

SAMPLE CORRELATIONS OF INFINITE VARIANCE TIME SERIES MODELS: AN EMPIRICAL AND THEORETICAL STUDY¹

JASON COHEN, SIDNEY RESNICK
and GENNADY SAMORODNITSKY

Cornell University

School of Operations Research and Industrial Engineering

Ithaca, New York 14853 USA

jason@orie.cornell.edu, sid@orie.cornell.edu, gennady@orie.cornell.edu

(Received September, 1997; Revised January, 1998)

When the elements of a stationary ergodic time series have finite variance the sample correlation function converges (with probability 1) to the theoretical correlation function. What happens in the case where the variance is infinite? In certain cases, the sample correlation function converges in probability to a constant, but not always. If within a class of heavy tailed time series the sample correlation functions do not converge to a constant, then more care must be taken in making inferences and in model selection on the basis of sample autocorrelations. We experimented with simulating various heavy tailed stationary sequences in an attempt to understand what causes the sample correlation function to converge or not to converge to a constant. In two new cases, namely the sum of two independent moving averages and a random permutation scheme, we are able to provide theoretical explanations for a random limit of the sample autocorrelation function as the sample grows.

Key words: Sample Correlations, Infinite Variance, Time Series, Stable Processes, Moving Average.

AMS subject classifications: 62M10, 60E07, 60G70.

1. Introduction

For a stationary sequence $\{X_n, n = 0, \pm 1, \pm 2, \dots\}$ the classical definition of the sample correlation function is

¹Resnick's and Samorodnitsky's research was partially supported by NSF Grant DMS-97-04982 and NSA grant MDA904-H-1036 at Cornell University. Jason Cohen was supported by NSF DMS8-94-00535 at Cornell.

$$\hat{\rho}(h) = \frac{\sum_{j=1}^{n-h} (X_j - \bar{X})(X_{j+h} - \bar{X})}{\sum_{j=1}^n (X_j - \bar{X})^2}, \quad h = 0, 1, \dots,$$

where $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ is the sample mean. When the variance of X_n is finite, and the sequence is ergodic, $\hat{\rho}(h) \rightarrow \text{Correlation}(X_0, X_h)$ with probability 1, for every h . See Brockwell and Davis [1]. When heavy tails are present, and the variance of X_n is infinite, it makes little sense to center at \bar{X} and the following heavy tailed version of the sample correlation function is often used:

$$\hat{\rho}_H(h) = \frac{\sum_{j=1}^{n-h} X_j X_{j+h}}{\sum_{j=1}^n X_j^2}.$$

Many common models for the infinite variance case are based on α -stable random variables, $0 < \alpha < 2$, or, more generally, on random variables in the domain of attraction of α -stable random variables. Recall that a random variable Z is in the domain of attraction of an α -stable law if it has appropriate regularly varying tails; that is, if

$$\lim_{t \rightarrow \infty} \frac{P[|Z| > tx]}{P[|Z| > t]} = x^{-\alpha}, \quad x > 0$$

and the tails balance

$$\lim_{t \rightarrow \infty} \frac{P[Z > t]}{P[|Z| > t]} = p, \quad 0 \leq p \leq 1$$

$$\lim_{t \rightarrow \infty} \frac{P[Z \leq -t]}{P[|Z| > t]} = q = 1 - p,$$

where $0 < \alpha < 2$. See e.g. Feller [8]. If $\{Z_k\}$ is an iid sequence of random variables in the domain of attraction of an α -stable law, $0 < \alpha < 2$, then for an infinite order moving average

$$X_n = \sum_{j=0}^{\infty} \psi_j Z_{n-j}, \tag{1.1}$$

Davis and Resnick [3, 4] have shown under appropriate summability conditions on the coefficients $\{\psi_j\}$ that

$$\hat{\rho}_H(h) \xrightarrow{P} \rho(h)$$

where

$$\rho(h) = \frac{\sum_{j=0}^{\infty} \psi_j \psi_{j+h}}{\sum_{j=0}^{\infty} \psi_j^2}. \tag{1.2}$$

Observe that $\rho(h)$ in (1.2) is not the theoretical correlation which does not exist. However, Davis and Resnick [5] have produced an example of a bilinear time series where

$$(\hat{\rho}_H(1), \dots, \hat{\rho}_H(h)) \Rightarrow (L(1), \dots, L(h)),$$

in \mathbb{R}^h for any $h > 0$ where “ \Rightarrow ” denotes weak convergence and where $L(h)$ is a non-degenerate random variable.

For finite variance time series models, infinite order moving averages are dense in the sense that any empirical sample correlation function can be mimicked for any fixed number of lags by an appropriate autoregression. However, for infinite variance time series this is no longer the case and in fact most heavy tailed stationary processes are nonlinear, and, in many senses, very far from linear processes. See for example, Rosiński [11]. However, the study of the sample correlation of more general nonlinear heavy tailed stationary processes is only beginning, the required point process

and regular variation tools are still being polished, and researchers have only a limited intuition into the question of which classes of processes have the property that the sample correlations converge to a non-random limit. It is precisely to develop this kind of intuition that we undertook an experimental project. In Section 2 we describe the models we have simulated. In Section 3 the simulation results are presented. Section 4 deals theoretically with two of the models and shows why a random limit occurs for the sample acf for the models under considerations.

The statistical significance of whether the sample autocorrelation is asymptotically random is profound. For example, model selection techniques for heavy tailed autoregressions based on the AIC criterion as well as coefficient estimation techniques based on Yule-Walker estimation all rely on the sample autocorrelation function converging to a constant and when this is not the case, the mischief potential for misspecifying a model is great. When the sample acf is asymptotically random, new statistical tools and parametric models need to be developed. This difficulty, as discussed in Feigin and Resnick [7], is not academic as all examples known to us of non-simulated, real, heavy tailed data exhibits the disturbing characteristic that the sample acf plot computed for a subset of the data set is not stable as the subset varies within the full data set. For example, splitting the data into three disjoint subsets and plotting the sample acf for all three produces plots which look quite different from each other.

2. Models

There are several classes of heavy tailed processes used in literature. One is based on various modifications of linear time series. We present two such examples here, in subsections 2.1 and 2.5; note that the former example is much more “standard” than the latter. In subsection 2.2 we consider the standard ARCH(1) process. Finally, our remaining examples are those of stationary symmetric α -stable (S α S) processes, $0 < \alpha < 2$. The structure of these processes is fairly well understood, which makes them an attractive source of examples. See Samorodnitsky and Taqqu [13]. A S α S process can be represented in the form

$$X_n = \int_{\mathbb{E}} f_n(X) M(dx), \quad n = 0, \pm 1, \pm 2, \dots, \quad (2.1)$$

where M is a S α S random measure on \mathbb{E} with a σ -finite control measure m , and $f_n \in L^\alpha(m)$ for all n . Only very special choices of the kernel f_n will produce a stationary S α S process (Rosiński [11]), and even more special kernels are needed to produce ergodic stationary S α S processes (Rosiński and Samorodnitsky [12]). Two examples of stationary ergodic S α S processes are presented in subsections 2.3 and 2.4; once again the former example is much more “standard” than the latter.

2.1 Sum of Two Moving Averages

The simplest possible departure from the linear moving average process is, of course, just a sum of two such independent processes. We simulate

$$X_n = \sum_{j=1}^{10} \frac{1}{(j+1)^2} Z_{n-j}^{(1)} + \sum_{j=1}^{10} \frac{1}{(j+1)^3} Z_{n-j}^{(2)}, \quad n = 1, 2, \dots, \quad (2.2)$$

where $\{Z_j^{(i)}, j = 0, \pm 1, \dots\}$, $i = 1, 2$ are independent sequences of iid Pareto (α) random variables. That is, $P(Z_j^{(i)} > \lambda) = \lambda^{-\alpha}$, $\lambda \geq 1$. We have chosen $\alpha = 1.5$. Of course, we could have used α -stable random variables in place of Pareto ones but this will not change the nature of the results, and so the ease with which Pareto random variables can be generated determined our choice.

Note that the choice of the coefficients in (2.2) is, basically, arbitrary. One only has to make sure that the two sequences are not proportional to each other since in that case the process reduces to the usual moving average. Any other choice of the coefficients does not change the nature of the results. Theoretical discussion of this example appears in Section 4.1.

2.2 ARCH(1)

The ARCH(1) process is defined by

$$X_n = \xi_n(a + bX_{n-1}^2)^{\frac{1}{2}}, \quad n = 1, 2, \dots, \tag{2.3}$$

where $\{\xi_n, n = 1, 2, \dots\}$ are iid standard normal random variables, independent of X_0 , $a > 0$ and $0 < b < 1$. Of course, only a particular choice of the initial distribution (that of X_0) will make this random recursion stationary. Instead, in simulation we start the process at 0, and discard the first 1000 observations to eliminate the initial transient in the system.

For this simulation, we used $a = 1$ and $b = 0.99$ which gives $P(X_n > \lambda) \sim C\lambda^{-1.014}$ as $\lambda \rightarrow \infty$. See de Haan et al. [6]. One of the major differences between this process and those based on linear models is that, in an ARCH process, heavy tails appear not because of an innovation with heavy tails, but due to the combined effect of infinitely many light tailed innovations.

2.3 Mixed Moving Average

A mixed moving average process represents yet another step away from a linear moving average process. We present it in the context of S α S processes. A mixed moving average S α S process can be written in the form

$$X_n = \int_{-\infty}^{\infty} \int_{\mathbb{E}} f(n-x, s) M(dx, ds), \tag{2.4}$$

$n = 0, \pm 1, \pm 2, \dots$, and M is now a S α S random measure on $\mathbb{R} \times \mathbb{E}$ with a σ -finite control measure $\text{Leb} \times m$, m being a σ -finite measure on \mathbb{E} . The function f is in $L^\alpha(\text{Leb} \times M)$. This process is ergodic (even mixing), see Surgailis et al. [14].

