PARAMETER ESTIMATION FOR MOVING AVERAGES WITH POSITIVE INNOVATIONS

Paul D. Feigin, Marie F. Kratz and Sidney I. Resnick

Technion, Israel Institute of Technology, Université René Descartes, Paris V and Cornell University

January 1994

Abstract. This paper continues the study of time series models generated by non-negative innovations which was begun in Feigin and Resnick (1992,1994). We concentrate on moving average processes. Estimators for moving average coefficients are proposed and consistency and asymptotic distributions established for the case of an order one moving average assuming either the right or left tail of the innovation distribution is regularly varying. The rate of convergence can be superior to that of the Yule–Walker or maximum likelihood estimators.

1. Introduction.

This paper continues the study of time series models generated by non-negative innovations which was begun in Feigin and Resnick (1992,1994). This program is motivated by the need to model teletraffic and hydrologic data sets where quantities such as holding times and stream flows are inherently positive and hence possibly unsuited to the usual time series methods which are based on Gaussian models. In Feigin and Resnick (1994), we showed how to estimate parameters of a pure autoregression using linear programming (lp) techniques. Such lp estimators have a good rate of convergence which is frequently superior to those achieved by Yule Walker or maximum likelihood estimators. Such estimators can be used for model selection and for testing for independence (Feigin, Resnick and Starica, 1994). In this paper, we focus on estimation of moving average coefficients. This is a necessary step along the road to being able to estimate parameters in more general ARMA processes which combine both autoregressive and moving average components.

The process under consideration is the finite order moving average of order q, denoted MA(q) and specified as follows: Let $\{Z_t\}$ be an iid sequence of non-negative random variables. For a positive integer $q \geq 1$, suppose we have parameters $\theta_1, \dots, \theta_q$ such that $\theta_i \geq 0$ for $1 \leq i \leq q$. The MA(q) process $\{X_t\}$ is

$$X_t = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i}, \quad t = 0, \pm 1, \pm 2, \dots \tag{1.1}$$

and we are interested in estimating $\theta_1, \dots, \theta_q$. n It is convenient to be able to write (1.1) compactly and to achieve this we define the moving average polynomial

$$\Theta(z) = \sum_{i=0}^{q} \theta_i z^i,$$

Key words and phrases. Poisson processes, linear programming, autoregressive processes, moving average processes, parameter estimation, weak convergence, consistency, time series analysis.

Paul Feigin and Sidney Resnick were partially supported by the United States-Israel Binational Science Foundation (BSF) Grant No. 89-00323/1. Sidney Resnick was also partially supported by NSF Grant DMS-9100027 at Cornell University, which also supported the travel to Cornell by Marie Kratz. Marie Kratz acknowledges with gratitude the hospitality of Cornell's School of Operations Research and Industrial Engineering during fall 1993 .

where $\theta_0 = 1$ and the backward shift operator B is defined symbolically by

$$BX_t = X_{t-1}, \quad BZ_t = Z_{t-1}.$$

With this notation we may write the MA(q) as

$$X_t = \Theta(B)Z_t, \quad t = 0, \pm 1, \pm 2, \ldots.$$

For a pure autoregressive process of order p, denoted by AR(p), with positive innovations $\{Z_t\}$, and with autoregressive coefficients $\phi_1, \ldots, \phi_p$, $(\phi_p \neq 0, \sum_{i=1}^{p} \phi_i < 1)$, of the form

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + Z_t \ ; \ t = 0, \pm 1, \pm 2, \ldots \tag{1.2}$$

Feigin and Resnick (1994) defined the linear programming estimators $\hat{\boldsymbol{\phi}}$ based on observing $X_1, \ldots, X_n$ as

$$\hat{\boldsymbol{\phi}} = \arg \max_{\boldsymbol{\delta} \in D_n} \boldsymbol{\delta}' \mathbf{1} \tag{1.3}$$

where $\mathbf{1}' = (1, \ldots, 1)$ and where the feasible region D_n is defined as

$$D_n = \{\boldsymbol{\delta} \in \mathbb{R}^p : X_t - \sum_{i=1}^{p} \delta_i X_{t-i} \geq 0, t = p+1, \ldots, n\}. \tag{1.4}$$

Assuming regular variation conditions on either the left or right tails of the innovations was sufficient to show that a limit distribution existed for $\hat{\boldsymbol{\phi}}$ and that rates of convergence were often superior to the Yule–Walker estimators. So a natural approach to the estimation problem for moving averages is to see what results from the autoregressive case can be brought to bear and thus we assume the moving average in (1.1) is *invertible* which according to Brockwell and Davis (1991) means $\Theta(z) \neq 0, \ |z| \leq 1$. This allows us to write

$$\Pi(z) := \frac{1}{\Theta(z)} = \sum_{i=0}^{\infty} \pi_k z^k, \quad |z| \leq 1,$$

and we hope we can convert (1.1) into an infinite order autoregression

$$\Pi(B)X_t = Z_t, \quad t = 0, \pm 1, \pm 2, \ldots.$$

If we now try to apply the lp estimators we find we have a nice objective function but the constraints involve an infinite number of variables. If we truncate the constraints suitably, we should obtain an estimator with worthwhile properties. The precise definition of our estimator of the moving average coefficients in the MA(q) process is

$$\hat{\theta} := \arg \ \max_{D_n} \sum_{i=1}^{q} \theta_i \tag{1.5}$$

where

$$D_n := \{\boldsymbol{\theta} : [\sum_{i=0}^{2l} (I - \Theta(B))^i] X_t \geq 0, t = 2lq+1, \ldots, n\} \tag{1.6}$$

and l is the first integer such that $2l \geq q$. Further motivation and discussion of this estimator is the subject of Section 2.

Here is a precise statement of the assumptions which will allow discussion of properties of our estimators. We need conditions which specify the model. In order to obtain a limit distribution for our estimators, we impose regular variation and moment conditions on the distribution of the innovation sequence. We recall that a function $U : [0, \infty) \mapsto (0, \infty)$ is regularly varying with exponent $\rho \in \mathbb{R}$ if

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\rho}, \quad x > 0.$$

(1) **Condition M** (model specification): The process $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ satisfies the equations (1.1)

$$X_t = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-1}, \quad t = 0, \pm 1, \pm 2, \ldots. \tag{1.1}$$

where $\{Z_t\}$ is an independent and identically distributed sequence of random variables with essential infimum (left endpoint) equal to 0 and common distribution function F. The coefficients $\theta_1, \ldots, \theta_q$ satisfy the invertibility condition that the moving average polynomial $\Theta(z) = \sum_{i=0}^{q} \theta_i z^i$ has no roots in the unit disk $\{z : |z| \leq 1\}$.

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

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

(3) **Condition R** (right tail): 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. $E(Z_t^{-\beta}) = \int_0^{\infty} u^{-\beta} F(du) < \infty$ for some $\beta > \alpha$.

Our results have as hypotheses M, and either L or R. Condition L is rather mild. It is satisfied if a density f of F exists which is continuous at 0 and with $f(0) > 0$. In this case $\alpha = 1$. Other common cases where Condition L holds are the Weibull distributions of the form $F(x) = 1 - \exp\{-x^{\alpha}\}$ where $F(x) \sim x^{\alpha}$, as $x \downarrow 0$ and the gamma densities $f(x) = ce^{-x}x^{r-1}$, $r > 0, x > 0$ so that $f(x) \sim cx^{r-1}$ as $x \downarrow 0$ and therefore the associated Gamma distribution function satisfies $F(x) \sim cr^{-1}x^{r}$, as $x \downarrow 0$. Examples of distributions satisfying condition R include positive stable densities and the Pareto density.

Section 2 further discusses motivation and properties of the mathematical programming estimator given in (1.6). Section 3 assumes Condition R and engages the point process limit theory (Resnick, 1987) which underlies discussion of the limit distributions for $\hat{\theta}$ carried out for the case $q = 1$. Section 4 parallels section 3 but assumes Condition L. In Section 5 we present some concluding remarks which emphasize the point that in contrast to the autoregressive case, the moving average estimators in the left tail case suffer a performance degradation depending on the order q of the model; no such degradation is present under condition R. Some future issues to be resolved are also considered.

2. The Parameter Estimator for MA(q).

Assume we have the invertible model $\{X_t\}$ specified by Condition M. Suppose the true value of the moving average coefficients is $\boldsymbol{\theta}^{(0)}$. In inverted form, the model can be written as the AR(∞) process

$$\Pi(B)X_t = Z_t, \quad t = 0, \pm 1, \pm 2, \dots \tag{2.1}$$

where

$$\frac{1}{\Theta(z)} = \Pi(z), \quad |z| \le 1.$$

For a finite order autoregression (1.2), the linear programming estimator of autoregressive coefficients is given by (1.3) and (1.4). If in (1.2) we write as usual the autoregressive polynomial as

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

then the objective function in (1.3) can be written as $1 - \Phi(1)$ and the constraints in (1.4) can be expressed as

$$\Phi(B)X_t \ge 0,\ t = p+1, \dots, n. \tag{2.2}$$

If we try to write down an analogous expression for the parameter estimators for the AR(∞) process in (2.1), we obtain as objective function

$$1 - \frac{1}{\Theta(1)} = \frac{\sum_{i=1}^q \theta_i}{1 + \sum_{i=1}^q \theta_i}$$

which is monotone in $\sum_{i=1}^q \theta_i$. So we try to maximize $\sum_{i=1}^q \theta_i$. For the constraints, (2.2) suggests the set of conditions

$$\Pi(B)X_t \ge 0, \quad t = 1, \dots, n.$$

A problem arises in that this constraint set requires knowledge of $X_t, X_{t-1}, \dots$ with the index extending back to $-\infty$ and since we only have knowledge of $X_1, \dots, X_n$ we must somehow truncate this constraint set.

