^{(n)}$ to satisfy these inequalities, just like those in (4.16,4.17), we must have

$$\Delta_1^{(n)} + \Delta_2^{(n)} \leq L_1^{(n)} + L_2^{(n)} \quad . \tag{4.26}$$

This fact together with the first inequality in (4.14) proves that

$$\frac{\delta}{1+\delta}(V_1^{(n)} + V_2^{(n)}) < \Delta_1^{(n)} + \Delta_2^{(n)} < (L_1^{(n)} + L_2^{(n)}) \tag{4.27}$$

and therefore, from (4.22)

$$|\Delta_k^{(n)}| < \left(\frac{1+\delta}{\delta}\right)^2 (L_1^{(n)} + L_2^{(n)}) \; ; \; k = 1,2 \quad . \tag{4.28}$$

However from Lemma 1, $q_n L_k^{(n)}$ coverges weakly and so we conclude that

$$q_n |\Delta_k^{(n)}| = O_p(1) \; ; \; k = 1,2 \quad . \tag{4.29}$$

This completes the proof of Theorem 1. $\square$

Proof of Theorem 2: This proof comes again from Proposition 3.4. Since the convergence in (3.77) is in $M_p([0,\infty))$, the mapping from the point measures to the minima of the points is an a.s. continuous functional (cf. Resnick, 1987, page 214 for related discussion) and hence we get

$$b_n \bigwedge_{t=1}^{n} \frac{Z_t}{X_{t-1}} \Rightarrow \bigwedge_{k,l} \frac{Y_{kl}}{c_l j_k} \tag{4.30}$$

Based on the fact that $\sum_k \epsilon_{j_k}$ is a Poisson process with mean measure $\alpha x^{-\alpha-1}$ and $\{Y_{kl}\}$ is iid and independent of the Poisson process, we can easily compute the distribution of the limit in (4.30): For any $x > 0$ we have

$$\begin{aligned}
\Pr[\bigwedge_{k,l} \frac{Y_{kl}}{c_l j_k} > x] &= E \Pr[\frac{Y_{kl}}{c_l j_k} > x, \quad \text{for all } \; k,l | \{j_n\}] && (4.31)\\
&= E \prod_{k,l} \Pr[Y_{kl} > c_l j_k x | \{j_n\}] && (4.32)\\
&= E \prod_{k,l} (1 - F(c_l j_k x)) && (4.33)\\
&= E \prod_{k} g(x j_k) && (4.34)
\end{aligned}$$

where we have defined a new function g by

$$g(u) = \prod_{l=0}^{\infty} (1 - F(c_l u)).$$

Now write

$$E\prod_k g(xj_k) = Ee^{-\sum_{l=0}^{\infty}(-\log g(xj_k))}$$

which is the Laplace functional of the Poisson process with points $\{j_k\}$ at the function $-\log g(x\cdot)$ and therefore

$$\begin{aligned} E\prod_k g(xj_k) &= \exp\{-\int_0^\infty (1 - e^{-(-\log g(xu))})\alpha u^{-\alpha-1}du\} && (4.35) \\ &= \exp\{-\int_0^\infty (1 - g(xu))\alpha u^{-\alpha-1}du\} && (4.36) \\ &= \exp\{-cx^\alpha\} && (4.37) \end{aligned}$$

where

$$c = \int_0^\infty (1 - g(xu))\alpha u^{-\alpha-1}du = \int_0^\infty \left(1 - \prod_{l=0}^{\infty}(1 - F(c_l s))\right) \alpha s^{-\alpha-1}ds.$$

Recall in the AR(1) case that $c_l = \phi^l$. The gives us exactly the required result since

$$\hat{\phi}^{(n)} - \phi^{(0)} = \bigwedge_t \frac{Z_t}{X_{t-1}} \quad . \qquad (4.38)$$

$\square$

References

[1] Andel, J. (1989), Non-negative autoregressive processes *J. Time Series Analysis* **10**, 1–11

[2] Breiman, L. (1965), On some limit theorems similar to the arc-sin law *Theor. Probab. Appl.* **10**, 323–331

[3] Brockwell, P. and Davis, R. (1991), *Time Series: Theory and Methods*, 2nd edition Springer–Verlag , New York

[4] Cline, D. (1983), Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data, Thesis, Dept of Statistics, Colorado State University, Ft. Collins CO 80521 USA

[5] Davis, R. and McCormick, W. (1989), Estimation for first-order autoregressive processes with positive or bounded innovations, *Stochastic Processes and their Applic.* **31**, 237–250

[6] Davis, R. and Resnick, S. (1985), Limit theory for moving averages of random variables with regularly varying tail probabilities, *Ann. Probability* **13** , 179–195.

[7] Davis, R. and Resnick, S. (1988), Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution, *Stochastic Processes and their Applications* **30**, 41–68.

[8] Rosenblatt, M. and Davis, R. (1991), Parameter estimation for some time series models without continuity, *Statistics and Probability Letters* **11**, 515–521.

[9] Feigin, P. (1991), Inference and martingale estimating equations for stochastic processes on a semigroup, Preprint

[10] Resnick, Sidney (1987), *Extreme Values, Regular Variation, and Point Processes*, Springer-Verlag, New York.