We have simulated a mixed moving average process with $\mathbb{E} = (0, 1)$, $m = \text{Leb}$ and

$$f(x, s) = e^{-|x|s^{\frac{1}{2}}}, \quad 0 \leq s \leq 1, \quad x \in \mathbb{R}.$$

It is not straightforward to simulate a general S α S process; we use the series representation. If G is an arbitrary real random variable with an everywhere positive density h , and U is an \mathbb{E} -valued random variable whose law λ is equivalent to the measure m , with $g = \frac{d\lambda}{dm}$, then the mixed moving average process (2.4) can be (in distribution) represented as

$$X_n = a_\alpha \sum_{j=1}^\infty \varepsilon_j \Gamma_j^{-\frac{1}{\alpha}} f(n - G_j, U_j) (h(G_j)g(U_j))^{-\frac{1}{\alpha}} \tag{2.5}$$

where a_α is a positive constant that depends only on α , and $\{\varepsilon_j\}$ are iid Rademacher, i.e., $P(\varepsilon_j = 1) = P(\varepsilon_j = -1) = \frac{1}{2}$, $\{\Gamma_j\}$ are the arrival times of a unit rate Poisson process; i.e. $\Gamma_j = \Gamma_{j-1} + e_j$ where $\{e_j\}$ are iid $\text{Exp}(1)$, $\{U_j\}$ are iid λ -distributed, $\{G_j\}$ are iid with density h .

All 4 sequences of random variables above are independent.

Even though the representation (2.5) is valid for every choice of h and λ as above, the practical necessity of truncating the sum in (2.5) at a finite number of terms makes the choice of h and λ an important one. In our case, for instance, the choice of the normal density $h(x) = (\sqrt{2\pi})^{-1} \exp(-x^2/2)$ initially looks attractive, but it has a major drawback: it takes an *extremely large* number of terms in the series to make sure that we observe a G_j close to even a moderately large n , and so the simulation program is likely to return values of X_n for such n 's as almost zero. To rectify the situation we choose heavy tailed G_j 's. For the purpose of this simulation we have chosen a Cauchy density

$$h(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}.$$

Moreover, we have chosen U_j 's to be uniformly distributed in $(0, 1)$. Then our simulation program uses the approximation

$$X_n = \sum_{j=1}^M \varepsilon_j \Gamma_j^{-\frac{1}{\alpha}} e^{-|n-G_j|U_j^{1/2}} (1+G_j^2)^{\frac{1}{\alpha}}$$

for a large M . We have dropped multiplicative constants that do not affect sample correlations.

2.4 Random Walk

There is only one class of mixing stationary S α S processes different from the mixed moving average processes considered in the previous subsection, that has been discussed in the literature. A representative of this class is simulated in this subsection. We refer the reader to Rosiński and Samorodnitsky [12] for more information.

Let m be the σ -finite measure on $\mathcal{Z}^{\mathcal{Z}}$ induced by mixing with respect to the counting measure on $m \in \mathcal{Z}$ probability measures induced on $\mathcal{Z}^{\mathcal{Z}}$ by a simple symmetric random walk passing through m at time 0. Let M be a S α S random measure on $\mathcal{Z}^{\mathcal{Z}}$ with control measure m . For any set $A \subset \mathcal{Z}$ with $m\{\omega \in \mathcal{Z}^{\mathcal{Z}}: \omega_n \in A\} < \infty$, the S α S process

$$X_n = \int_{\mathcal{Z}^{\mathcal{Z}}} \mathbf{1}(\omega_n \in A) M(d\omega), \quad n = 0, \pm 1, \pm 2, \dots$$

is well defined mixing stationary process.

For the purpose of this simulation we have chosen $A = \{0\}$. Since the process has to be, once again, simulated through its series representation, there is an issue of

selecting a probability measure on $\mathcal{Z}^{\mathcal{Z}}$ equivalent to m . In the present context this has an intuitive interpretation of selecting a probability law on \mathcal{Z} with all positive probabilities, according to which the initial position of the random walk is chosen. Unlike the previous example, in this case it is not so obvious why one choice of such a distribution will perform, when the series representing the process is truncated, better than another such distribution. Nevertheless, it still seems that choosing the initial state according to a heavy tailed distribution will “mix” the random walks better than a light tailed initial distribution will, and so the approximate process will be closer to true stationarity in the former case than in the latter. To get a feeling of whether this is, in fact, so, we have simulated this process twice, once with the initial state chosen according to

$$p_m = \frac{1}{3} 2^{-|m|}, \quad m = 0, \pm 1, \pm 2, \dots, \tag{2.6}$$

leading to simulating the series

$$X_n = \sum_{j=1}^M \varepsilon_j \Gamma_j^{-\frac{1}{\alpha}} 2^{|Y_0^j|/\alpha} \mathbf{1}(Y_n^j = 0), \quad n = 0, 1, 2, \dots,$$

where, as before,

- $\{\varepsilon_j\}$ are iid Rademacher,
- $\{\Gamma_j\}$ are the arrival times of a unit rate Poisson processes,
- $\{Y_n^j, n \in \mathcal{N}\}$ are iid simple symmetric random walks with initial distribution given by (2.6).

The number of terms M is large. The second choice of the initial distribution is that of a heavy tailed one, with

$$p_m = \frac{c}{(|m| + 1)^2}, \quad m = 0, \pm 1, \pm 2, \dots, \tag{2.7}$$

with $c = 3/(\pi^2 - 3)$. This leads to having to simulate the series

$$X_n = \sum_{j=1}^M \varepsilon_j \Gamma_j^{-\frac{1}{\alpha}} (|Y_0^j| + 1)^{\frac{2}{\alpha}} \mathbf{1}(Y_n^j = 0), \quad n = 0, 1, 2, \dots,$$

where, this time, the initial state Y_0^j of a simple symmetric random walk has distribution (2.7). Once again, the number of terms M in the series is large.

2.5 Coefficient Permutation

Our final example represents yet another modification of the linear time series. Let $\{\psi_j\}$ be a doubly infinite sequence of coefficients and $\{Z_j\}$ a random noise sequence such that the series $\sum_{j=-\infty}^{\infty} \psi_j Z_{-j}$ converges. The linear process (1.1) can be viewed as follows. Start with a sequence $\{\psi_j\}$ such that $\psi_j = 0$, for $j < 0$ and define

$$X_0 = \sum_{j=-\infty}^{\infty} \psi_j Z_{-j}.$$

To find X_1 we apply a shift to the sequence of coefficients

$$\psi^{(0)} = \{\dots, \psi_{-1}, \psi_0, \psi_1, \dots\} \mapsto \psi^{(1)} = \{\dots, \psi_0, \psi_1, \psi_2, \dots\}$$

so that

$$\psi_j^{(1)} = \psi_{j-1}^{(0)}$$

and compute

$$X_1 = \sum_{j=-\infty}^{\infty} \psi_j^{(1)} Z_{-j}.$$

To compute then X_2 , apply the shift to $\psi^{(1)}$, etc. Our idea was to use an operation on the sequence of coefficients different from the pure shift.

The operation on the sequences we have chosen for this example is a combination of a shift with a randomly chosen rearrangement of the coefficients. Since, in theory, we are dealing with an infinite sequence of coefficients, which makes it difficult to deal with permutations, we rearrange the coefficients by moving the first (non-zero) coefficient into a random position. Specifically, let $\{K_j\}$ and $\{M_j\}$ be two independent sequences of iid positive integer valued random variables. Given a sequence $\psi^{(0)} = \{\psi_0, \psi_1, \dots\}$, we define recursively

$$\psi^{(j)} = \{\psi_1^{(j-1)}, \dots, \psi_{K_j-1}^{(j-1)}, \psi_0^{(j-1)}, \psi_{K_j}^{(j-1)}, \dots\}, \tag{2.8}$$

for $j = 1, \dots, M_1 - 1$, and $\psi^{(M_1)} = \psi^{(0)}$. So for $j < M_1$, $\psi^{(j)}$ is obtained from $\psi^{(j-1)}$ by taking the initial entry of the sequence $\psi^{(j-1)}$ and moving it to the K_j th spot after displacing that entry one step to the right to clear room. We then continue the recursion (2.8) for $j = M_1 + 1, \dots, M_1 + M_2 - 1$, set $\psi^{(M_1 + M_2)} = \psi^{(0)}$, etc. The reasons for “resetting” the coefficients back to their initial state $\psi^{(0)}$ from time to time is that without such an action, the vector of coefficients tends to zero and the resulting process would be very difficult to simulate.

Having constructed the sequences $\psi^{(j)}$, $j \geq 0$, we define the *permutation process* by

$$X_n = \sum_{i=0}^{\infty} \psi_i^{(n)} Z_{n-i}, \quad n = 0, 1, 2, \dots \tag{2.9}$$

The random noise in (2.9) is independent of the randomness involved in constructing $\psi^{(j)}$, $j \geq 0$.

For the purpose of this simulation we have chosen

$$\psi_i^{(0)} = \frac{1}{(1+i)^2}, \quad i = 0, 1, 2, \dots,$$

$$P(K_j = k) = \frac{1}{k(k+1)}, \quad k = 1, 2, \dots,$$

$$P(M_j = k) = 2^{-k}, \quad k = 1, 2, \dots,$$

$\{Z_j\}$ are iid symmetrized Pareto random variables with $\alpha = 1.5$ (i.e., a product of a Pareto random variable and an independent Rademacher random variable).

Of course, the series in (2.9) has to be truncated as well, so we actually simulate

$$X_n = \sum_{i=0}^M \psi_i^{(n)} Z_{n-i}, \quad n = 0, 1, 2, \dots$$

for some large M . In particular, only the first M coefficients get permuted. If a particular K_j takes a value exceeding M , we discard this value and generate K_j anew.

3. Results

For all the examples we present time series plots of several runs and the corresponding sample correlations computed from these runs.

3.1 Sum of Two Moving Averages

The 9 sample autocorrelations shown in Figure 1 show enough variation that one must suspect that for the sum of two independent moving average processes, the sample correlations do not converge to a constant limit. The 9 time series plots are also given in Figure 2 and look rather different. This result may be somewhat counterintuitive for some since the sum of two independent linear processes behaves, in many respects, similarly to a linear process. A theoretical analysis of this case is presented in the next section where we verify that the sample correlation function converges in law to a nondegenerate limit.

We have generated 10000 observations in each run.