A suggestion for how to construct a truncated set of constraints comes from symbolically expanding $1/\Theta$:

$$\frac{1}{\Theta(B)} = \frac{1}{I - (I - \Theta(B))} = \sum_{k=0}^{\infty}(I - \Theta(B))^k \approx \sum_{k=0}^{2l}(I - \Theta(B))^k,$$

where $l \ge 1$ is an integer to be specified. Note that

$$\sum_{k=0}^{2l}(I - \Theta(B))^k \Theta(B) = I - (I - \Theta(B))^{2l+1} = I + Q(B)^{2l+1},$$

where $Q(B) = \Theta(B) - I = \sum_{i=1}^q \theta_i B^i$. Let $\Theta^{(0)}(B) = \sum_{i=0}^q \theta_i^{(0)} B^i$ and $Q_0(B) = \sum_{i=1}^q \theta_i^{(0)} B^i$ and thus

$$\begin{aligned}\sum_{k=0}^{2l}(I - \Theta^{(0)}(B))^k X_t =& \sum_{k=0}^{2l}(I - \Theta^{(0)}(B))^k \Theta^{(0)}(B) Z_t \\ =& (I + Q_0^{2l+1}(B))Z_t \ge 0,\end{aligned}$$

since all ${\theta^{(0)}}_i$'s are assumed non-negative. So by truncating the series expansion for $1/\Theta$ in a judicious manner, the truncated expansion is always positive at the correct value of the parameter vector.

Thus our estimator is

$$\hat{\boldsymbol{\theta}} = \arg\max_{D_n} \sum_{i=1}^{q} \eta_i \tag{2.3}$$

where the constraint set is

$$D_n = \{\boldsymbol{\eta} \in \mathbb{R}_+^q : \sum_{k=0}^{2l} (I - \sum_{i=0}^{q} \eta_i B^i)^k X_t \geq 0, t = 2lq+1, \ldots, n;\ \eta_0 = 1, \sum_{i=0}^{q} \eta_i z^i \neq 0,\ |z| \leq 1\}. \tag{2.4}$$

The choice of l suggested by the limit theory is to choose l to be the first integer such that $2l \geq q$.

Change of variable: We seek a limit distribution for $q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)})$ where q_n is an appropriate scaling satisfying $q_n \to \infty$. It will turn out that under Condition R, the right choice of q_n is

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

and under Condition L the appropriate choice of q_n is

$$q_n = a_n = F^{\leftarrow}(\frac{1}{n}).$$

We observe that $q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)})$ satisfies

$$q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)})'\mathbf{1} \geq q_n(\boldsymbol{\eta} - \boldsymbol{\theta}^{(0)})'\mathbf{1}$$

for all $\boldsymbol{\eta} \in D_n$. Let $\boldsymbol{\delta} = q_n(\boldsymbol{\eta} - \boldsymbol{\theta}^{(0)})$ so that $q_n^{-1}\boldsymbol{\delta} + \boldsymbol{\theta}^{(0)} = \boldsymbol{\eta}$. Then $q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)})$ satisfies $q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)})'\mathbf{1} \geq \boldsymbol{\delta}'\mathbf{1}$ for all $\boldsymbol{\delta}$ such that

$$1 + \sum_{i=1}^{q} \left(\frac{\delta_i}{q_n} + \theta_i^{(0)}\right) z^i \neq 0, \quad |z| \leq 1$$

and

$$\sum_{k=0}^{2l} (-1)^k \left(Q_0(B) + \frac{\delta(B)}{q_n}\right)^k X_t \geq 0$$

for $t = 2lq+1, \ldots, n$, where $\delta(B) = \sum_{i=1}^{q} \delta_i B^i$. Thus

$$q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)}) = \arg\max_{\Lambda_n} \boldsymbol{\delta}'\mathbf{1}$$

where

$$\Lambda_n = \{\boldsymbol{\delta} \in \mathbb{R}_+^q : 1 + \sum_{i=1}^{q} \left(\frac{\delta_i}{q_n} + \theta_i^{(0)}\right) z^i \neq 0,\ |z| \leq 1, \sum_{k=0}^{2l} (-1)^k \left(Q_0(B) + \frac{\delta(B)}{q_n}\right)^k X_t \geq 0, t = 2lq+1, \ldots, n\}.$$

In case $q = 1$, we also have $l = 1$ and $\Theta(B) = I + \theta B$ so $I - \Theta(B) = -\theta B$ and the estimator is

$$\begin{aligned}\hat{\theta} &= \sup\{\eta \in [0,1) : X_t - \eta X_{t-1} + \eta^2 X_{t-2} \geq 0,\ t = 3, \ldots, n\} \\ &= \bigwedge_{t=3}^{n} \sup\{\eta \in [0,1) : X_t - \eta X_{t-1} + \eta^2 X_{t-2} \geq 0\}.\end{aligned}$$

Also

$$q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)}) = \arg \max_{\Lambda_n} \delta = \sup\{\delta \geq 0 : \delta \in \Lambda_n\}$$

where

$$\Lambda_n = \{\delta \geq 0 : 1 + \left(\frac{\delta}{q_n} + \theta^{(0)}\right) z \neq 0,\ |z| \leq 1, X_t - \left(\theta^{(0)} + \frac{\delta}{q_n}\right) X_{t-1} + \left(\theta_0 + \frac{\delta}{q_n}\right)^2 X_{t-2} \geq 0, t = 3, \dots, n\}.$$

Let $\eta = \delta / q_n$ and recall that $X_t = Z_t + \theta^{(0)} Z_{t-1}$. If we set

$$\begin{aligned} A_t &= Z_{t-2} + \theta^{(0)} Z_{t-3}, \\ B_t &= -Z_{t-1} + \theta^{(0)} Z_{t-2} + 2(\theta^{(0)})^2 Z_{t-3}, \\ C_t &= Z_t + (\theta^{(0)})^3 Z_{t-3}, \end{aligned} \tag{2.5}$$

then

$$\Lambda_n = q_n\{0 \leq \eta < 1 - \theta^{(0)} : A_t\eta^2 + B_t\eta + C_t \geq 0, t = 3, \dots, n\}$$

and we find in case $q = 1$ that

$$q_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(0)}) = q_n \bigwedge_{t=3}^{n} \sup\{0 \leq \eta < 1 - \theta^{(0)} : A_t\eta^2 + B_t\eta + C_t \geq 0\}. \tag{2.6}$$

So the limit distribution depends on the behavior of random parabolas and from extreme value theory we expect the limit distribution to be in the Weibull family. (Cf. Resnick, 1987, pages 14, 15.)

To analyze the limit distribution in (2.6), we intend to proceed as follows: Denote the random parabola by

$$p_t(\eta) = A_t\eta^2 + B_t\eta + C_t.$$

Only those parabolas such that $p_t(1 - \theta^{(0)}) < 0$ are of interest since if $p_t(1 - \theta^{(0)}) \geq 0$, then

$$\sup\{0 \leq \eta \leq 1 - \theta^{(0)} : p_t(\eta) \geq 0\} = 1 - \theta^{(0)},$$

which is an uninteresting contribution to the minimum in (2.6). Note that the condition $p_t(1-\theta^{(0)}) < 0$ also implies $B_t < 0$ and that the discriminant of the quadratic is positive so that the two roots of the quadratic are real. (The product of the roots of $p_t(\eta)$ is $C_t/A_t \geq 0$ so that both roots have the same sign. The sum of the roots is $-B_t/2A_t$. If $p_t(1 - \theta^{(0)}) < 0$, then the bigger root is positive which implies both roots are positive and hence $B_t < 0$.) Thus the smaller root r_t^- is the desired root and in (2.6)

$$q_n(\hat{\theta} - \theta^{(0)}) = q_n \bigwedge_{\substack{1 \leq t \leq n \\ p_t(1-\theta^{(0)}) < 0}} r_t^-.$$

Now

$$\begin{aligned} r_t^- &= \frac{-B_t - \sqrt{B_t^2 - 4A_tC_t}}{2A_t} \\ &= \left(|B_t| - |B_t|\left(1 - \frac{4A_tC_t}{B_t^2}\right)^{1/2}\right) / 2A_t \\ &\approx \frac{C_t}{|B_t|} \end{aligned} \tag{2.7}$$

where the last step resulted from expanding the function $(1+x)^{1/2}$ and neglecting remainder terms. Thus the behavior of

$$q_n \bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} r_t^-$$

should be determined by

$$q_n \bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} C_t/|B_t|$$

and the behavior of this quantity can be determined by using a point process argument which depends on whether Condition L or R is assumed.

The approximation in (2.7) can be justified by the following mechanism: Assume $p_t(1-\theta^{(0)})<0$. Consider the roots of the lines

$$L_1 : B_t\eta + C_t, \quad L_2 : (A_t(1-\theta^{(0)}) + B_t)\eta + C_t,$$

L_2 being the line passing through C_t and $p_t(1-\theta^{(0)})$. The roots of the two lines are

$$x_{t1} = \frac{C_t}{|B_t|}, \quad x_{t2} = \frac{C_t}{|B_t| - (1-\theta^{(0)})A_t}$$

and

$$x_{t1} \le r_t^- \le x_{t2}.$$

If we know the limit behavior of the point process depending on $\{x_{t1}\}$, and if x_{t2} is sufficiently close to x_{t1}, then the sandwiched piece r_t^- will behave properly and give us the limit distribution.

Details are in the next two sections which assume Condition R and then L.

3. The limit distribution in the right tail case for $q=1$.

In this section we assume Conditions M and R hold. We assume we are dealing with MA(1) so that $q=1$. The goal of this section is to present the limit distribution for $\hat\theta$. We will prove the following theorem.

