J. Appl. Prob. **41,** 639–653 (2004)
Printed in Israel

RECURRENCE PROPERTIES OF AUTOREGRESSIVE PROCESSES WITH SUPER-HEAVY-TAILED INNOVATIONS

ASSAF ZEEVI,* *Columbia University*
PETER W. GLYNN,** *Stanford University*

Abstract

This paper studies recurrence properties of autoregressive (AR) processes with 'super-heavy-tailed' innovations. Specifically, we study the case where the innovations are distributed, roughly speaking, as log-Pareto random variables (i.e. the tail decay is essentially a logarithm raised to some power). We show that these processes exhibit interesting and somewhat surprising behaviour. In particular, we show that AR(1) processes, with the usual root assumption that is necessary for stability, can exhibit null-recurrent as well as transient dynamics when the innovations follow a log-Cauchy-type distribution. In this regime, the recurrence classification of the process depends, somewhat surprisingly, on the value of the constant pre-multiplier of this distribution. More generally, for log-Pareto innovations, we provide a positive-recurrence/null-recurrence/transience classification of the corresponding AR processes.

Keywords: Autoregressive process; heavy tail; Markov process; recurrence; stochastic stability; transience

2000 Mathematics Subject Classification: Primary 68M20; 60J10; 60J15

1. Introduction

Autoregressive (AR) models are one of the most widely used stochastic models in existence, and play a central role in many areas of research. For example, in the realm of time-series modelling and analysis, the focus has been on linear models, namely autoregressive and moving average models (see e.g. [4]). In signal processing applications, AR models have been used extensively in filter design, modelling noise processes and quantization effects, and spectral analysis (see e.g. [10]). The simple structure of AR models has led to a good theoretical understanding of various properties that pertain to stability, estimation, and representation of these processes.

Despite the importance of this class of models, it turns out that the recurrence properties of AR processes have not yet been fully worked out. To be concrete, consider the scalar AR(1) process, $X = (X_n : n \geq 0)$, given by

$$X_{n+1} = \alpha X_n + Z_{n+1}, \tag{1}$$

where $(Z_n : n \geq 1)$ is a sequence of i.i.d. random variables independent of X_0, often referred to as the *innovation process*, representing additive noise. When $\alpha = 1$, the Markov chain X

Received 20 December 2002; revision received 16 December 2003.
* Postal address: Graduate School of Business, Columbia University, 3022 Broadway, New York, NY 10027-6902, USA. Email address: assaf@gsb.columbia.edu
** Postal address: Department of Management Science and Engineering, Stanford University, Stanford, CA 94305-4026, USA. Email address: glynn@stanford.edu

corresponds to a random walk, and the recurrence theory is well known. When $\alpha > 1$, X_n tends to grow geometrically with n, precluding recurrence (see Proposition 1 below and related discussion). On the other hand, when $|\alpha| < 1$, the chain tends to 'contract', and we expect it to be positive recurrent. The great majority of the literature on autoregressive processes therefore assumes that $|\alpha| < 1$, which we shall henceforth refer to as the 'usual stability condition'.

If, in addition to the usual stability condition, we assume that $\mathrm{E}\log(1 + |Z_1|) < \infty$, then X is a positive-recurrent Markov chain (see [1, Proposition 1]). The latter moment assumption will be referred to as the 'log-moment condition'. This raises an intriguing theoretical question:

> What are the recurrence properties of AR processes when we assume the usual stability condition (i.e. $|\alpha| < 1$) and $\mathrm{E}\log(1 + |Z_1|) = \infty$?

The above question essentially concerns the behaviour of AR processes that are subject to innovations that assume 'very large' values. If Z satisfies the condition $\mathrm{E}\log(1 + |Z_1|) = \infty$, we say that its distribution has *super-heavy tails*. In this regime, the process has dynamics that are contractive due to the magnitude of α; however, this is potentially off-set by the extremely large values introduced by the innovations. To illustrate the main point, observe that X_n can be written as

$$X_n = \alpha^n X_0 + \sum_{i=0}^{n-1} \alpha^i Z_{n-i}. \tag{2}$$

Note that in the super-heavy-tailed regime, $\max\{Z_k : 1 \le k \le n\}$ can grow roughly like $\exp(cn)$ for some $c > 0$. We therefore anticipate that this growth can potentially cause the process to be transient, or null recurrent, in spite of the contraction caused by exponentially decaying weights.