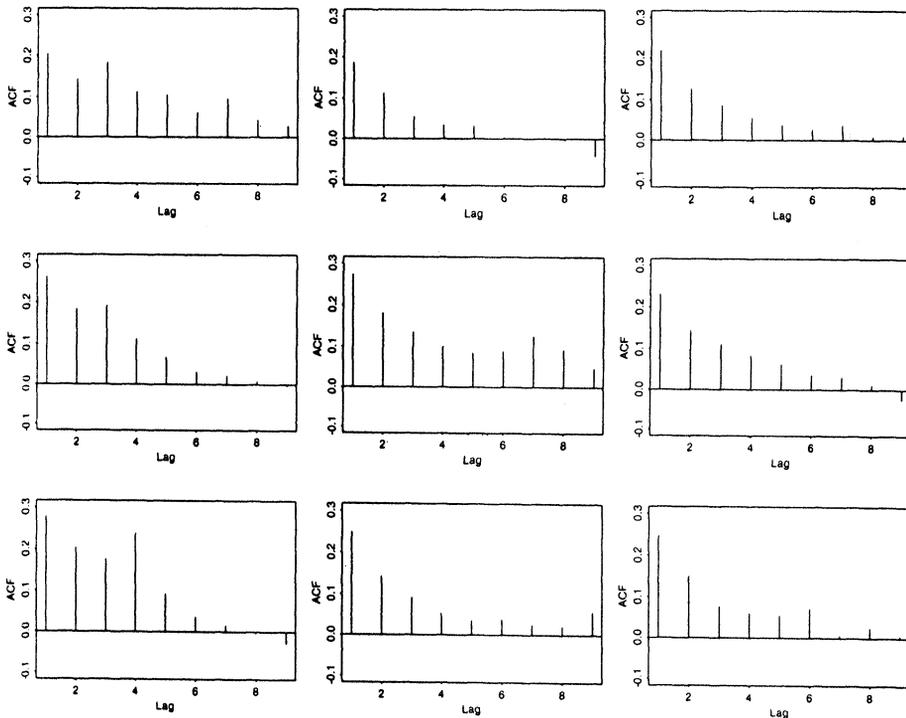


Figure 1. Sum of Two Moving Averages Auto-Correlation Function

Sum MA: Length = 1000 Alpha = 1.5

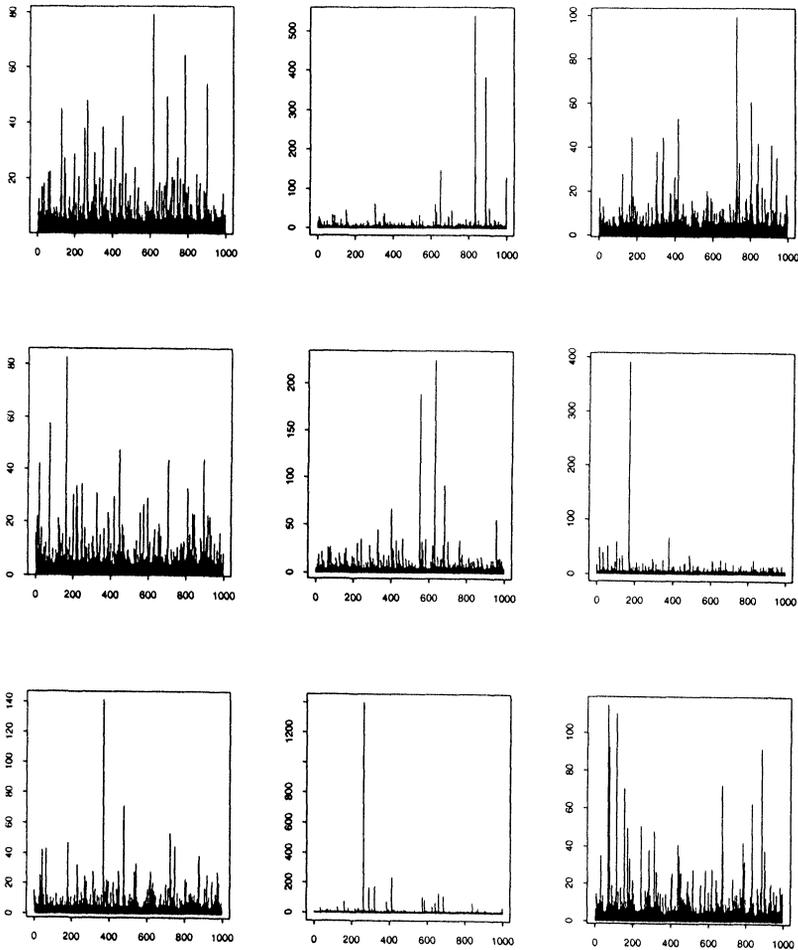


Figure 2. Sum of Two Moving Averages Time Series

3.2 ARCH(1)

For the process (2.3) we generated 100000 observations for each run, while dropping the first 1000 observations to reduce the effect of the possible initialization bias. The sample correlations in Figure 3 seem to indicate, in this case, that the convergence, if present, is to a non-degenerate (random) limit. This fact has recently been investigated by Davis and Mikosch [2].

Arch: Length = 100000 A = 1.00 B = 0.99 Extra = 1000

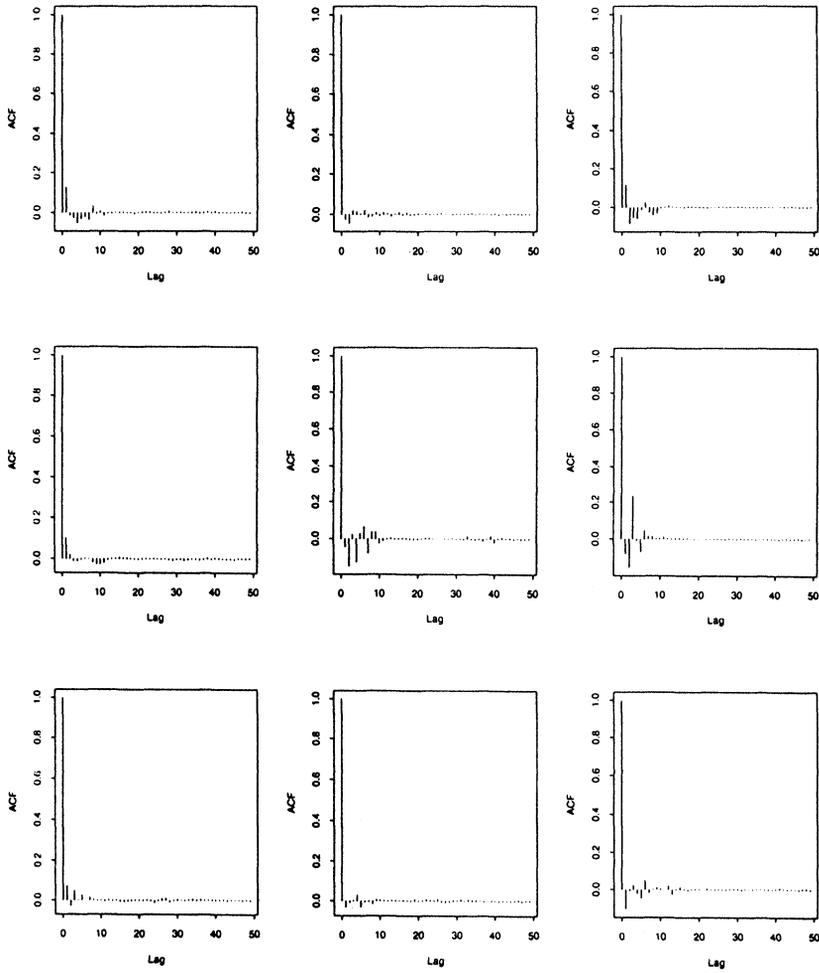


Figure 3. Arch Auto-Correlation Function

Arch: Length = 1000 A = 1.00 B = 0.99 Extra = 1000

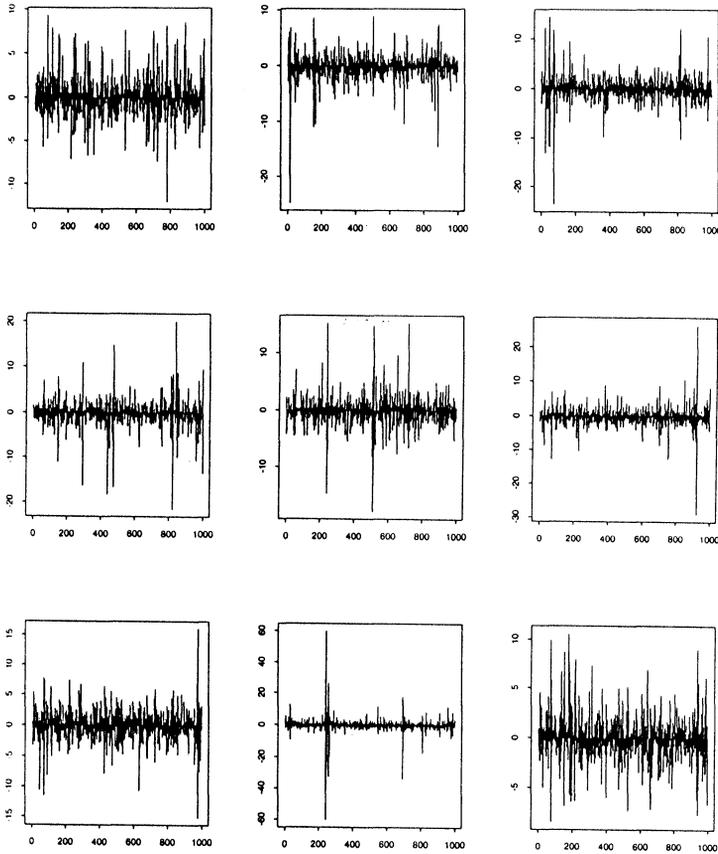


Figure 4. Arch Time Series

3.3 Mixed Moving Average

To simulate this process we have chosen the number of terms $M = 10^5$, and each of the 9 simulation runs was of length 1000. The results, presented in Figures 5 and 6, seem to clearly indicate that the sample acf does not converge to a constant. To check this unexpected conclusion we have generated also 4 additional simulation runs with $M = 10^6$, each run having length 10^5 . The resulting ACF, presented in Figure 7, seems to support the above conclusion. We have used $\alpha = 1.5$ throughout.