Theorem 3.1. *Suppose $\{X_t\}$ is the MA(1) process given in (1.1) and that Conditions M, R hold. Suppose the true parameter is $\theta^{(0)} \in (0,1)$ and that F is continuous. Let $q_n = b_n$ be the quantile function*

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

where F is the distribution of Z_1. The estimator $\hat\theta$ given in Section 2 has a Weibull limit distribution: In $[0,\infty)$

$$b_n(\hat\theta - \theta^{(0)}) \Rightarrow \bigwedge_{k=1}^{\infty} \Gamma_k^{1/\alpha}(Y_k + (\theta^{(0)})^3 Y_k'), \tag{3.1}$$

where $\{Y_k, Y_k', k\ge 1\}$ are iid with common distribution F and

$$\Gamma_k = E_1 + \cdots + E_k, \quad k \ge 1,$$

is a sum of iid unit exponentially distributed random variables independent of $\{(Y_k, Y_k')\}$. The limit distribution of $\hat\theta$ is Weibull:

$$\lim_{n\to\infty} P[b_n(\hat\theta - \theta^{(0)}) \le x] = 1 - \exp\{-cx^\alpha\}, \quad x > 0, \tag{3.2}$$

where

$$c = E|Y_k + (\theta^{(0)})^3 Y_k'|^{-\alpha}$$

which is finite by the second statement of Condition R.

Before discussing the limit theory which leads to the asymptotic distribution of $\hat{\theta}$, we review rapidly some facts about point processes.

For a locally compact, Hausdorff topological space E, we let $M_p(E)$ be the space of Radon point measures on E. This means $m \in M_p(E)$ is of the form

$$m = \sum_{i=1}^{\infty} \epsilon_{x_i}$$

where $x_i \in E$ are the point masses of m and where

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

We emphasize that we assume that all measures in $M_p(E)$ are Radon which means that for any $m \in M_p(E)$ and any compact $K \subset E$, $m(K) < \infty$. On the space $M_p(E)$ we use the vague metric $\rho(\cdot,\cdot)$. Its properties are discussed for example in Resnick (1987, Section 3.4) or Kallenberg (1983). Note that a sequence of measures $m_n \in M_p(E)$ converge vaguely to $m_0 \in M_p(E)$ if for any continuous function $f : E \mapsto [0,\infty)$ with compact support we have $m_n(f) \to m_0(f)$ where $m_n(f) = \int_E f dm_n$. The non-negative continuous functions with compact support will be denoted $C_K^+(E)$.

A Poisson process on E with mean measure μ will be denoted PRM(μ). Two examples of the space E that interest us are $E = [0,\infty)$, where compact sets are those closed sets bounded away from ∞ and $E = [0,\infty]^p \setminus \{\mathbf{0}\}$, where compact sets are closed subsets of $[0,\infty]^p$ which are bounded away from $\mathbf{0}$. Other examples of the space E will be needed as well.

The fact that (3.1) implies (3.2) is a standard fact in extreme value theory. Since

$$\sum_{k=1}^{\infty} \epsilon_{(\Gamma_k, Y_k + (\theta^{(0)})^3 Y_k')}$$

is a Poisson process with mean measure $du \times P[Y_1 + (\theta^{(0)})^3 Y_1' \in dv]$ (Resnick, 1987, page 135), if we set

$$\Lambda := \{(u,v) \in [0,\infty)^2 : u^{1/\alpha} v \le x\},$$

then for $x > 0$

$$\begin{aligned} P[\bigwedge_{k=1}^{\infty} \Gamma_k^{1/\alpha}(Y_k + (\theta^{(0)})^3 Y_k') > x] &= P[\sum_{k=1}^{\infty} \epsilon_{(\Gamma_k, Y_k + (\theta^{(0)})^3 Y_k')}(\Lambda) = 0] \\ &= \exp\{-\iint_{\Lambda} du \times P[Y_1 + (\theta^{(0)})^3 Y_1' \in dv]\} \\ &= \exp\{-\int_{v=0}^{\infty} \int_{u \le (\frac{x}{v})^{\alpha}} du P[Y_1 + (\theta^{(0)})^3 Y_1' \in dv]\} \\ &= \exp\{-x^{\alpha} \int_0^{\infty} v^{-\alpha} P[Y_1 + (\theta^{(0)})^3 Y_1' \in dv]\} \end{aligned}$$

which is (3.2).

The main point process limit theorem which underlies our work in this section now follows. It is more general than we need for considering the asymptotic behavior of $\hat{\theta}$ in the MA(1) case but is stated in full generality for application to future work.

Proposition 3.2. *Suppose Conditions M and R hold and that $P[Z_1 = 0] = 0$. Define the measure*

$$\nu(dx) = \alpha x^{-\alpha-1}dx, \quad x > 0.$$

For any positive integer m we have

$$\begin{aligned} &nP[Z_0 \in dy_0, \frac{Z_i}{b_n} \in dx_i, i = 1, \dots, m; Z_j \in dy_j, j = 1, \dots, m] \\ &\xrightarrow{v} F(dy_0) \sum_{i=1}^{m} \nu(dx_i)\epsilon_\infty(dy_i) \prod_{\substack{j \neq i \\ 1 \leq j \leq m}} \epsilon_0(dx_j)F(dy_j), \end{aligned} \tag{3.3}$$

in

$$E := [0, \infty] \times ([0, \infty]^m \setminus \{\mathbf{0}\}) \times [0, \infty]^m. \tag{3.4}$$

Furthermore let E be defined as in (3.4) and for $l = 1, \dots, m$ set

$$\mathbf{e}_l = (0, \dots, 1, \dots, 0) \in \mathbb{R}^m$$

where the 1 appears in the lth spot. Suppose $\{Y_{k,l}, Y'_{k,l},\ k \geq 1, l \geq 1\}$ are iid with distribution F. Then

$$\begin{aligned} \sum_{t=1}^{n} \epsilon_{(Z_t, b_n^{-1}(Z_{t-i}, i=1,\dots,m), Z_{t-j}, j=1,\dots,m)} &\Rightarrow \sum_{k=1}^{\infty} \epsilon_{(Y_{k,1}, \Gamma_k^{-1/\alpha}\mathbf{e}_1, \infty, Y'_{k,1}, \dots, Y'_{k,m-1})} \\ &+ \sum_{k=1}^{\infty} \epsilon_{(Y_{k,2}, \Gamma_k^{-1/\alpha}\mathbf{e}_2, Y_{k,1}, \infty, Y'_{k,1}, \dots, Y'_{k,m-2})} + \dots + \sum_{k=1}^{\infty} \epsilon_{(Y_{k,m}, \Gamma_k^{-1/\alpha}\mathbf{e}_m, Y_{k,m-1}, \dots, Y_{k,1}, \infty)} \end{aligned} \tag{3.5}$$

in $M_p(E)$.

Proof. The proof of (3.3) is based on the following two simple results. For $y, a, b \geq 0$, since $b_n \to \infty$

$$nP[b_n^{-1}Z_t > y, Z_t \in [a, b]] \to \begin{cases} 0, & \text{if } b < \infty, \\ y^{-\alpha}, & \text{if } b = \infty, \end{cases}$$

and for $x > 0, y \geq 0$

$$nP[b_n^{-1}Z_t \geq x, b_n^{-1}Z_{t-1} \geq y] \to \begin{cases} x^{-\alpha}, & \text{if } y = 0, \\ 0, & \text{if } y > 0, \end{cases}$$

and furthermore we have checked the vague convergence of

$$nP[\left(\frac{Z_t}{b_n}, Z_t\right) \in \cdot] \xrightarrow{v} \nu \times \epsilon_0$$

as measures on $(0, \infty] \times [0, \infty]$. To prove (3.5) we let ρ be the vague metric on $M_p(E)$. Then we can show that

$$\rho(\sum_{t=1}^{n} \epsilon_{(Z_t, b_n^{-1}(Z_{t-1}, Z_{t-2}, \dots, Z_{t-m}), Z_{t-1}, Z_{t-2}, \dots, Z_{t-m})}, \sum_{t=1}^{n}\sum_{i=1}^{m} \epsilon_{(Z_t, b_n^{-1}Z_{t-i}\mathbf{e}_i, Z_{t-1}, Z_{t-2}, \dots, Z_{t-m})}) \xrightarrow{P} 0 \tag{3.6}$$

and

$$\rho(\sum_{t=1}^{n}\sum_{i=1}^{m}\epsilon_{(Z_t, b_n^{-1}Z_{t-i}\mathbf{e}_i, Z_{t-1}, Z_{t-2}, \dots, Z_{t-m})}, \sum_{t=1}^{n}\sum_{i=1}^{m}\epsilon_{(Z_{t+i}, b_n^{-1}Z_t\mathbf{e}_i, Z_{t+i-1}, Z_{t+i-2}, \dots, Z_{t+i-m})}) \xrightarrow{P} 0. \tag{3.7}$$

The proof of (3.6) is almost exactly the same as that of the Proposition 4.26 of Resnick (1987) and rests on the fact that for any $\delta > 0$

$$I_n := \sum_{t=1}^{n} \epsilon_{(Z_t, b_n^{-1}(Z_{t-1}, Z_{t-2}, \dots, Z_{t-m}), Z_{t-1}, Z_{t-2}, \dots, Z_{t-m})}$$

has the property that

$$EI_n(\bigcup_{1\le i<j\le m}\{(y_0, x_1, \dots, x_m, y_1, \dots, y_m) : x_i > \delta, x_j > \delta\}) \le \binom{m}{2} n \Pr[Z_1 > \delta b_n, Z_2 > \delta b_n]$$

which tends to 0 as $n \to \infty$ (cf. Feigin and Resnick (1992), (3.37)). The proof of (3.7) is identical to that of (3.40) in Feigin and Resnick (1992). From (3.7) and (3.6) we see that to prove (3.5), it suffices to show that the right most point process in (3.7) converges to the limit in (3.5). Towards showing this, we assert that

$$\sum_{t=1}^{n} \epsilon_{(Z_{t+m}, \dots, Z_{t+1}, b_n^{-1}Z_t, Z_t, Z_{t-1}, \dots, Z_{t-m+1})} \Rightarrow \sum_{k\ge 1} \epsilon_{(Y_{k,m}, \dots, Y_{k,1}, \Gamma_k^{-1/\alpha}, \infty, Y'_{k,1}, \dots, Y'_{k,m-1})}$$

where $\{Y'_{k,i}, Y_{k,j}, k \ge 1, 1 \le i, j \le m\}$ are iid random variables with the distribution of Z_1 and independent of $\{\Gamma_k\}$. To see this, let

$$X_{n,t} = (Z_{t+m}, \dots, Z_{t+1}, \frac{Z_t}{b_n}, Z_t, Z_{t-1}, \dots, Z_{t-m+1})$$

and observe that $\{X_{n,t}, -\infty < t < \infty\}$ satisfies

(1) $\{X_{n,t}, -\infty < t < \infty\}$ is stationary and $2m$–dependent.