The theory that we develop will be largely focused on the following subclass of super-heavy-tailed distributions. We say that a nonnegative random variable Z has *log-Pareto tails* if $\log(1 + Z)$ has a Pareto distribution with parameters (p, β), i.e. for some $p > 0$ and $\beta > 0$,

$$\mathrm{P}(\log(1 + Z) > z) = \frac{1}{(1 + \beta z)^p} \tag{3}$$

when $z \ge 0$. In the case $p = 1$, we shall refer to Z as having 'log-Cauchy tails' with scale parameter $\beta > 0$. By imposing a similar restriction on the left tail we can extend this to the case of random variables taking on positive and negative values.

The main contributions of this paper are the following:

(i) We show that the known sufficient condition for positive recurrence, $\mathrm{E}\log(1+|Z_1|) < \infty$, is also necessary (see Proposition 1). The latter generalizes in a straightforward manner to vector AR processes (see Proposition 2).

(ii) We show that recurrence classification using return times to compact intervals is equivalent to using the potential function (see Proposition 3). Necessary and sufficient conditions for classifying recurrent/transient behaviour are established using properties of the first exit time of the 'backward iterated process' from an interval (see Proposition 4). We also provide a sufficient condition for recurrence that is more easy to verify in practice (see Lemma 1).

(iii) We obtain a complete characterization of transience/null-recurrence/positive-recurrence for AR processes with log-Pareto innovations (see Theorem 1). We also derive a more general sufficient condition for transience when the innovations are super-heavy tailed but are not distributed as log-Pareto (see Proposition 5). Note that positive recurrence, transience, and even null recurrence can all be exhibited by AR processes with $|\alpha| < 1$.

(iv) We show that when positive recurrence fails and the innovations are not log-Pareto, no simple moment condition appears to exist for differentiating between null recurrence and transience. In particular, when $\mathrm{E}(\log(1 + |Z|))^p = \infty$ for some $p < 1$, we can have either transience or null recurrence, and therefore cannot differentiate the dynamics based on log-moments (see Example 1).

When $p = 1$ in (3) above, the distinction between null recurrence and transience relies on the pre-multiplier, β, in the corresponding log-Cauchy distribution of the innovations. This is another indication that moment conditions are not sufficient for this type of classification.

Thus, the conclusions of this study are that log-Pareto tails can give rise to a surprisingly broad range of possible dynamics, even in the presence of the usual stability condition $|\alpha| < 1$. Moreover, the moment-based classification that holds for log-Pareto innovations does not hold in general. Thus, it seems that fine-grain properties of the tail are crucial in determining recurrence properties in the super-heavy-tailed regime.

As mentioned previously, sufficiency of the log-moment condition for positive recurrence was proved explicitly by Athreya and Pantula [1]. Vervaat [12] investigated stability of linear stochastic recursions under more general dependence conditions (see also [3]). Recurrence classification for a class of nonlinear generalizations of AR processes was investigated in continuous time by Brockwell *et al.* [5] and in discrete time by Rai *et al.* [11]. The latter paper studies stochastic recursions of the form $X_{n+1} = X_n - aX_{n+1}^b + Z_{n+1}$, where Z_{n+1} represents the inflow and aX_{n+1}^b the outflow over the interval of time $[n, n + 1)$. Note that when $b = 1$ we recover an autoregressive sequence. Recurrence properties of Markov chains were discussed more generally by Borovkov [3] and Meyn and Tweedie [9].

2. Main results

Let us first define more rigorously what we mean by recurrence properties. Consider a Markov chain $X = (X_n : n \geq 0)$ on a state space $E \subseteq \mathbb{R}^d$. Let $\mathcal{E} = \mathcal{B}(E)$ denote the associated Borel sigma-field over E. Let P_x denote the underlying probability measure conditional on $X_0 = x$, and let $\mathrm{E}_x[\,\cdot\,] := \mathrm{E}[\,\cdot \mid X_0 = x]$. For any $A \in \mathcal{E}$, let $T_A := \inf\{n \geq 1 : X_n \in A\}$ denote the hitting time of A. Let $\|\cdot\|$ denote the usual Euclidean norm and $|\cdot|$ denote absolute value. In what follows, '$\Rightarrow$' is used to denote weak convergence (i.e. convergence in distribution).

We will use the following notions of recurrence for an E-valued Markov chain X.

Definition 1. (*Recurrence.*) (i) The Markov chain X is said to be *recurrent* if there exists a compact set A such that

$$\mathrm{P}_x(T_A < \infty) = 1 \quad \text{for all } x \in E$$

and *transient* otherwise.

(ii) The Markov chain X is said to be *positive recurrent* if there exists a compact set A such that

$$\mathrm{E}_x[T_A] < \infty \quad \text{for all } x \in E.$$

(iii) The Markov chain X is said to be *null recurrent* if it recurrent but not positive recurrent.