Mixed MA: Length = 1000 Sum-Len = 100000 Alpha = 1.5

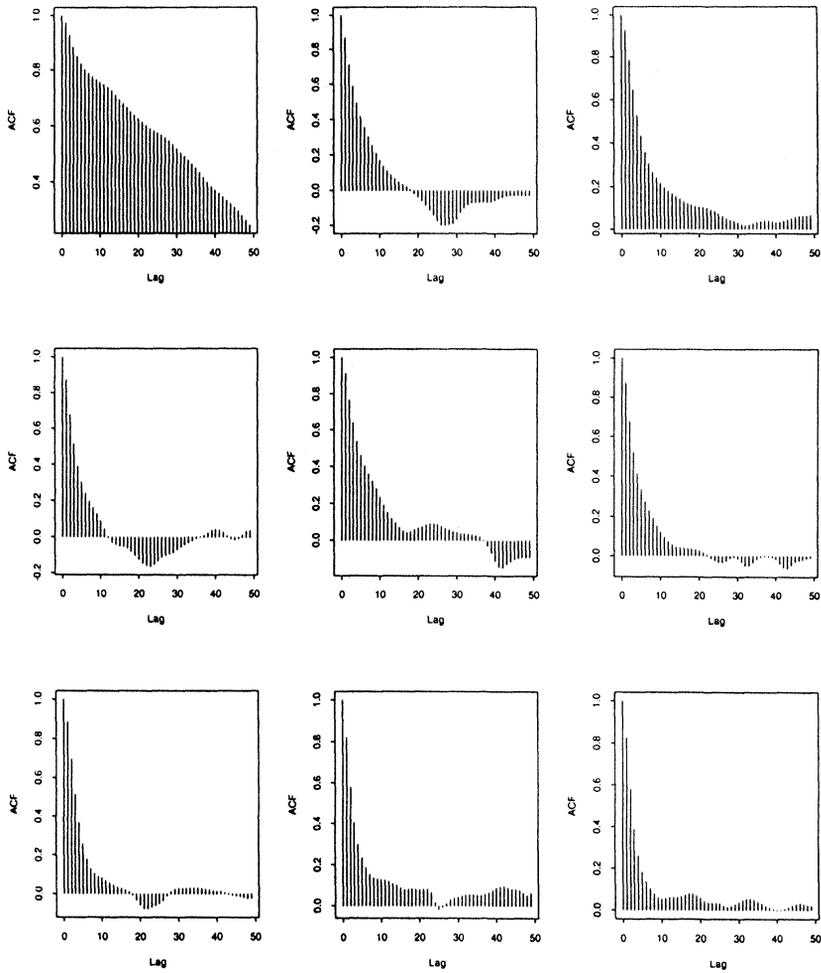


Figure 5. Mixed Moving Average Auto-Correlation Function

Mixed MA: Length = 1000 Sum-Len = 100000 Alpha = 1.5

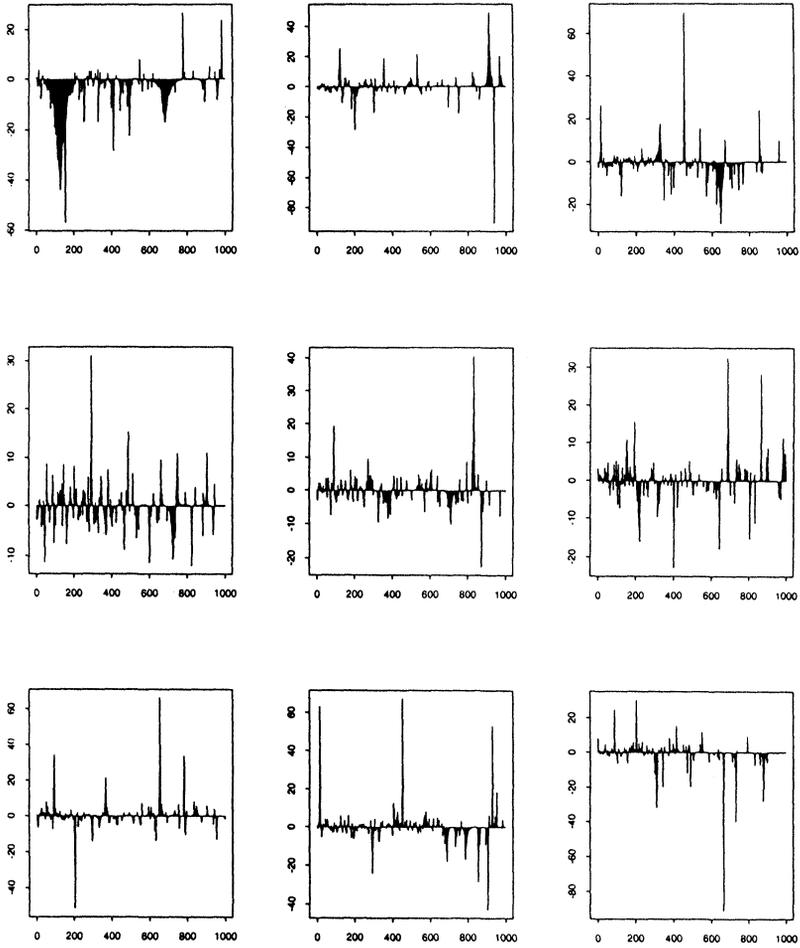


Figure 6. Moving Average Time Series

Mixed MA: Length = 10000 Sum-Len = 1000000 Alpha = 1.5

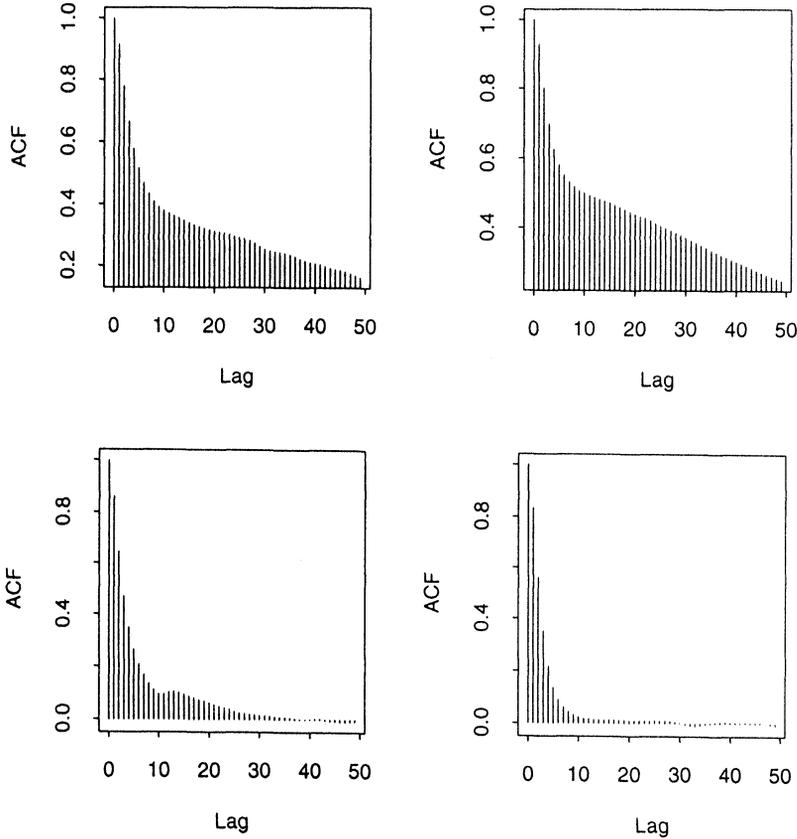


Figure 7. Mixed Moving Average Auto-Correlation Function

3.4 Random Walk

For this model we generated two batches of runs; the first one used the light tailed initial distribution (2.6), while the second one used the heavy tailed initial distribution (2.7). Once again we have used $M = 10^5$ terms in the series representation of the process, simulated each run of length 10^4 , and used $\alpha = 1.5$ throughout. One of the conclusions is that, as the inspection of the time series plots seems to indicate, we see less “action” going on towards the end of the plot than at its beginning when using the light tailed initial distribution (2.6). This phenomenon is not seen in the case of the heavy tailed initial distribution (2.7). Secondly, the sample correlations appear, once again, to converge to a non-degenerate limit, even though this phenomenon is not as obvious here, as it is in the previous example. See Figures 8-11.

3.5 Coefficient Permutation We have chosen here $M = 10^5$, and each simulation run was of length 2000. The sample correlations look “very random” in Figure 12, implying that, for this new class of stochastic processes, certain phenomena occur which require explanation. As we will show in Section 4.2, the randomness in the limit is due to the random coefficients.

RW 1: Length = 10000 Sum-Len = 100000 Alpha = 1.5

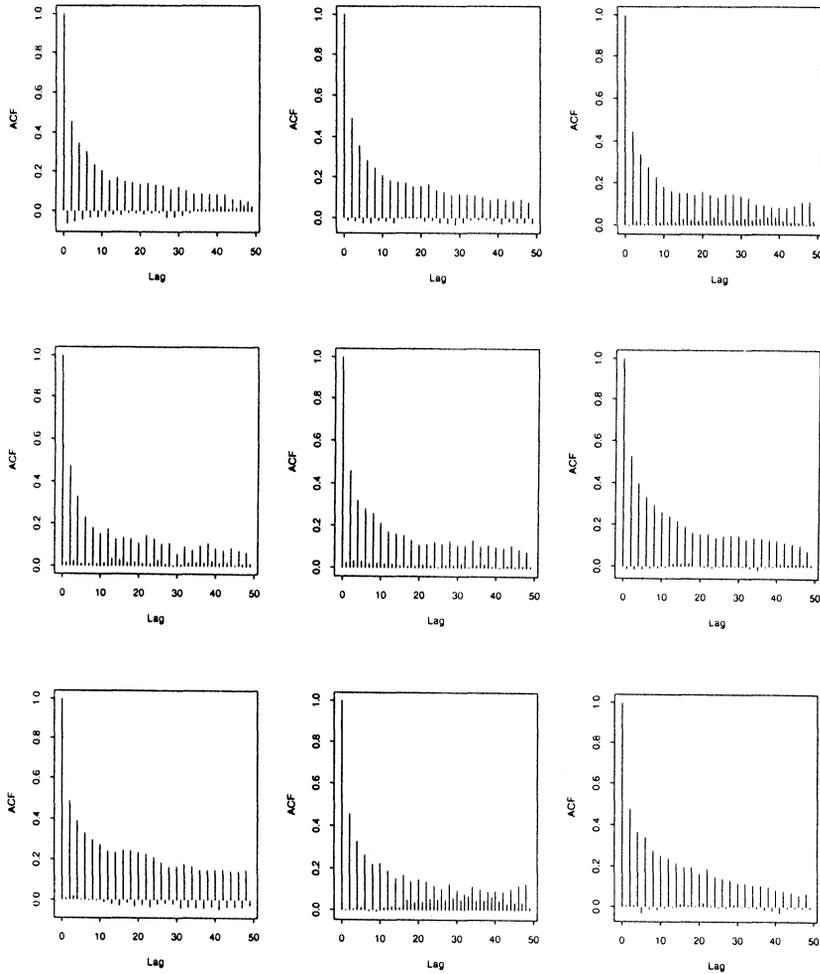


Figure 8. Random Walk Auto-Correlation Function
Light Tailed Initial Distribution