(2) We have

$$nP[X_{n,1} \in (dz_m, \dots, dz_1, dx, dy, du_1, \dots, du_{m-1})] \xrightarrow{v} \prod_{i=1}^{m} F(dz_i)\nu(dx)\epsilon_\infty(dy)\prod_{j=1}^{m-1}F(du_j),$$

on $[0, \infty]^m \times (0, \infty] \times [0, \infty]^m := E'$.

(3) For $g \in C_K^+(E')$

$$\lim_{k\to\infty}\limsup_{n\to\infty} n\sum_{t=2}^{[n/k]} Eg(X_{n,1})g(X_{n,t}) = 0.$$

The desired result then follows from Theorem 2.1 in Davis and Resnick (1988).

By projecting, we then get in $(M_p([0, \infty] \times (0, \infty] \times [0, \infty]^{m-1}))^m$ that

$$\begin{aligned}&(\sum_{t=1}^{n}\epsilon_{(Z_{t+1}, b_n^{-1}Z_t, Z_t, \dots, Z_{t+1-m})}, \sum_{t=1}^{n}\epsilon_{(Z_{t+2}, b_n^{-1}Z_t, Z_{t+1}, \dots, Z_{t+2-m})},\\ &\quad \dots, \sum_{t=1}^{n}\epsilon_{(Z_{t+m}, b_n^{-1}Z_t, Z_{t+m-1}, \dots, Z_t)}) \Rightarrow\\ &(\sum_k \epsilon_{(Y_{k,1}, \Gamma_k^{-1/\alpha}, \infty, Y'_{k,1}, \dots, Y'_{k,m-1})}, \sum_k \epsilon_{(Y_{k,2}, \Gamma_k^{-1/\alpha}, Y_{k,1}, \infty, Y'_{k,1}, \dots, Y'_{k,m-2})}, \dots,\\ &\quad \sum_k \epsilon_{(Y_{k,m}, \Gamma_k^{-1/\alpha}, Y_{k,m-1}, \dots, Y_{k,1}, \infty)}).\end{aligned}$$

Hence by using a mapping argument, we have

$$(\sum_{t=1}^{n} \epsilon_{(Z_{t+i}, b_n^{-1} Z_t \mathbf{e}_i, Z_{t+i-1}, Z_{t+i-2}, \dots, Z_{t+i-m})}; i = 1, \dots, m) \Rightarrow$$
$$(\sum_{k} \epsilon_{(Y_{k,i}, \Gamma_k^{-1/\alpha} \mathbf{e}_i, Y_{k,i-1}, \dots, Y_{k,1}, \infty, Y'_{k,1}, \dots, Y_{k,m-i})}.$$

Since addition is vaguely continuous we finally obtain that the right most point process in (3.7) converges weakly to the limit in (3.5) and therefore by using (3.6) and (3.7), we deduce the result of the Proposition. □

Before proceeding with the proof of Theorem 3.1 based on Proposition 3.2, we state some preliminaries. Suppose $E' \subset E$ and give E' the relative topology inherited from E. The compact subsets of E' are those subsets $K' \subset E'$ such that K' is compact when considered as a subset of E. To see this, suppose K' is compact in E'. Suppose $O_\beta, \beta \in A$ is an open covering of K' in E, so that $K' \subset \cup_{\beta \in A} O_\beta$. Since $K' \subset E'$, we also have

$$K' \subset \left(\bigcup_{\beta \in A} O_\beta \right) \bigcap E' = \bigcup_{\beta \in A} \left(O_\beta \bigcap E' \right).$$

Since $O_\beta \cap E'$ is open in E' and K' is compact in E', we have

$$K' \subset \bigcup_{\beta \in I} \left(O_\beta \bigcap E' \right) \subset \bigcup_{\beta \in I} O_\beta,$$

where I is a finite index set. Thus K' is compact in E. The converse is similar.

Proposition 3.3. *Suppose E' is a measurable subset of E and give E' the relative topology inherited from E. For a set $B \subset E'$ denote by $\partial_{E'} B$ the boundary of B in E' and denote by $\partial_E B$ the boundary of B in E.*

(a) Define

$$\hat{T} : M_p(E) \mapsto M_p(E')$$

by

$$\hat{T} m = m(\cdot \cap E').$$

If $m \in M_p(E)$ and $m(\partial_E E') = 0$, then $\hat{T}$ is continuous at m so that if $m_n \xrightarrow{v} m$ in $M_p(E)$, then $\hat{T} m_n \xrightarrow{v} \hat{T} m$ in $M_p(E')$.

(b) The same conclusion holds if we define $\hat{T}$ the same way but consider it as a mapping

$$\hat{T} : M_p(E) \mapsto M_p(E).$$

(c) Conversely, suppose $m_n \in M_p(E)$ for $n \geq 0$ and that $m_n \xrightarrow{v} m_0$ in $M_p(E')$. If

$$m_n((E')^c) = 0, \quad n \geq 0$$

and $m(\partial_E E') = 0$, then $m_n \xrightarrow{v} m_0$ in $M_p(E)$ as well.

Proof. (a) Suppose $B \subset E'$ is relatively compact in E' and $m(\partial_{E'} B) = 0$. It suffices to show $m_n(B) \to m(B)$ (Resnick, 1987, page 142). Since $m_n \xrightarrow{v} m$ in $M_p(E)$, it suffices to show that $m(\partial_E B) = 0$. One can readily check the inclusion

$$\partial_E B \subset \partial_{E'} B \bigcup \partial_E E' \tag{3.8}$$

for $B \subset E'$. Thus,

$$m(\partial_E B) \leq m(\partial_{E'} B) + m(\partial_E E') = 0.$$

The proof of (b) and (c) is very similar except one needs the inclusion

$$\partial_E (B \cap E') \subset (\partial_E B) \cap \bar{E}' \bigcup \partial_E E'. \tag{3.9}$$ □

The following simple result allows us to discard components in a point process convergence result.

Lemma 3.4. *Suppose $E_i, i = 1, 2$ are locally compact, Hausdorff topological spaces and that E_2 is compact. If $m_n \in M_p(E_1 \times E_2)$ for $n \geq 0$ and $m_n \xrightarrow{v} m_0$ in $M_p(E_1 \times E_2)$, then*

$$m_n(\cdot \times E_2) \xrightarrow{v} m_0(\cdot \times E_2)$$

in $M_p(E_1)$.

Proof. Let $f_1 \in C_K^+(E_1)$ and define $f : E_1 \times E_2 \mapsto \mathbb{R}_+$ by $f(x, y) = f_1(x)$. The support of f is contained in $K_1 \times E_2$ where K_1 is the compact support of f_1. Thus the support of f is compact and $m_n(f) \to m_0(f)$ which translates to

$$\int_{E_1} f_1(x) m_n(dx \times E_2) \to \int_{E_1} f_1(x) m_0(dx \times E_2)$$

which is equivalent to the desired convergence in $M_p(E_1)$. $\square$

We now proceed to prove Theorem 3.1. From Proposition 3.2 and Lemma 3.4 we have

$$\begin{aligned} N_n &:= \sum_{t=1}^{n} \epsilon_{(Z_t, b_n^{-1}(Z_{t-1}, Z_{t-2}, Z_{t-3}), Z_{t-1}, Z_{t-2}, Z_{t-3})} \\ \Rightarrow N_\infty &:= \sum_{k=1}^{\infty} \epsilon_{(Y_{k,1}, \Gamma_k^{-1/\alpha}, 0, 0, \infty, Y'_{k,1}, Y'_{k,2})} + \sum_{k=1}^{\infty} \epsilon_{(Y_{k,2}, 0, \Gamma_k^{-1/\alpha}, 0, Y_{k,1}, \infty, Y'_{k,1})} \\ &\quad + \sum_{k=1}^{\infty} \epsilon_{(Y_{k,3}, 0, 0, \Gamma_k^{-1/\alpha}, Y_{k,2}, Y_{k,1}, \infty)} \end{aligned} \tag{3.10}$$

in $M_p(E)$ where recall $E := [0, \infty] \times ([0, \infty]^3 \setminus \{\mathbf{0}\}) \times [0, \infty]^3$.