Null recurrence, positive recurrence, and transience as defined above are not generally a mutually exclusive partition of possibilities; however, this is true in the context of AR processes (for a more general treatment the reader is referred to [9]). An alternative definition of recurrence uses the notion of *potential*. In particular, let

$$\eta_A = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_n \in A\}}$$

denote the *occupation measure* of the set A, where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function, and let

$$U(x, A) := \mathrm{E}_x[\eta_A] = \sum_{n=1}^{\infty} \mathrm{P}_x(X_n \in A)$$

denote the *potential* or *mean occupation measure* of the set A. We say that the chain X is *recurrent* if $U(x, A) = \infty$ for every initial state x and compact set A. In what follows we establish an equivalence between this characterization and the one given in Definition 1 (see Proposition 3).

2.1. Positive recurrence under log-moment conditions

Our first result concerns the necessity and sufficiency of $\mathrm{E} \log(1 + |Z_1|) < \infty$, the log-moment condition, for positive recurrence of the chain X.

Proposition 1. *The process X is positive recurrent if and only if $|\alpha| < 1$ and $\mathrm{E} \log(1+|Z_1|) < \infty$, in which case X admits a unique stationary distribution, π, and $\mathrm{P}_x(X_n \in \cdot) \Rightarrow \pi(\cdot)$ as $n \to \infty$.*

Proposition 1 indicates that the stability condition $|\alpha| < 1$ and the finiteness of the log-moment of the innovation process provide 'sharp' conditions for positive recurrence. The sufficiency of the log-moment condition appears in [1] and necessity for the case where Z is nonnegative is given in [5] and [11].

We now briefly illustrate how the necessary and sufficient conditions for positive recurrence given in Proposition 1 can be extended to the vector case. Consider the recursion

$$X_{n+1} = AX_n + Z_{n+1},$$

where $A \in \mathbb{R}^{d \times d}$ is a nonsingular matrix, and X_n and Z_n take values in $\mathbb{R}^d$. Assume that the innovation process $Z = (Z_n : n \geq 1)$ consists of i.i.d. random vectors and the components of these vectors are independent of each other. We also assume that Z is independent of X_0. Let the *spectral radius* of A be defined as

$$\rho(A) = \max\{|\lambda| : Av = \lambda v, v \neq 0\}.$$

The following is an analogue of Proposition 1.

Proposition 2. *If A is nonsingular, then X is positive recurrent if and only if $\rho(A) < 1$ and $\mathrm{E} \log(1 + \|Z_1\|) < \infty$, in which case X admits a unique stationary distribution, π, and $\mathrm{P}_x(X_n \in \cdot) \Rightarrow \pi(\cdot)$ as $n \to \infty$.*

The above result is clearly applicable to AR(d) processes. Specifically, consider the process

$$X_{n+1} = \alpha_1 X_n + \alpha_2 X_{n-1} + \cdots + \alpha_d X_{n-d+1} + Z_{n+1},$$

where $Z = (Z_n : n \geq 1)$ is a process of i.i.d. innovations independent of $\{X_0, \dots, X_{d-1}\}$. (Implicit here is the assumption that $\alpha_d \neq 0$.) Recasting this as a vector-valued Markov chain, let $Y_n = (X_n, \dots, X_{n-d+1})^\top$; then $Y_{n+1} = AY_n + W_{n+1}$, where $W_n = (Z_n, \dots, Z_{n-d+1})^\top$ and

$$A = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{d-1} & \alpha_d \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}_{d\times d}.$$

We also set $Y_0 = (X_{d-1}, \dots, X_0)^\top$. Applying Proposition 2 we have that the process $Y = (Y_n : n \geq 0)$ admits a stationary distribution if and only if $\mathrm{E}\log(1 + |Z_1|) < \infty$ and $\mathcal{P}(z) = \det(A - zI)$ has all its roots strictly inside the unit disc in the complex plane. The latter condition can be simplified by a more explicit description of $\mathcal{P}(z)$, namely

$$\mathcal{P}(z) = z^d - \alpha_1 Z^{d-1} - \cdots - \alpha_d.$$

The sufficiency of these conditions can be found, for example, in [1, p. 891].

2.2. Recurrence classification with super-heavy-tailed innovations

As mentioned previously, our goal here is to investigate what happens when the log-moment condition is violated; in particular, our main objective is to illustrate the range of behaviour that AR processes can exhibit in this super-heavy-tailed setting. To this end, we will restrict attention to nonnegative innovation processes and suppose that $0 < \alpha < 1$ in order to reduce technical complications. (Clearly if $\alpha = 0$ the classification is vacuous.)

The next result establishes a correspondence between the properties of the first hitting time of an interval $[0, c]$, $T_{[0,c]}$, and the potential function, $U(x, [0, c])$, for the process started at $X_0 = x$. Proposition 3 asserts that the notion of recurrence articulated in Definition 1 is equivalent to nonfiniteness of the potential function.