RW 1: Length = 1000 Sum-Len = 10000 Alpha = 1.5

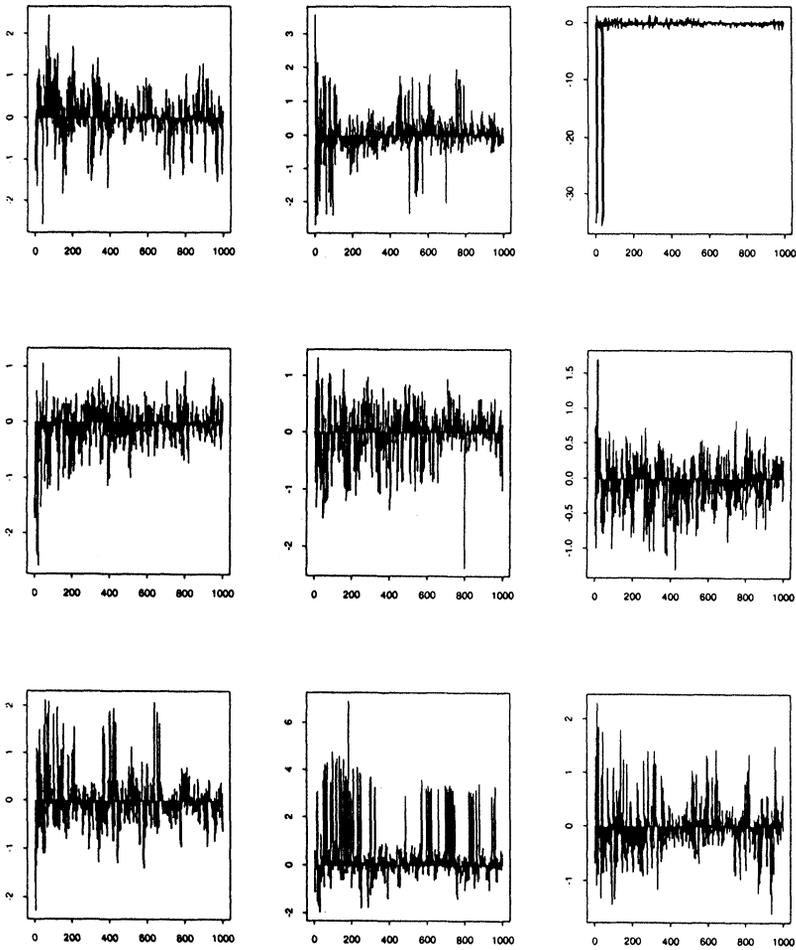


Figure 9. Random Walk Time Series
Light Tailed Initial Distribution

RW 2: Length = 10000 Sum-Len = 100000 Alpha = 1.5

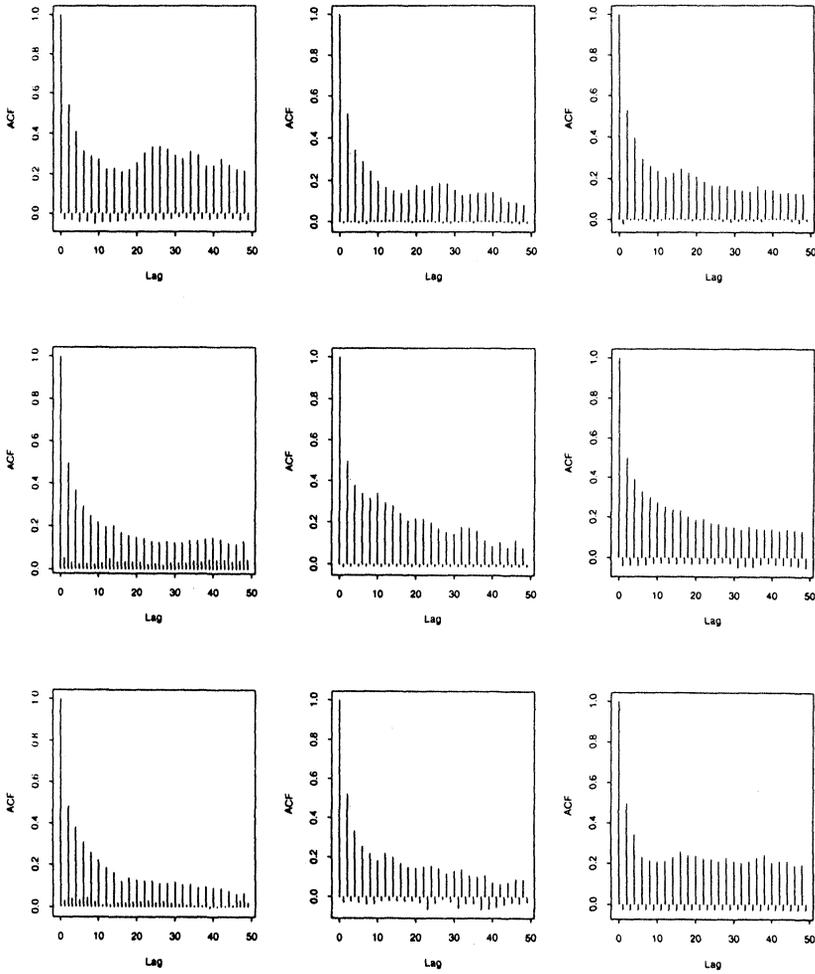


Figure 10. Random Walk Auto-Correlation Function Heavy Tailed Initial Distribution