Referring back to the end of Section 2 and (2.5), recall that we are interested only in the case where $p_t(1 - \theta^{(0)}) < 0$. However, it is initially easier to deal with the restriction of the point process convergence in (3.10) for the case that $B_t < 0$ which is implied by $p_t(1 - \theta^{(0)}) < 0$. So we define the region $[B < 0] \subset E$ by

$$\begin{aligned} [B < 0] :=& \{(x_0, \dots, x_6) \in E : (x_4 > \theta^{(0)} x_5 + 2(\theta^{(0)})^2 x_6, x_1 \neq \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3\} \\ =& \{(x_0, \dots, x_6) \in [0, \infty] \times ((0, \infty] \times [0, \infty)^2)^2 : (x_4 > \theta^{(0)} x_5 + 2(\theta^{(0)})^2 x_6, x_1 > \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3\} \\ & \cup \{(x_0, \dots, x_6) \in [0, \infty] \times [0, \infty) \times ([0, \infty]^2 \setminus \{(0, 0)\}) \times (0, \infty] \times [0, \infty)^2 : \\ & \qquad x_4 > \theta^{(0)} x_5 + 2(\theta^{(0)})^2 x_6, x_1 < \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3\} \\ =& [B < 0]_> \cup [B < 0]_<. \end{aligned}$$

Note that for $b > 0, c > 0, b_1 > 0, b_4 > 0, a_j < \infty, j = 2, 3, 5, 6$ the set

$$\begin{aligned} \{(x_0, \dots, x_6) \in& [0, \infty] \times ((0, \infty] \times [0, \infty)^2)^2 : \\ & (x_4 > \theta^{(0)} x_5 + 2(\theta^{(0)})^2 x_6, x_1 > \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3, x_i \geq b_i, i = 1, 4; x_j \leq a_j, j = 2, 3, 5, 6\} \end{aligned}$$

is a compact subset of $[0, \infty] \times ((0, \infty] \times [0, \infty)^2)^2$. Define the mappings

$$\begin{aligned} T_1(x_0, \dots, x_6) :=& x_0 + (\theta^{(0)})^3 x_6, \\ T_2(x_0, \dots, x_6) :=& |-x_1 + \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3|, \\ T(x_0, \dots, x_6) :=& (T_1(x_0, \dots, x_6), T_2(x_0, \dots, x_6), x_4, x_5, x_6) \end{aligned} \tag{3.11}$$

and think of T as

$$\begin{aligned}T:[B<0]\mapsto E' :=&[0,\infty]\times(0,\infty]\times\{(x_4,x_5,x_6)\in[0,\infty]^3 : x_4>\theta^{(0)}x_5+2(\theta^{(0)})^2x_6\}\\ =&[0,\infty]\times(0,\infty]\times\{(x_4,x_5,x_6)\in(0,\infty]\times[0,\infty)^2 : x_4>\theta^{(0)}x_5+2(\theta^{(0)})^2x_6\}.\end{aligned}$$

Note that

$$T(Z_t,b_n^{-1}(Z_{t-1},Z_{t-2},Z_{t-3}),Z_{t-1},Z_{t-2},Z_{t-3})=(C_t,\frac{|B_t|}{b_n},Z_{t-1},Z_{t-2},Z_{t-3}).$$

We seek to show

$$\sum_{t=1}^n 1_{[B_t<0]}\epsilon_{(C_t,b_n^{-1}|B_t|,Z_{t-1},Z_{t-2},Z_{t-3})}\Rightarrow N_\infty^{\#}:=\sum_{k=1}^\infty \epsilon_{(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2},\Gamma_k^{-1/\alpha},\infty,Y'_{k,1},Y'_{k,2})},\tag{3.12}$$

in $M_p([0,\infty]\times(0,\infty]\times[0,\infty]^3)$.

Since

$$P[N_\infty(\partial_E([B<0])=0]=1,$$

we obtain from Proposition 3.3 that if we take restrictions to $[B<0]$ in (3.10), we get

$$N_n|_{[B<0]}\Rightarrow N_\infty|_{[B<0]}=\sum_{k=1}^\infty\epsilon_{(Y_{k,1},\Gamma_k^{-1/\alpha},0,0,\infty,Y'_{k,1},Y'_{k,2})},\tag{3.13}$$

in $M_p([B<0])$. It is tempting to try to apply Proposition 3.18, page 148, of Resnick, 1987. To apply this theorem, we need to check that T is continuous (no problem) and that T^{-1} maps compact sets into compact sets (problem). This last compactness property fails, so to get around the problem, truncation of the domain is necessary. For $M>0$, let

$$\begin{aligned}[B<0]_M=&[B<0]_>\cap\{(x_0,\dots,x_6):x_1\ge M^{-1},x_2\vee x_3\le M\}\\ &\cup[B<0]_<\cap\{(x_0,\dots,x_6):x_1\le M,x_2\vee x_3\ge M^{-1}\}.\end{aligned}$$

From (3.13) we get by restriction

$$N_n|_{[B<0]_M}\Rightarrow N_\infty|_{[B<0]_M}$$

in $M_p([B<0]_M)$. Considered as a mapping on $[B<0]_M$, T^{-1} maps compacta into compacta. For instance, for $a>0,b,b',c>0,a_i<\infty,i=5,6$

$$\begin{aligned}&T^{-1}([0,a]\times[b,\infty]\times\{(x_4,x_5,x_6):x_4\ge b',x_i\le a_i,i=5,6;x_4\ge\theta^{(0)}x_5+2(\theta^{(0)})^2x_6+c\})\\ &=\{(x_0,\dots,x_6)\in[0,\infty]\times((0,\infty]\times[0,\infty)^2)^2:x_4\ge b',x_i\le a_i,i=5,6;x_4\ge\theta^{(0)}x_5+2(\theta^{(0)})^2x_6+c,\\ &\qquad x_0+(\theta^{(0)})^3x_6\le a,x_2\vee x_3\le M,x_1\ge\theta^{(0)}x_2+2(\theta^{(0)})^2x_3+b\}\end{aligned}$$

is compact in $[0,\infty]\times((0,\infty]\times[0,\infty)^2)^2$ and hence in $[B<0]_M$. Thus we conclude from Proposition 3.13, page 148, Resnick, 1987 that

$$N_n|_{[B<0]_M}\circ T^{-1}\Rightarrow N_\infty|_{[B<0]_M}\circ T^{-1}$$

in $M_p(E')$. Part (c) of Proposition 3.3 then implies

$$\begin{aligned}N_{n,M}:=&\sum_{t=1}^n 1_{\{[B_t<0]\cap([Z_{t-1}>b_nM^{-1},Z_{t-2}\vee Z_{t-3}\le b_nM]\cup[Z_{t-1}\le b_nM,Z_{t-2}\vee Z_{t-3}>M^{-1}b_n])\}}\epsilon_{(C_t,b_n^{-1}|B_t|,Z_{t-1},Z_{t-2},Z_{t-3})}\\ &\Rightarrow N_\infty^{\#}\end{aligned}$$

in $M_p([0,\infty]\times(0,\infty]\times[0,\infty]^3)$. From Billingsley, 1968, Theorem 4.2, to show (3.12) we need to verify that for any $\eta>0$

$$\lim_{M\to\infty}\limsup_{n\to\infty} P[\rho(N_{n,M},N_n)>\eta]=0.$$

Let $g\in C_K^+([0,\infty]\times(0,\infty]\times[0,\infty]^3)$ and suppose the support of g is in $[0,\infty]\times[b,\infty]\times[0,\infty]^3$. It suffices to show for all $\eta>0$ that

$$\lim_{M\to\infty}\limsup_{n\to\infty} P[|N_{n,M}(g)-N_n(g)|>\eta]=0.$$

Now

$$\begin{aligned}|N_{n,M}(g)-N_n(g)|\le\sum_{t=1}^n &1_{\{[B_t<0]\cap([Z_{t-1}>b_nM^{-1},Z_{t-2}\vee Z_{t-3}\le b_nM]\cup[Z_{t-1}\le b_nM,Z_{t-2}\vee Z_{t-3}>b_nM^{-1}])^c\}}\\ &g(C_t,b_n^{-1}|B_t|,Z_{t-1},Z_{t-2},Z_{t-3})\end{aligned}$$

and so

$$\begin{aligned}&P[|N_{n,M}(g)-N_n(g)|>\eta]\\ &\le nP\{[B_t<0,|B_t|/b_n\ge b]\cap([Z_{t-1}>b_nM^{-1},Z_{t-2}\vee Z_{t-3}\le b_nM]\cup\\ &\qquad[Z_{t-1}\le b_nM,Z_{t-2}\vee Z_{t-3}>b_nM^{-1}])^c\}\end{aligned}$$

and since $B_t<0$ and $|B_t|/b_n\ge b$ imply $Z_{t-1}/b_n>b$ we have the bound

$$\begin{aligned}&\le nP\{[Z_{t-1}/b_n>b]\cap([Z_{t-1}\le b_nM^{-1}]\cup[Z_{t-2}\vee Z_{t-3}>b_nM])\}\\ &\le nP[Z_{t-1}>b_nb,Z_{t-2}\vee Z_{t-3}>b_nM]\\ &=nP[Z_{t-1}>b_nb]P[Z_{t-2}\vee Z_{t-3}>b_nM].\end{aligned}$$

As $n\to\infty$ this last expression is asymptotic to

$$b^{-\alpha}P[Z_{t-2}\vee Z_{t-3}>b_nM]$$

which goes to zero as $n\to\infty$.

Now corresponding to the condition $p_t(1-\theta^{(0)})$ define

$$\begin{aligned}NEG:=&\{(x_0,x_1,x_2,x_3,x_4)\in[0,\infty]\times(0,\infty]\times[0,\infty]^3:\\ &(x_3+\theta^{(0)}x_4)(1-\theta^{(0)})^2+(-x_2+\theta^{(0)}x_3+2(\theta^{(0)})^2x_4)(1-\theta^{(0)})+x_0+(\theta^{(0)})^3x_4<0\}.\end{aligned}$$

Since

$$P[N_\infty^{\#}(\partial(NEG))=0]=1,$$

and $p_t(1-\theta^{(0)})<0$ implies $B_t<0$, we get from parts (a) and (c) of Proposition 3.3 that

$$\sum_{t=1}^n 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{(C_t,b_n^{-1}|B_t|,Z_{t-1},Z_{t-2},Z_{t-3})}\Rightarrow N_\infty^{\#}=\sum_{k=1}^\infty \epsilon_{(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2},\Gamma_k^{-1/\alpha},\infty,Y'_{k,1},Y'_{k,2})}, \tag{3.14}$$

and applying Lemma 3.4 gives

$$\sum_{t=1}^n 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{(C_t,b_n^{-1}|B_t|)}\Rightarrow\sum_{k=1}^\infty \epsilon_{(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2},\Gamma_k^{-1/\alpha})}, \tag{3.15}$$

in $M_p([0,\infty]\times(0,\infty])$. Finally, another application of Proposition 3.3 (a) gives the desired result

$$\sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{(C_t,b_n^{-1}|B_t|)} \Rightarrow \sum_{k=1}^{\infty}\epsilon_{(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2},\Gamma_k^{-1/\alpha})}, \tag{3.16}$$

in $M_p([0,\infty)\times(0,\infty])$.