RW 2: Length = 1000 Sum-Len = 10000 Alpha = 1.5

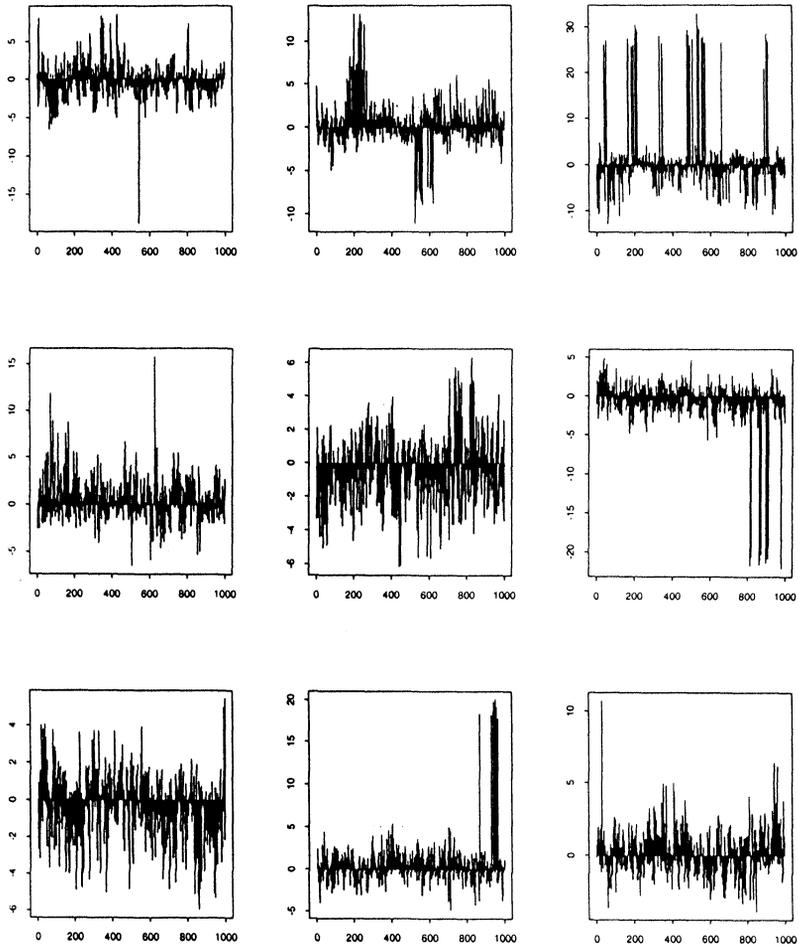


Figure 11. Random Walk Time Series
Heavy Tailed Initial Distribution

Reset: Length = 2000 Sum-Len = 100000 Alpha = 1.5

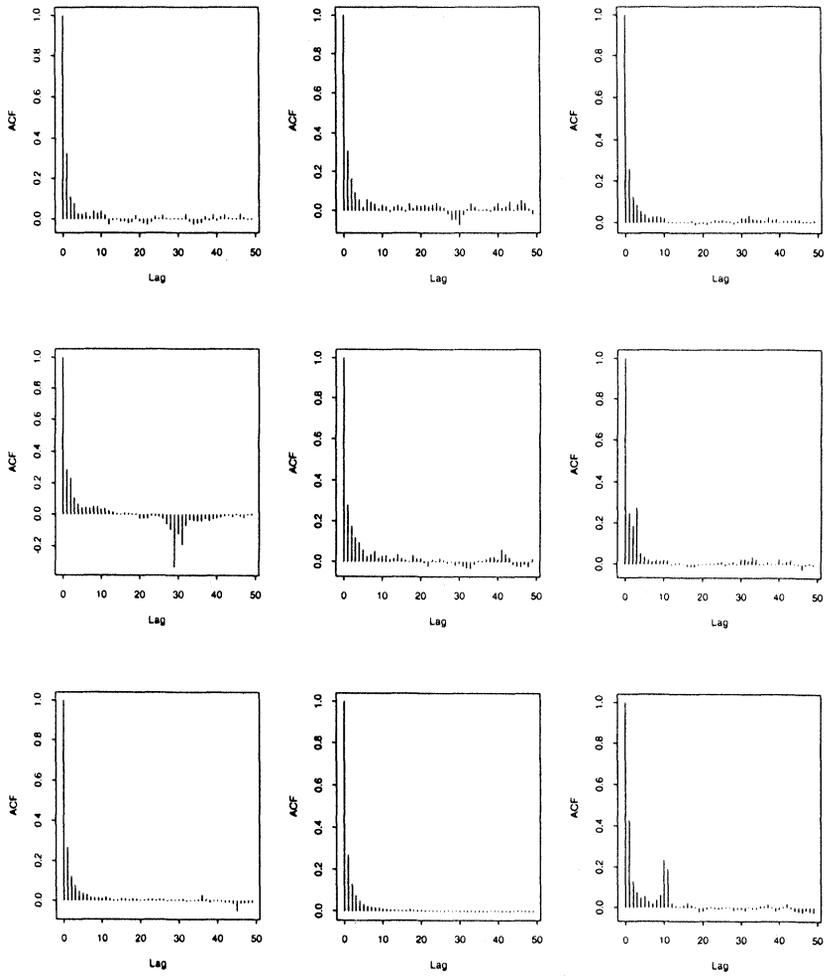


Figure 12. Coefficient Permutation with Reset Auto-Correlation Function

Reset: Length = 1000 Sum-Len = 100000 Alpha = 1.5

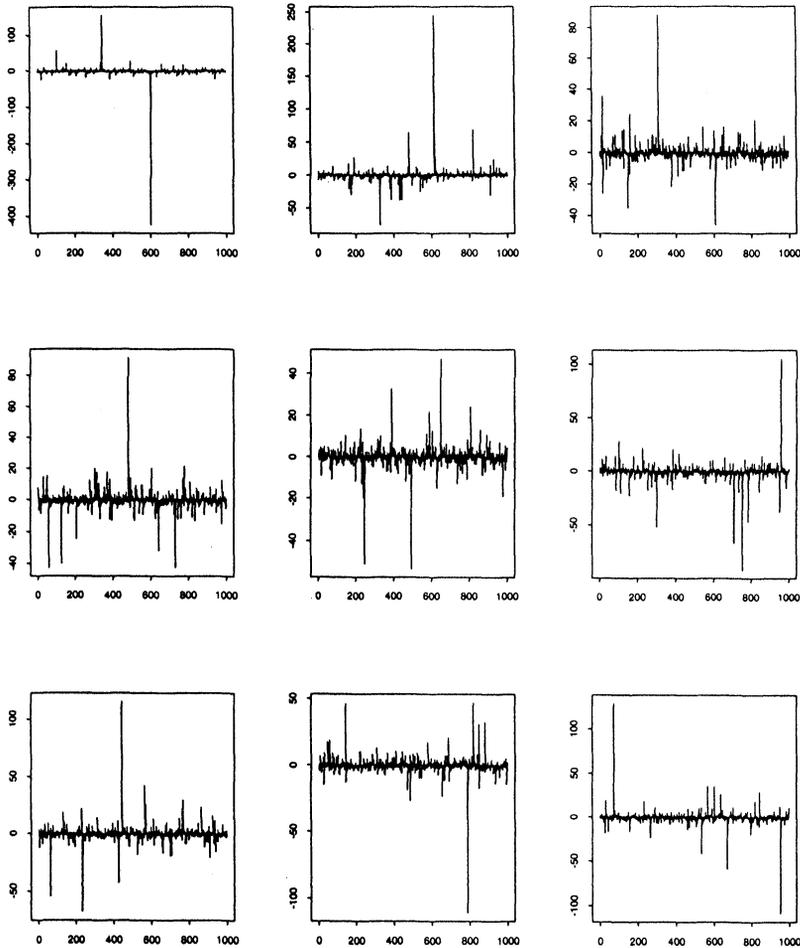


Figure 13. Coefficient Permutation with Reset Time Series

4. Analytical Results

This section is devoted to outlining explanations of why the sample correlation converges to a random limit for the cases discussed in Sections 2.1 and 2.5, namely for the sum of two moving averages and for the coefficient permutation with reset example.

The methods of proof are standard and use the connection between point processes and regular variation as outlined in Resnick [9], Section 4.5. The method for time series was developed by Davis and Resnick [3, 4].

We denote a Radon point measure on a nice space \mathbb{E} by

$$\sum_i \varepsilon_{x_i}$$

where $x_i \in \mathbb{E}$ and the collection of all such point measures is called $M_p(\mathbb{E})$. The topology is generated by the vague metric, vague convergence is denoted by “ \xrightarrow{v} ” and the Borel σ -algebra is denoted $\mathcal{M}_p(\mathbb{E})$. Let $C_K^{\downarrow}(\mathbb{E})$ be the non-negative continuous functions with compact support on \mathbb{E} .

4.1 Sum of Moving Average Processes

We examine in more detail, the example of the sum of two moving averages. The fact that the lag 1 correlation differs substantially from simulation run to simulation run raises doubts that the sample correlation function converges to a constant in this case. We will prove the following.

Theorem 1: *Suppose*

$$\left\{ \{Z_n^{(i)}, -\infty < n < \infty\}, i = 1, 2 \right\}$$

are two independent sequences of iid random variables with the same common distribution F satisfying

$$\lim_{x \rightarrow \infty} \frac{1 - F(x)}{1 - F(x) + F(-x)} = p, \quad 0 \leq p \leq 1, \tag{4.1}$$

$$1 - F(x) + F(-x) = x^{-\alpha} L(x), \quad 0 < \alpha < 2, \quad x \rightarrow \infty, \tag{4.2}$$

where L is slowly varying. Let the moving average coefficients

$$\left\{ \{c_n^{(i)}, -\infty < n < \infty\}, i = 1, 2 \right\}$$

satisfy the condition for $i = 1, 2$

$$\sum_{n=-\infty}^{\infty} |c_n^{(i)}|^\delta < \infty, \text{ for some } 0 < \delta < \alpha \wedge 1. \tag{4.3}$$

Define the sum of moving average process $\{X_n\}$ by

$$X_n = \sum_{j=-\infty}^{\infty} c_j^{(1)} Z_{n-j}^{(1)} + \sum_{j=-\infty}^{\infty} c_j^{(2)} Z_{n-j}^{(2)}.$$

Then, for any integer k , the heavy tailed sample correlation function $\{\widehat{\rho}_H(h), 1 \leq h \leq k\}$ converges in distribution in \mathbb{R}^k to a nondegenerate limit random vector

$$\left(\frac{\xi^{(1)} \sum_{l=-\infty}^{\infty} c_l^{(1)} c_{l+h}^{(1)} + \xi^{(2)} \sum_{l=-\infty}^{\infty} c_l^{(2)} c_{l+h}^{(2)}}{\xi^{(1)} \sum_{l=-\infty}^{\infty} (c_l^{(1)})^2 + \xi^{(2)} \sum_{l=-\infty}^{\infty} (c_l^{(2)})^2}, 1 \leq h \leq k \right),$$

where $\xi^{(i)}, i = 1, 2$ are two independent, identically distributed, positive, strictly $\alpha/2$ -stable random variables.

Proof: We proceed in a series of steps.

Step 1: Define the quantile function

$$b_n = \left(\frac{1}{P[|Z_1^{(i)}| > \cdot]} \right)^{\leftarrow} (n) \tag{4.4}$$

so that for $x > 0$

$$nP[|Z_1^{(i)}| > b_n x] \rightarrow x^{-\alpha}, \quad n \rightarrow \infty$$

and in terms of measures

$$nP[b_n^{-1}Z_1^{(i)} \in \cdot] \xrightarrow{v} \nu \tag{4.5}$$

where the vague convergence is in the space of Radon measures on $[-\infty, \infty] \setminus \{0\}$ and for $x > 0$

$$\nu(x, \infty] = px^{-\alpha}, \quad \nu(-\infty, -x] = q|x|^{-\alpha} \tag{4.6}$$

and $q = 1 - p$.

Step 2: Observe that for $i, j = 1, 2$ and $x > 0, y > 0$

$$\lim_{n \rightarrow \infty} nP[b_n^{-1}Z_1^{(i)} > x, b_n^{-1}Z_2^{(j)} > y] = 0. \tag{4.7}$$

Fix an integer m . For any n , define the vector of length $2m + 1$

$$\mathbf{Z}_n^{(m, i)} = (Z_{n+j}^{(i)}, |j| \leq m)'$$

Because (4.7) says two components cannot be simultaneously large, we get the following vague convergence in $[-\infty, \infty]^{4m+2} \setminus \{\mathbf{0}\}$:

$$nP[b_n^{-1}(\mathbf{Z}_1^{(m, 1)}, \mathbf{Z}_1^{(m, 2)}) \in \cdot] \xrightarrow{v} v^{(m)} \times \varepsilon_0 + \varepsilon_0 \times v^{(m)}. \tag{4.8}$$

Here ε_0 is the probability measure concentrating all mass on the zero vector in $2m + 1$ dimensional space and $\nu^{(m)}$ concentrates all mass on the axes through $\mathbf{0}$ so that

$$\nu^{(m)}(dx_j, |j| \leq m) = \sum_{j=-m}^m \nu(dx_j) \times \varepsilon_0(dx_{-m}, \dots, d\hat{x}_j, \dots, dx_m)$$

where the hat denotes the deleted variable.

Step 3: The regular variation and balance conditions (4.2) and (4.1) imply (Resnick [9], page 226) for $i = 1, 2$ as $n \rightarrow \infty$, the weak convergence in the space $M_p([-\infty, \infty] \setminus \{0\})$:

$$\sum_{l=1}^n \varepsilon_{Z_l^{(i)}/b_n} \Rightarrow \sum_k \varepsilon_{j_k^{(i)}}$$

where for $i = 1, 2$, the limits are independent Poisson processes on $[-\infty, \infty] \setminus \{0\}$ with mean measures equal to ν . This has an extension using Step 2 as follows:

$$\sum_{l=1}^n \varepsilon_{b_n^{-1}(\mathbf{Z}_l^{(m, 1)}, \mathbf{Z}_l^{(m, 2)})} \Rightarrow \sum_k \sum_{|l| \leq m} \varepsilon_{(j_k^{(1)}, \mathbf{e}_l^0)} + \sum_k \sum_{|l| \leq m} \varepsilon_{(\mathbf{0}, j_k^{(2)}, \mathbf{e}_l)} \tag{4.9}$$

where \mathbf{e}_l is a vector of length $2m + 1$ with 1 in the l th spot and 0's elsewhere. This follows because no two components can be large simultaneously.

Step 4: Define for $i = 1, 2$

$$X_n^{(m, i)} = \sum_{|j| \leq m} c_j^{(i)} Z_{n-j}^{(i)}$$

Consider the map $\mathbb{R}^{2m+1} \times \mathbb{R}^{2m+1} \mapsto \mathbb{R} \times \mathbb{R}$ defined by

$$(\mathbf{z}^{(1)}, \mathbf{z}^{(2)}) \mapsto \left(\sum_{|j| \leq m} c_j^{(1)} z_j^{(1)}, \sum_{|j| \leq m} c_j^{(2)} z_j^{(2)} \right).$$

Apply this to the convergence in (4.9) which yields after a continuous mapping arguments that

$$\sum_{l=1}^n \varepsilon_{b_n^{-1}(X_l^{(m,1)}, X_l^{(m,2)})} \Rightarrow \sum_{|l| \leq m} \sum_k \varepsilon_{(j_k^{(1)} c_l^{(1)}, 0)} + \sum_{|l| \leq m} \sum_k \varepsilon_{(0, j_k^{(2)} c_l^{(2)})}$$

and after an argument which allows $m \rightarrow \infty$ we get

$$\sum_{l=1}^n \varepsilon_{b_n^{-1}(X_l^{(1)}, X_l^{(2)})} \Rightarrow \sum_{|l| < \infty} \sum_k \varepsilon_{(j_k^{(1)} c_l^{(1)}, 0)} + \sum_{|l| < \infty} \sum_k \varepsilon_{(0, j_k^{(2)} c_l^{(2)})}.$$

Add the components and we get our intermediate goal, that

$$\sum_{l=1}^n \varepsilon_{b_n^{-1} X_l} \Rightarrow \sum_{|l| < \infty} \sum_k \left(\varepsilon_{j_k^{(1)} c_l^{(1)}} + \varepsilon_{j_k^{(2)} c_l^{(2)}} \right). \tag{4.10}$$

Step 5: we now modify the procedure in Step 4. For a nonnegative integer h , we define a map from $\mathbb{R}^{2m+1} \times \mathbb{R}^{2m+1} \mapsto \mathbb{R}^4$ defined by

$$(\mathbf{z}^{(1)}, \mathbf{z}^{(2)}) \mapsto \left(\sum_{|i| \leq m} c_i^{(1)} z_i^{(1)}, \sum_{|i| \leq m} c_{i+h}^{(1)} z_i^{(1)}, \sum_{|i| \leq m} c_i^{(2)} z_i^{(2)}, \sum_{|i| \leq m} c_{i+h}^{(2)} z_i^{(2)} \right).$$

Apply this map to (4.9) and after letting $m \rightarrow \infty$ we obtain

$$\begin{aligned} & \sum_{l=1}^n \varepsilon_{b_n^{-1}(X_l^{(1)}, X_{l+h}^{(1)}, X_l^{(2)}, X_{l+h}^{(2)})} \\ & \Rightarrow \sum_{l,k} \varepsilon_{(j_k^{(1)} c_l^{(1)}, j_k^{(1)} c_{l+h}^{(1)}, 0, 0)} + \sum_{l,k} \varepsilon_{(0, 0, j_k^{(2)} c_l^{(2)}, j_k^{(2)} c_{l+h}^{(2)})} \end{aligned}$$

which yields after adding the first and third and then the second and fourth components that

$$\sum_{l=1}^n \varepsilon_{b_n^{-1}(X_l, X_{l+h})} \Rightarrow \sum_{l,k} \varepsilon_{(j_k^{(1)} c_l^{(1)}, j_k^{(1)} c_{l+h}^{(1)})} + \sum_{l,k} \varepsilon_{(j_k^{(2)} c_l^{(2)}, j_k^{(2)} c_{l+h}^{(2)})} \tag{4.11}$$

in $M_p([-\infty, \infty]^2 \setminus \{\mathbf{0}\})$.

The convergence in (4.11) holds for $h = 1, \dots, k$ for any k and in fact, examining the outline of the proof, we see that the convergence holds jointly. Apply the functional which takes the product of the components and then sums the points of the point measure to (4.11) when $h = 0$ and when h is arbitrary. We get after a truncation and continuity argument that in \mathbb{R}^{k+1}

$$\left(b_n^{-2} \sum_{l=1}^n X_l X_{l+h}, 0 \leq h \leq k \right)$$

$$\Rightarrow \left(\sum_{k,l} \left[(j_k^{(1)})^2 c_l^{(1)} c_{l+h}^{(1)} + (j_k^{(2)})^2 c_l^{(2)} c_{l+h}^{(2)} \right], 0 \leq h \leq k \right). \tag{4.12}$$

Divide the h th component by the zero-th component in (4.12) and we achieve

$$\begin{aligned} \hat{\rho}_H(h) &= \frac{\sum_{l=1}^n X_l X_{l+h}}{\sum_{l=1}^n X_l^2} \\ &\Rightarrow \frac{\xi^{(1)} \sum_l c_l^{(1)} c_{l+h}^{(1)} + \xi^{(2)} \sum_l c_l^{(2)} c_{l+h}^{(2)}}{\xi^{(1)} \sum_l (c_l^{(1)})^2 + \xi^{(2)} \sum_l (c_l^{(2)})^2} \end{aligned}$$

where

$$\xi^{(i)} = \sum_k (j_k^{(i)})^2.$$

This is the desired result since for $i = 1, 2, \xi^{(i)}$ is a positive strictly stable random variable with index $\alpha/2$. □

4.2 Coefficient Permutation with Reset

In this subsection we outline a justification for the empirical results of Section 3.5 and show why the sample correlation function converges to a random limit.

Recall that we start with a sequence $\psi = \{\psi_j, j \geq 0\}$. We assume this sequence satisfies the analogue of (4.3), namely

$$\sum_{n=0}^{\infty} |\psi_n|^\delta < \infty, \text{ for some } 0 < \delta < \alpha \wedge 1. \tag{4.13}$$

We also have an iid innovation sequence $\{Z_n\}$ whose common distribution F satisfies (4.1) and (4.2). Remember we get $\psi^{(1)}$ from ψ by shifting the initial entry of the sequence of the K_1 -th spot, displacing the entries with index greater or equal to K_1 one slot to the right. This is repeated until the reset time M_1 and so on.

Lemma 1: *Let $\{\psi^{(j)}, j \geq 0\}$ be the random elements of \mathbb{R}^∞ created by the random shift and reset scheme of Section 2.5 where $\psi^{(0)} = \psi$ and ψ satisfies (4.13). If $E(M_1) < \infty$ then there exists a random element $\psi^{(\infty)}$ of \mathbb{R}^∞ such that $n \rightarrow \infty$*

$$\psi^{(n)} \Rightarrow \psi^{(\infty)} \tag{4.14}$$

in \mathbb{R}^∞ where $\psi^{(\infty)}$ is specified as follows: Let $\{\tilde{\psi}^{(j)}\}$ be the sequence of random elements of \mathbb{R}^∞ constructed with shifts and no resets. Then for $B \in \mathfrak{B}(\mathbb{R}^\infty)$ we have

$$P[\psi^{(\infty)} \in B] = \sum_{j=0}^{\infty} P[\tilde{\psi}^{(j)} \in B] \left(\frac{P[M_1 > j]}{E(M_1)} \right).$$

Proof: The sequence of random elements $\{\tilde{\psi}^{(j)}\}$ is a regenerative sequence with expected cycle length $E(M_1)$ and the limit distribution follows from Smith's Theorem. See Resnick [10], Section 3.12.2. □

Before we explain why the sample correlation has a random limit for the shift and reset process, we need the following lemma. For this result, we set

$$\mathbb{E}_1 = [-\infty, \infty]^m \setminus \{\mathbf{0}\}, \quad \mathbb{E}_2 = \mathbb{R}^m.$$

Lemma 2: *Suppose*

$$x_i^{(n)} \in \mathbb{E}_1, \quad n \geq 1, \quad 1 \leq i \leq n,$$

and $\psi^{(n)} \in \mathbb{E}_2$.
 Suppose

(i) There exist $\{x_i^{(\infty)}, i \geq 1\}$ such that

$$\sum_{i=1}^n \varepsilon_{x_i^{(n)}} \xrightarrow{v} \sum_i \varepsilon_{x_i^{(\infty)}} \tag{4.15}$$

(ii) in $M_p(\mathbb{E}_1)$;
 For some $\psi^{(\infty)} \in \mathbb{E}_2$

$$\lim_{n \rightarrow \infty} \psi^{(n)} = \psi^{(\infty)}; \tag{4.16}$$

(iii) For any compact $K \subset \mathbb{E}_2$ and any fixed integer L we have

$$\lim_{n \rightarrow \infty} \sum_{i=1}^L \varepsilon_{x_i^{(n)}}(K) = 0, \tag{4.17}$$

so that for any $i, x_i^{(n)} \in K^c$ for all n large enough.

Then,

$$\sum_{i=1}^n \varepsilon_{(x_i^{(n)}, \psi^{(i)})} \xrightarrow{v} \sum_i \varepsilon_{(x_i^{(\infty)}, \psi^{(\infty)})} \tag{4.18}$$

in $M_p(\mathbb{E}_1 \times \mathbb{E}_2)$.

Proof: Let $f \in C_K^\dagger(\mathbb{E}_1 \times \mathbb{E}_2)$. Then there exists a large number k such that

$$f(x, y) = 0 \text{ if either } \|y\| > k \text{ or } \|x\| < k^{-1}$$

and if we set $K_1^c = \{x: \|x\| < k^{-1}\}$ we require additionally that k be chosen to satisfy

$$\sum_i \varepsilon_{x_i^{(\infty)}}(\partial K_1) = 0.$$

The desired conclusion (4.18) will follow if we show

$$\sum_{i=1}^n f(x_i^{(n)}, \psi^{(i)}) \rightarrow \sum_{i=1}^\infty f(x_i^{(\infty)}, \psi^{(\infty)}). \tag{4.19}$$

Write

$$\begin{aligned} & \left| \sum_{i=1}^n f(x_i^{(n)}, \psi^{(i)}) - \sum_{i=1}^\infty f(x_i^{(\infty)}, \psi^{(\infty)}) \right| \\ & \leq \left| \sum_{i=1}^n f(x_i^{(n)}, \psi^{(i)}) - \sum_{i=1}^n f(x_i^{(n)}, \psi^{(\infty)}) \right| \\ & \quad + \left| \sum_{i=1}^n f(x_i^{(n)}, \psi^{(\infty)}) - \sum_{i=1}^\infty f(x_i^{(\infty)}, \psi^{(\infty)}) \right| = I + II. \end{aligned}$$

However, for fixed $\psi^{(\infty)}$,

$$f(\cdot, \psi^{(\infty)}) \in C_K^\dagger(\mathbb{E}_1)$$

and so from (4.15) we get $II \rightarrow 0$ as $n \rightarrow \infty$.