The next step is to show that we can take the ratio of the components in (3.16) and we show

$$\nu_{0,n} := \sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{b_nC_t/|B_t|} \Rightarrow \nu_{0,\infty} := \sum_{k=1}^{\infty}\epsilon_{\Gamma_k^{1/\alpha}(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2})}, \tag{3.17}$$

in $M_p([0,\infty))$. The map

$$T_3 : (u,v)\in[0,\infty)\times(0,\infty]\mapsto u/v\in[0,\infty),$$

although continuous, does not have the property that T_3^{-1} carries compacta into compacta and so a truncation of the domain must be done. For small $\delta>0$, we restrict attention to $[0,\delta^{-1}]\times[\delta,\infty]$ and then apply T_3 to get as $n\to\infty$

$$\begin{aligned}\nu_{\delta,n} :=& \sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0,C_t\le\delta^{-1},b_n^{-1}|B_t|\ge\delta]}\epsilon_{b_nC_t/|B_t|}\\ \Rightarrow&\nu_{\delta,\infty} := \sum_{k} 1_{[Y_{k,1}+(\theta^{(0)})^3Y'_{k,2}\le\delta^{-1},\Gamma_k^{-1/\alpha}\ge\delta]}\epsilon_{\Gamma_k^{1/\alpha}(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2})},\end{aligned}$$

in$M_p([0,\infty))$. As $\delta\to0$ we have

$$\nu_{\delta,\infty} := \sum_{k} 1_{[Y_{k,1}+(\theta^{(0)})^3Y'_{k,2}\le\delta^{-1},\Gamma_k^{-1/\alpha}\ge\delta]}\epsilon_{\Gamma_k^{1/\alpha}(Y_{k,1}+(\theta^{(0)})^3Y'_{k,2})} \Rightarrow \nu_{0,\infty} := \sum_{k}\epsilon_{\Gamma_k^{1/\alpha}(Y_{k1}+(\theta^{(0)})^3Y'_{k13})},$$

in $M_p([0,\infty))$, which is the right side of (3.17). To show (3.17), therefore, it suffices by Theorem 4.2 of Billingsley, 1968 to prove that for any $\eta>0$

$$\lim_{\delta\to0}\limsup_{n\to\infty}P[\rho(\nu_{\delta,n},\nu_{0,n})>\eta]=0, \tag{3.18}$$

where ρ is the vague metric.

Let $g\in C_K^+([0,\infty))$ and suppose the support of g is in $[0,a]$. For (3.18), it suffices to show that

$$\lim_{\delta\to0}\limsup_{n\to\infty}P[|\nu_{\delta,n}(g)-\nu_{0,n}(g)|>\eta]=0,$$

or equivalently

$$\lim_{\delta\to0}\limsup_{n\to\infty}P[|\sum_{t=1}^{n} g(b_n\frac{C_t}{|B_t|})1_{[\frac{|B_t|}{b_n}\ge\delta,C_t\le\delta^{-1},p_t(1-\theta^{(0)})<0]} - \sum_{t=1}^{n} g(b_n\frac{C_t}{|B_t|})1_{[p_t(1-\theta^{(0)})<0]}|>\eta]=0. \tag{3.19}$$

The absolute value of the difference between the two sums is bounded by

$$\begin{aligned}&\sum_{t=1}^{n} g(b_n\frac{C_t}{|B_t|})1_{[p_t(1-\theta^{(0)})<0]}1_{[|B_t|\ge\delta b_n,C_t\le\delta^{-1}]^c}\\ &\le\sum_{t=1}^{n} g(b_n\frac{C_t}{|B_t|})1_{[B_t<0]\cap[|B_t|\ge\delta b_n,C_t\le\delta^{-1}]^c}\\ &\le\sum_{t=1}^{n} g(b_n\frac{C_t}{|B_t|})(1_{[B_t<0,|B_t|\le\delta b_n,C_t\le\delta^{-1}]}\\ &\quad+\sum_{t=1}^{n} g(b_n\frac{C_t}{|B_t|})(1_{[B_t<0,C_t>\delta^{-1}]}\end{aligned}$$

Taking into account the compact support of g, the probability in (3.19) is thus bounded by

$$\begin{aligned}
P\{\bigcup_{t=1}^{n}[b_n\frac{C_t}{|B_t|} \le a, B_t < 0, \frac{|B_t|}{b_n} \le \delta, C_t \le \delta^{-1}]\}& \\
+ P\{\bigcup_{t=1}^{n}[b_n\frac{C_t}{|B_t|} \le a, B_t < 0, C_t > \delta^{-1}]\}& \\
\le nP[b_n\frac{C_t}{|B_t|} \le a, B_t < 0, \frac{|B_t|}{b_n} \le \delta, C_t \le \delta^{-1}]& \\
+ nP[b_n\frac{C_t}{|B_t|} \le a, B_t < 0, C_t > \delta^{-1}]& \\
= I + II.&
\end{aligned}$$

Now II is bounded as follows:

$$II \le nP[\frac{\delta^{-1}}{a} \le \frac{|B_t|}{b_n}, B_t < 0]$$

and because $B_t < 0$ implies $Z_{t-1} > \theta^{(0)}Z_{t-2} + 2(\theta^{(0)})^3 Z_{t-3}$, the previous expression is dominated by

$$\begin{aligned}
&nP[\frac{\delta^{-1}}{a} \le \frac{Z_{t-1}}{b_n}] \to \left(\frac{\delta^{-1}}{a}\right)^{-\alpha}, \quad (n \to \infty)\\
=&\delta^{\alpha}a^{\alpha} \to 0, \quad (\delta \to 0).
\end{aligned}$$

To bound I observe that for arbitrarily small $\omega < \delta^{-1}$ we have

$$\begin{aligned}
I \le& nP[b_n\frac{C_t}{|B_t|} \le a, B_t < 0, \frac{|B_t|}{b_n} \le \delta, C_t \le \omega]\\
&+ nP[b_n\frac{C_t}{|B_t|} \le a, B_t < 0, C_t > \omega]\\
=&Ia + Ib.
\end{aligned}$$

Now Ib has a double limit which is zero by the argument that handled II so we concentrate on Ia. From the definition of C_t we have $C_t \ge Z_t$, and $C_t \ge (\theta^{(0)})^3 Z_{t-3}$. Recall that $Z_{t-1} \ge \theta^{(0)}Z_{t-2} + 2(\theta^{(0)})^3 Z_{t-3}$, when $B_t < 0$. Thus Ia is bounded by

$$\begin{aligned}
Ia \le& nP[a^{-1} \le \left(Z_t^{-1}\right)\frac{|B_t|}{b_n}, B_t < 0, C_t \le \omega]\\
\le& nP[a^{-1} \le \left(Z_t^{-1}\right)\frac{Z_{t-1}}{b_n}, (\theta^{(0)})^3 Z_{t-3} \le \omega]\\
\to& E(Z_t)^{-\alpha}a^{\alpha}P[(\theta^{(0)})^3 Z_{t-3} \le \omega], \quad (n \to \infty)
\end{aligned}$$

since by Condition R, $E(Z_t)^{-\beta} < \infty$ for some $\beta > \alpha$ and thus

$$\lim_{n\to\infty}\limsup_{\delta\to 0} Ia \le (constant)P[(\theta^{(0)})^3 Z_{t-3} \le \omega]$$

and since ω can be picked arbitrarily small, the double limit must be zero as desired. This completes the verification of (3.17).

From (3.13), a standard argument mapping the points of $\nu_{0,n}$ into the minimum (see, for example, Resnick, 1987, page 214) yields

$$b_n \bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} \frac{C_t}{|B_t|} \Rightarrow \bigwedge_{k=1}^{\infty} \Gamma_k^{1/\alpha}(Y_{k,1}+(\theta^{(0)})^3 Y'_{k,2}). \tag{3.20}$$

The rest of the proof consists in showing that in fact $b_n(\hat\theta - \theta^{(0)})$ has the same limit distribution as the limit random variable in (3.20).

To do this, recall the outline presented at the end of Section 2. When $p_t(1-\theta^{(0)}) < 0$, we approximate the polynomial $p_t(\eta)$ with two lines

$$L_1 : B_t + \eta + C_t, \quad L_2 : (A_t(1-\theta^{(0)}) + B_t)\eta + C_t,$$

which have roots

$$x_{t1} = \frac{C_t}{|B_t|}, \quad x_{t2} = \frac{C_t}{|B_t| - (1-\theta^{(0)})A_t}$$

and therefore the smallest root r_t^- of the random polynomial satisfies

$$x_{t1} \le r_t^- \le x_{t2}. \tag{3.21}$$

We know from (3.13) that

$$\sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{b_n x_{t1}} \Rightarrow \nu_{0,\infty} := \sum_k \epsilon_{\Gamma_k^{1/\alpha}(Y_{k,1}+(\theta^{(0)})^3 Y'_{k,2})},$$

in $M_p([0,\infty))$, and we now propose to show that the same holds true with x_{t2} replacing x_{t1} and we show

$$\rho\Big(\sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{b_n x_{t1}}, \sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]}\epsilon_{b_n x_{t2}}\Big) \xrightarrow{P} 0, \tag{3.22}$$

as $n \to \infty$. Suppose $f \in C_K^+([0,\infty))$ with support in $[0,k]$. It is enough to show for any η

$$\lim_{n\to\infty} P\Big[\sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[p_t(1-\theta^{(0)})<0]} > \eta\Big] = 0.$$

For any small $\delta > 0$, this probability is bounded by

$$\begin{aligned}
&P\Big[\sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0]} > \eta\Big] \\
&\quad \le P\Big[\sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0, b_n x_{t2}>k+\delta, b_n|x_{t2}-x_{t1}|<\delta]} > \eta/3\Big] \\
&\quad\quad + P\Big[\sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0, b_n x_{t2}>k+\delta, b_n|x_{t2}-x_{t1}|\ge\delta]} > \eta/3\Big] \\
&\quad\quad + P\Big[\sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0, b_n x_{t2}\le k+\delta]} > \eta/3\Big] \\
&\quad = I + II + III.
\end{aligned}$$

Note that I is zero because both arguments of f are outside the support of f.