It remains to show $I \rightarrow 0$. Recall $K_1 = \{x: \|x\| \geq k^{-1}\}$ is compact. From (4.15) there exists n_0 such that all $n \geq n_0$

$$m := \sum_{i=1}^n \varepsilon_{x_i^{(n)}}(K_1) = \sum_{i=1}^{\infty} \varepsilon_{x_i^{(\infty)}}(K_1). \tag{4.20}$$

Suppose $\varepsilon > 0$ is given. Since f is continuous with compact support, it is uniformly continuous and so there exists $\delta > 0$ such that

$$\sup\{ |f(x_1, y_1) - f(x_2, y_2)| : \|(x_1, y_1) - (x_2, y_2)\| \leq \delta\} \leq \frac{\varepsilon}{m}. \tag{4.21}$$

Choose L so big that for all $i \geq L$ we have

$$\|\psi^{(i)} - \psi^{(\infty)}\| \leq \delta.$$

We have

$$\begin{aligned} I &\leq \sum_{i=1}^n |f(x_i^{(n)}, \psi^{(i)}) - f(x_i^{(n)}, \psi^{(\infty)})| \\ &\leq \sum_{i=1}^L |f(x_i^{(n)}, \psi^{(i)}) - f(x_i^{(n)}, \psi^{(\infty)})| \\ &\quad + \sum_{i=L+1}^n |f(x_i^{(n)}, \psi^{(i)}) - f(x_i^{(n)}, \psi^{(\infty)})| = Ia + Ib. \end{aligned}$$

Suppose f is bounded above by $\|f\|$. Then

$$Ia \leq \|f\| \sum_{i=1}^L \varepsilon_{x_i^{(n)}}(K_1) \rightarrow 0$$

as $n \rightarrow \infty$ from (4.17). For Ib we get from (4.21) and (4.20) that

$$\begin{aligned} Ib &\leq \sup\{ |f(x_1, y_1) - f(x_2, y_2)| : \|(x_1, y_1) - (x_2, y_2)\| \leq \delta\} \cdot m \\ &\leq \frac{\varepsilon}{m} m = \varepsilon. \end{aligned}$$

We thus get

$$\limsup_{n \rightarrow \infty} I \leq \varepsilon,$$

which finishes the verification of (4.19). □

We are now in a position to show why the sample acf converges, in general, to a random limit.

Theorem 2: *Suppose $\{X_n\}$ is the shift-reset process described in Section 2.5. For any k , as $n \rightarrow \infty$*

$$(\hat{\rho}_H(h), 1 \leq h \leq k) \Rightarrow \left(\frac{\sum_{l=0}^{\infty} \psi_l^{(\infty)} \psi_{l+h}^{(\infty)}}{\sum_{l=0}^{\infty} (\psi_l^{(\infty)})^2}, 1 \leq h \leq k \right)$$

in \mathbb{R}^k where $\{\psi_l^{(\infty)}, l \geq 0\}$ is the process described in Lemma 1.

Note all the randomness in the limit is caused by the random coefficients and not by the innovation sequence. We will see that in the limit, the randomness caused by the Z 's cancels out.

Proof: As in the proof of Theorem 1, we proceed in a series of steps.

Step 1: For an integer m , define a vector of length $m + 1$ by

$$\mathbf{Z}_j^{(m)} = (Z_j, Z_{j-1}, \dots, Z_{j-m})'.$$

Define b_n as in (4.4) so that

$$b_n := \left(\frac{1}{P[|Z_1| > \cdot]} \right)^{\leftarrow} (n).$$

Then

$$\sum_{j=1}^n \varepsilon_{Z_j^{(m)}/b_n} \Rightarrow \sum_{l=0}^m \sum_k \varepsilon_{j_k \mathbf{e}_l} \tag{4.22}$$

where $\{j_k\}$ are the points of a Poisson process on $[-\infty, \infty] \setminus \{0\}$ with mean measure ν given in (4.6) and \mathbf{e}_l is a basis vector of length $m+1$ all of whose entries are zero except for a 1 in the l -th spot. (See Resnick [9], page 232 or Davis and Resnick [3]).

Step 2: Define

$$\boldsymbol{\psi}^{(m, j)} = (\psi_i^{(j)}, i \leq m)',$$

$j = 0, 1, \dots, \infty$. Apply Lemma 2 to the convergences (4.22) and (4.14) and we get

$$\sum_{j=1}^n \varepsilon_{(Z_j^{(m)}/b_n, \boldsymbol{\psi}^{(m, j)})} \Rightarrow \sum_{l=0}^m \sum_k \varepsilon_{(j_k \mathbf{e}_l, \boldsymbol{\psi}^{(m, \infty)})}. \tag{4.23}$$

Step 3: Define

$$X_j^{(m)} = \sum_{i=0}^m \psi_i^{(j)} Z_{j-i} = \langle Z_j^{(m)}, \boldsymbol{\psi}^{(m, j)} \rangle$$

so that $X_j^{(m)}$ is an inner product. Apply the inner product to the components of the points in (4.23) and by a continuous mapping argument we get

$$\sum_{j=1}^n \varepsilon_{X_j^{(m)}/b_n} \Rightarrow \sum_{l=0}^m \sum_k \varepsilon_{j_k \psi_l^{(m, \infty)}}$$

and after an argument which justifies replacing m by ∞ , we get

$$\sum_{j=1}^n \varepsilon_{X_j/b_n} \Rightarrow \sum_{l=0}^{\infty} \sum_k \varepsilon_{j_k \psi_l^{(\infty)}}. \tag{4.24}$$

Step 4: We now redo Step 2. We again apply Lemma 2 to the convergences (4.22) and (4.14) to get for any integer h

$$\sum_{j=1}^{n-h} \varepsilon_{(Z_{j+h}^{(m)}/b_n, \boldsymbol{\psi}^{(m, j+h)}, \boldsymbol{\psi}^{(m, j)})} \Rightarrow \sum_{l=0}^m \sum_k \varepsilon_{(j_k \mathbf{e}_l, \boldsymbol{\psi}^{(m, \infty)}, \boldsymbol{\psi}^{(m, \infty)})}. \tag{4.25}$$

Now define a map

$$[-\infty, \infty]^{m+1} \setminus \{\mathbf{0}\} \times \mathbb{R}^{m+1} \times \mathbb{R}^{m+1} \mapsto [-\infty, \infty]^2 \setminus \{\mathbf{0}\}$$

by

$$(\mathbf{z}, \boldsymbol{\psi}, \boldsymbol{\psi}') \mapsto (\langle \mathbf{z}, \boldsymbol{\psi} \rangle, z_0 \psi'_h + z_1 \psi'_{h+1} + \dots + z_{m-h} \psi'_m).$$

Note the second component on the right is obtained by ignoring the first h components of $\boldsymbol{\psi}'$ and then taking the inner product with the correspondingly truncated version of \mathbf{z} .

Apply this mapping to the components of (4.25). After a continuous mapping argument and a justification of the replacement of m by ∞ we get

$$\begin{aligned} & \left(\sum_{j=1}^n \varepsilon_{b_n^{-1}(X_{j+h}, X_j)}, 0 \leq h \leq k \right) \\ & \Rightarrow \left(\sum_{l=0}^{\infty} \sum_k \varepsilon_{(j_k \psi_l^{(\infty)}, j_k \psi_{l-h}^{(\infty)})}, 0 \leq h \leq k \right). \end{aligned} \quad (4.26)$$

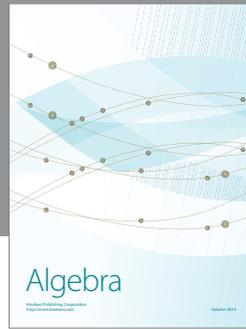
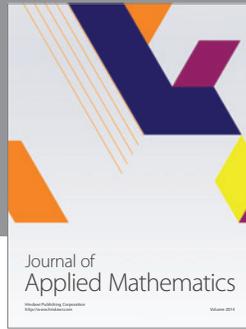
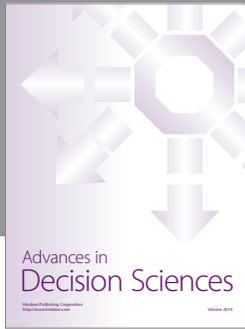
Note that any ψ with a negative subscript can be interpreted as 0.

Taking products of components and summing points and then taking ratios yield the result:

$$(\widehat{\rho}_H(h), 1 \leq h \leq k) \Rightarrow \left(\frac{\sum_{l=0}^{\infty} \psi_l^{(\infty)} \psi_{l+h}^{(\infty)}}{\sum_{l=0}^{\infty} (\psi_l^{(\infty)})^2}, 1 \leq h \leq k \right). \quad \square$$

References

- [1] Brockwell, P. and Davis, R., *Time Series: Theory and Methods*, (2nd Edition), Springer-Verlag, New York 1991.
- [2] Davis, R. and Mikosch, T., Sample correlations of ARCH processes, (1997), in preparation.
- [3] Davis, R. and Resnick, S., Limit theory for moving averages of random variables with regularly varying tail probabilities, *Ann. Probability* **13** (1985), 179-195.
- [4] Davis, R. and Resnick, S., Limit theory for the sample covariance and correlation functions of moving averages, *Annals of Statistics* **14** (1986), 533-558.
- [5] Davis, R. and Resnick, S., Limit theory for bilinear processes with heavy-tailed noise, *Annals of Applied Probability* **6** (1997), 1191-1210.
- [6] De Haan, L., Resnick, S., Rootzén, H. and de Vries, C., Extremal behavior of solutions to a stochastic difference equation with application to ARCH processes, *Stochastic Processes and Their Applications* **32** (1989), 213-224.
- [7] Feigin, P. and Resnick, S., Pitfalls of fitting autoregressive models for heavy-tailed time series, <http://www.orie.cornell.edu/trlist/trlist.html> at **TR1163.ps.Z** (1996).
- [8] Feller, W., *An Introduction to Probability Theory and its Applications*, Vol. 2, (2nd edition), Wiley, New York 1971.
- [9] Resnick, S., *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, New York 1987.
- [10] Resnick, S., *Adventures in Stochastic Processes*, Birkhauser, Boston 1992.
- [11] Rosiński, J., On the structure of stationary stable processes, *Ann. Probab.* **23** (1995), 1163-1187.
- [12] Rosiński, J. and Samorodnitsky, G., Classes of mixing stable processes, *Bernoulli* **2** (1996), 365-378.
- [13] Samorodnitsky, G. and Taqqu, M., *Stable Non-Gaussian Random Processes*, Chapman and Hall, New York 1994.
- [14] Surgailis, D., Rosiński, J., Mandrekar, V. and Cambanis, S., Stable mixed moving averages, *Probab. Theory Related Fields* **97** (1993), 543-558.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