For II, we have the bounds

$$\begin{aligned}II \leq& P[\sum_{t=1}^{n} f(b_n x_{t1}) 1_{[b_n x_{t1}\leq k, B_t<0, |x_{t1}-x_{t2}|\geq \delta]} > \eta/3]\\ \leq& P\{\cup_{t=1}^{n}[[b_n x_{t1} \leq k, B_t < 0, b_n|x_{t1}-x_{t2}| \geq \delta]\}\\ \leq& nP[[b_n x_{t1} \leq k, B_t < 0, b_n|x_{t1}-x_{t2}| \geq \delta].\end{aligned}$$

Now

$$b_n|x_{t1}-x_{t2}| = b_n x_{t1}\left|1 - \frac{1}{1-(1-\theta^{(0)})\frac{A_t}{|B_t|}}\right|$$

and on the set $[b_n x_{t1} \leq k]$ this difference is bounded by

$$k\left|1 - \frac{1}{1-(1-\theta^{(0)})\frac{A_t}{|B_t|}}\right|.$$

Thus for some $\delta' > 0$

$$[B_t < 0, b_n x_{t1} \leq k, |x_{t1}-x_{t2}| \geq \delta] \subset [B_t < 0, b_n x_{t1} \leq k, \frac{A_t}{|B_t|} > \delta'].$$

On $[b_n x_{t1} \leq k]$, we have

$$\frac{1}{|B_t|} \leq \frac{k}{b_n C_t}$$

which implies

$$\frac{A_t}{|B_t|} \leq k\frac{A_t}{b_n C_t}.$$

Thus using the definitions of A_t, C_t we have the bound for some new δ''

$$\begin{aligned}II \leq& nP[B_t < 0, \frac{Z_{t-2}+\theta^{(0)}Z_{t-3}}{b_n(Z_t+(\theta^{(0)})^3 Z_{t-3})} > \delta'']\\ \leq& nP[B_t < 0, \frac{Z_{t-2}+\theta^{(0)}Z_{t-3}}{b_n Z_t} > \delta'', \frac{Z_{t-2}+\theta^{(0)}Z_{t-3}}{b_n(\theta^{(0)})^3 Z_{t-3}} > \delta''].\end{aligned}$$

On the set $[B_t < 0]$, we have $(\theta^{(0)})^{-1}Z_{t-1} \geq Z_{t-2}+\theta^{(0)}Z_{t-3}$ and so

$$\begin{aligned}II \leq& nP[\frac{Z_{t-1}}{b_n\theta^{(0)}Z_t} > \delta'', \frac{Z_{t-2}+\theta^{(0)}Z_{t-3}}{b_n(\theta^{(0)})^3 Z_{t-3}} > \delta'']\\ \leq& nP[\frac{Z_{t-1}}{b_n\theta^{(0)}Z_t} > \delta'']P[\frac{Z_{t-2}+\theta^{(0)}Z_{t-3}}{b_n(\theta^{(0)})^3 Z_{t-3}} > \delta'']\\ \sim& (const)(\delta'')^{-\alpha}P[\frac{Z_{t-2}+\theta^{(0)}Z_{t-3}}{b_n(\theta^{(0)})^3 Z_{t-3}} > \delta'']\end{aligned}$$

(by a result of Breiman, 1965 which is applicable from the second part of Condition R)

$$\to 0, \qquad (n \to \infty)$$

which shows $II \to 0$.

Observe that what was proven in the treatment of II is that for any constants $k > 0$ and $\delta > 0$ that

$$\lim_{n\to\infty} nP[B_t < 0, b_n|x_{t1} - x_{t2}| > \delta, b_n x_{t1} \le k] = 0. \tag{3.23}$$

For III, we note that

$$\begin{aligned}
&\sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0, b_n x_{t2} \le k+\delta]} \\
&\qquad \le \sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0, b_n x_{t2} \le k+\delta, b_n|x_{t2}-x_{t1}|\le\delta]} \\
&\qquad\quad + \sum_{t=1}^{n} |f(b_n x_{t1}) - f(b_n x_{t2})| 1_{[B_t<0, b_n x_{t2} \le k+\delta, b_n|x_{t2}-x_{t1}|>\delta]} \\
&\qquad = S_1 + S_2.
\end{aligned}$$

So

$$III \le P[S_1 > \eta/6] + P[S_2 > \eta/6] = IIIa + IIIb.$$

Letting $\omega_f(\delta)$ be the modulus of continuity for f:

$$\omega_f(\delta) := \sup_{|x-y|\le\delta} |f(x) - f(y)|,$$

we have for $IIIa$ the bound

$$IIIa \le P[\omega_f(\delta) \sum_{t=1}^{n} 1_{[B_t<0, b_n x_{t1} \le k+2\delta]} > \eta/6],$$

and by the argument which showed (3.17), this probability converges to $(n \to \infty)$

$$P[\omega_f(\delta)\nu_{0,\infty}([0, k + 2\delta]) > \eta/6]$$

which converges to 0 as $\delta \to 0$ since $\omega_f(\delta) \to 0$.

For $IIIb$ we have by (3.23) that

$$\begin{aligned}
IIIb \le & nP[B_t < 0, b_n x_{t2} \le k + \delta, b_n|x_{t2} - x_{t1}| > \delta, b_n x_{t1} > 2k] + o(1) \\
= & o(1),
\end{aligned}$$

since $x_{t1} < x_{t2}$.

The proof of Theorem 3.1 is concluded by recalling that when $p_t(1 - \theta^{(0)}) < 0$

$$x_{t1} \le r_t^- \le x_{t2}$$

and

$$b_n(\hat\theta - \theta^{(0)}) = \bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} b_n r_t^- .$$

Since

$$\bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} b_n x_{t1} \le \bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} b_n r_t^- \le \bigwedge_{\substack{1\le t\le n \\ p_t(1-\theta^{(0)})<0}} b_n x_{t2},$$

and the two extremes converge to the same weak limit, the minimum in the middle converges to the same weak limit. The proof of Theorem 3.1 is complete.

4. The limit distribution in the left tail case for $q = 1$.

In this section we assume Conditions M and L hold. We continue to assume the order is $q = 1$ and we present the limit distribution for $\hat{\theta}$. We will discuss the following theorem whose proof parallels that of Theorem 3.1 for the right tail case.

Theorem 4.1. *Suppose $\{X_t\}$ is the MA(1) process given in (1.1) and that Conditions M, L hold. Suppose the true parameter is $\theta^{(0)} \in (0,1)$ and that F, the distribution of Z_1, is continuous. Let $q_n = a(n)^{-1}$ where $a(n)$ is the quantile function*

$$a(n) = F^{\leftarrow}(1/n).$$

Note $a(n) \to 0$. The estimator $\hat{\theta}$ given in Section 2 has a Weibull limit distribution: In $[0, \infty)$

$$a(\sqrt{n})^{-1}(\hat{\theta} - \theta^{(0)}) \Rightarrow \frac{(\theta^{(0)})^{3/2\alpha}}{c(\alpha)^{1/2\alpha}} \bigwedge_{\substack{1 \le k < \infty \\ Y_{k,1} > Y_{k,2}}} \frac{\Gamma_k^{1/2\alpha}}{|Y_{k,1} - (\theta^{(0)})^3 Y_{k,}|}, \tag{4.1}$$

where $\{Y_{k,1}, Y_{k,2}, k \ge 1\}$ are iid with common distribution F and

$$\Gamma_k = E_1 + \cdots + E_k, \quad k \ge 1,$$

is a sum of iid unit exponentially distributed random variables. The constant $c(\alpha)$ is defined by the Beta integral

$$c(\alpha) = \int_0^1 (1-s)^\alpha \alpha s^{\alpha - 1} ds.$$

The limit distribution of $\hat{\theta}$ is Weibull:

$$\lim_{n \to \infty} P[a(\sqrt{n})^{-1}(\hat{\theta} - \theta^{(0)}) \le x] = 1 - \exp\{-k x^{2\alpha}\}, \quad x > 0, \tag{4.2}$$

where

$$k = (\theta^{(0)})^{-3\alpha} c(\alpha) E\left(|Y_{k,1} - (\theta^{(0)})^3 Y_{k,2}|^{2\alpha} 1_{[Y_{k,1} > Y_{k,2}]} \right)$$

which is finite by Condition L. The convergence rate is $1/a(\sqrt{n})$.

Remark. In the right tail case the convergence rate was b_n which up to a slowly varying multiplicative factor is of order $n^{1/\alpha}$. However, under Condition L, the convergence rate is only $1/a(\sqrt{n})$ which up to a slowly varying multiplicative factor is of order $n^{1/2\alpha}$. The convergence rate is slowed by the presence of a moving average component.

The proof of Theorem 4.1 parallels that of Theorem 3.1 and is only outlined. Our plan of attack is to show first that $\min\{C_t/|B_t| : 1 \le t \le n, p_t(1 - \theta^{(0)}) < 0\}$ has the limit distribution given in (4.2) and then we show that $\hat{\theta}$ has in fact this limit distribution.

We begin with the following limit theorem which parallels Proposition 3.2. It is built on the observations that for $x > 0, y > 0$,

$$\begin{aligned} \lim_{n \to \infty} \sqrt{n} P[Z_t \le x, \frac{Z_t}{a(\sqrt{n})} \le y] &= y^\alpha \\ \lim_{n \to \infty} \sqrt{n} P[Z_t \ge x, \frac{Z_t}{a(\sqrt{n})} \le y] &= 0 \end{aligned}$$

and

$$\begin{aligned} \lim_{n \to \infty} n P[\frac{Z_t}{a(\sqrt{n})} \le x, \frac{Z_{t-1}}{a(\sqrt{n})} \le y] = \lim_{n \to \infty} \sqrt{n} P[\frac{Z_t}{a(\sqrt{n})} \le x] P[\frac{Z_{t-1}}{a(\sqrt{n})} \le y] \\ = x^\alpha y^\alpha. \end{aligned}$$

Proposition 4.2. *Suppose Conditions L and R hold. Define the measure*

$$\mu(dx) = \alpha x^{\alpha-1} dx, \quad x > 0.$$

We have

$$\begin{aligned} nP[Z_{t-i} \in dy_i, i = 0, \dots, 3; \frac{Z_t}{a(\sqrt{n})} \in dx_0, \frac{Z_{t-3}}{a(\sqrt{n})} \in dx_3] \\ \xrightarrow{v} \epsilon_0(dy_0)F(dy_1)F(dy_2)\epsilon_0(dy_3)\mu(dx_0)\mu(dx_3), \end{aligned} \tag{4.3}$$

in

$$E := \{(y_0, y_1, y_2, y_3, x_0, x_3) \in [0,\infty]^4 \times ([0,\infty]^2\}. \tag{4.4}$$

Furthermore, with E as defined in (4.4), we have in $M_p(E)$ that

$$N_n := \sum_{t=1}^{n} \epsilon_{(Z_t, Z_{t-1}, Z_{t-2}, Z_{t-3}, \frac{Z_t}{a(\sqrt{n})}, \frac{Z_{t-3}}{a(\sqrt{n})})} \Rightarrow N_\infty := \sum_k \epsilon_{(0, Y_{k1}, Y_{k2}, 0, j_{k1}, j_{k2})}, \tag{4.5}$$

where

$$\sum_k \epsilon_{(j_{k1}, j_{k2})}$$

is PRM with mean measure $\mu \times \mu$ on $[0,\infty)^2$.

Again, as in Section 3, we wish to only consider points corresponding to $p_t(1-\theta^{(0)}) < 0$ but because it is easier, we start by restricting attention to the part of the state space corresponding to $B_t < 0$. So we define

$$[B < 0] := \{(x_0, \dots, x_5) \in E : x_1 > \theta_0 x_2 + 2(\theta^{(0)})^2 x_3\}.$$

Further we need the maps

$$\begin{aligned} T_1(x_0, \dots, x_5) :=& x_4 + (\theta^{(0)})^3 x_5, \\ T_2(x_0, \dots, x_5) :=& | - x_1 + \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3|, \\ T(x_0, \dots, x_5) :=& (x_0, x_1, x_2, x_3, T_1(x_0, \dots, x_5), T_2(x_0, \dots, x_5)) \end{aligned} \tag{4.6}$$

with domains and ranges

$$T_1 : [B < 0] \mapsto [0,\infty), \quad T_2 : [B < 0] \mapsto (0, \infty]$$

and

$$T : [B < 0] \mapsto E' := [0,\infty]^4 \times [0,\infty) \times (0,\infty].$$

Also, T is continuous and T^{-1} maps compact sets into compact sets since for instance

$$\begin{aligned} &T^{-1}(\times_{i=0}^{4}[0, a_i] \times [a_5, \infty]) \\ &\quad= \{(x_0, \dots, x_5) \in E : x_i \le a_i, i = 0, \dots 3; x_4 + (\theta^{(0)})^3 x_5 \le a_4, a_5 \le x_1 - (\theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3)\}. \end{aligned}$$

From Proposition 4.2 and Proposition 3.3 we have

$$N_n|_{[B<0]} \Rightarrow N_\infty|_{[B<0]}$$

in $M_p([B<0])$. Applying Proposition 3.18, page 148, of Resnick, 1987 yields

$$N_n|_{[B<0]} \circ T^{-1} \Rightarrow N_\infty|_{[B<0]} \circ T^{-1}$$

in $M_p(E')$ where remember $E' = [0,\infty]^4 \times [0,\infty) \times (0,\infty]$. Written another way, this is

$$\sum_{t=1}^{n} 1_{[B<0]} \epsilon_{(Z_t, Z_{t-1}, Z_{t-2}, Z_{t-3}, \frac{C_t}{a(\sqrt{n})}, |B_t|)} \Rightarrow \sum_k 1_{[Y_{k1} > \theta^{(0)} Y_{k2}]} \epsilon_{(0, Y_{k1}, Y_{k2}, 0, j_{k1} + (\theta^{(0)})^3 j_{k2}, |-Y_{k1} + \theta^{(0)} Y_{k2}|)}$$

in $M_p(E')$.

Now define

$$\begin{aligned} NEG := \{ & (x_0, x_1, x_2, x_3, x_5) \in E' : (x_2 + \theta^{(0)} x_3)(1-\theta^{(0)})^2 \\ & + (-x_1 + \theta^{(0)} x_2 + 2(\theta^{(0)})^2 x_3)(1-\theta^{(0)}) + (x_0 + (\theta^{(0)})^3 x_3) < 0 \}. \end{aligned}$$

We get from parts (a) and (c) of Proposition 3.3 that

$$\begin{aligned} & \sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]} \epsilon_{(Z_t, Z_{t-1}, Z_{t-2}, Z_{t-3}, C_t/a(\sqrt{n}), |B_t|)} \\ & \qquad \Rightarrow N_\infty^\# = \sum_{k=1}^{\infty} 1_{[Y_{k1} > Y_{k2}]} \epsilon_{(0, Y_{k1}, Y_{k2}, 0, j_{k1} + (\theta^{(0)})^3 j_{k2}, |-Y_{k1} + \theta^{(0)} Y_{k2}|)}. \end{aligned}$$

Note that in the indicator on the right, the condition $[Y_{k1} > Y_{k2}]$ is equivalent to the condition

$$Y_{k2}(1-\theta^{(0)})^2 + (-Y_{k1} + \theta^{(0)} Y_{k2})(1-\theta^{(0)}) < 0.$$

Applying Lemma 3.4 yields

$$\sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]} \epsilon_{(C_t/a(\sqrt{n}), |B_t|)} \Rightarrow N_\infty^\# = \sum_{k=1}^{\infty} 1_{[Y_{k1} > Y_{k2}]} \epsilon_{(j_{k1} + (\theta^{(0)})^3 j_{k2}, |-Y_{k1} + \theta^{(0)} Y_{k2}|)}.$$

After an argument that verifies division between the two components is permitted we get

$$\sum_{t=1}^{n} 1_{[p_t(1-\theta^{(0)})<0]} \epsilon_{\frac{C_t}{a(\sqrt{n})|B_t|)}} \Rightarrow \sum_k 1_{[Y_{k1} > Y_{k2}]} \epsilon_{\frac{j_{k1} + (\theta^{(0)})^3 j_{k2}}{|-Y_{k1} + \theta^{(0)} Y_{k2}|}}, \tag{4.7}$$

in $M_p(([0,\infty)$. Now one finishes the derivation with a comparison argument as in Section 3.

The form of the limit in (3.1) is based on the fact that

$$\sum_k \epsilon_{(j_{k1}, j_{k2})}$$

is PRM with mean measure of $[0,x]$ equal to $(\theta^{(0)})^{-3\alpha} c(\alpha) x^{2\alpha}$ and so is

$$\sum_{k=1}^{\infty} \epsilon_{((\theta^{(0)})^{-3} c(\alpha))^{-1/2\alpha} \Gamma_k^{1/2\alpha}}.$$

The form of the Weibull limit is gotten from the usual argument that the minimum of the points is greater than x iff the point process has no points in $[0,x]$.

5. Concluding Remarks.

It is noteworthy that in contrast to the autoregressive case, the moving average estimators in the left tail case suffer a performance degradation depending on the order q of the model; no such degradation is present under condition R. From the results of Section 4 we see that the convergence rate for the estimator of the MA(1) parameter is $1/a(\sqrt{n})$ which is a regularly varying function of index $\alpha/2$. Contrast this to the convergence rate of the lp estimators in the autoregressive case which is regularly varying of index α. We anticipate that the convergence rate in the left tail case for MA(q) parameters will have index α/q. Thus under Condition L, a sharp penalty is paid for using models which have moving average components and the penalty increases as the order of the model increases. This is in contrast to results under Condition R and to the results found for lp estimators for autoregressive parameters.

The challenge now is to extend these results from the MA(1) case to more general moving average processes and then on to the general ARMA model. We anticipate that for ARMA models, the rate of convergence under Condition L of the lp estimators will suffer depending on the order of the moving average component.

References

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

Billingsley, P., *Convergence of Probability Measures*, Wiley, New York, 1968.

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

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

Feigin, P. and Resnick, S., *Estimation for autoregressive processes with positive innovations*, Stochastic Models **8** (1992), 479–498.

Feigin, P. and Resnick, S., *Limit distributions for limear programming time series estimators*, Stochastic Processes and their Applications **51** (1994), 135–166.

Feigin, P. Resnick, S. and Stărică, Cătălin, *Testing for Independence in Heavy Tailed and Positive Innovation Time Series*, To appear (1994).

Kallenberg, O., *Random Measures*, Third edition, Akademie-Verlag, Berlin, 1983.

Resnick, Sidney, *Extreme Values, Regular Variation, and Point Processes*, Springer-Verlag, New York, 1987.

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

E-mail: paulf@ie.technion.ac.il

Marie Kratz, Université René Descartes, Paris V U.F.R. de Mathématiques et Informatique 45, rue des Saints-Pères 75270 Paris Cedex 06, France

E-mail: kratz@math-info.univ-paris5.fr

Sidney I. Resnick, Cornell University, School of Operations Research and Industrial Engineering, ETC Building, Ithaca, NY 14853 USA

E-mail: sid@orie.cornell.edu