Proposition 3. *Let X be an AR(1) process with nonnegative innovations and $\alpha \in (0, 1)$. Then, for all $x \geq 0$,*

$$U(x, [0, c]) = \infty \quad \textit{if and only if} \quad \mathrm{P}_x(T_{[0,c]} < \infty) = 1.$$

For the class of models we are concerned with in this paper, it is quite natural to focus on recurrence conditions that are driven by the study of dynamical systems, viewing Markov chains as iterated random maps (cf. [8] and the recent survey by Diaconis and Freedman [6]). In particular, for the scalar AR process (1) we define the *backward iterated process* $S = (S_n : n \geq 1)$ associated with X to be

$$S_n = \alpha^n X_0 + \sum_{i=1}^{n} \alpha^{i-1} Z_i. \tag{4}$$

To better understand the terminology 'backward iterated', contrast this with (2) and observe that S_n is derived from X_n by essentially substituting Z_i for Z_{n-i} in the latter. While S_n has the same distribution as X_n for all n, it is also clear that the backward process S has very different behaviour from the forward process X. In particular, under regularity conditions that ensure 'stability', the backward process converges *almost surely* (as n grows to infinity) to a limiting random variable whose distribution is the unique stationary distribution of the chain X, while the forward process converges *in distribution* to the same limiting random variable. (For further details on the requisite conditions see Theorem 1 of [6].)

For the purpose of recurrence classification of AR processes, we can use the *first exit time* of the backward process S from an interval $[0, c]$ defined as follows:

$$\tau(c) = \inf\{n \geq 1 : S_n > c\}. \tag{5}$$

The key observation, stated informally, is that a ‘finite’ exit time rules out positive recurrence, and the precise definition of ‘finite’ stands in one-to-one correspondence with a null-recurrent/transient classification.

Proposition 4. *Let X be an $AR(1)$ process with nonnegative innovations and $\alpha \in (0, 1)$. Then*

(i) *the chain is positive recurrent if and only if, for some $c > 0$, $\mathrm{P}_0(\tau(c) = \infty) > 0$;*

(ii) *the chain is null recurrent if and only if, for some $c > 0$, $\mathrm{E}_0[\tau(c)] = \infty$ and $\mathrm{P}_0(\tau(c) < \infty) = 1$;*

(iii) *the chain is transient if and only if, for some $c > 0$, $\mathrm{E}_0[\tau(c)] < \infty$.*

For the next result we impose the additional assumption that the innovations are distributed log-Pareto in the sense of (3). While this may seem quite restrictive at first glance, Example 1 clearly indicates that in the super-heavy-tailed regime it is necessary, in some sense, to have more control over the precise characteristics of the tail of the innovations distribution.

Theorem 1. *If $\alpha \in (0, 1)$ and the innovations have log-Pareto distribution as in (3) with parameters (p, β), then*

(i) *if $p > 1$, the chain is positive recurrent;*

(ii) *if $p < 1$, the chain is transient;*

(iii) *if $p = 1$ and $\beta \log(1/\alpha) < 1$, the chain is transient and, if $p = 1$ and $\beta \log(1/\alpha) \geq 1$, the chain is null recurrent.*

The results described in Proposition 1 and Theorem 1 indicate that a ‘general’ recurrence-classification theory under the usual stability condition might look roughly as follows: (i) if $\mathrm{E}\log(1 + |Z_1|) < \infty$, then the chain is positive recurrent (see Proposition 1); (ii) if Z_1 has log-Cauchy tails, both transient and null-recurrent dynamics are possible according to the classification result in Theorem 1; and (iii) if $\sup\{p : \mathrm{E}(\log(1 + |Z_1|))^p < \infty\} < 1$, then the chain is transient. Unfortunately, it turns out that only parts (i) and (ii) of the ‘general theory’ hold, and the transience implication in (iii) is generally false. In fact, if $\mathrm{E}(\log(1+|Z_1|))^p = \infty$ for some $p < 1$, the chain could still be null recurrent as the following example illustrates.

Example 1. (*Null recurrence and insufficiency of log-moment conditions.*) Our construction uses the following auxiliary lemma that provides a relatively simple way to establish recurrence.

Lemma 1. *Let X be an $AR(1)$ process with nonnegative innovations and $\alpha \in (0, 1)$. If, for some $\delta > 0$,*

$$\int_0^\infty \exp\biggl\{-(1+\delta)\int_0^t \mathrm{P}(\log(1 + Z_1) > y)\biggr\}\,\mathrm{d}t = \infty, \tag{6}$$

then X is recurrent.

Let Z be such that $\mathrm{E}(\log(1+Z_1))^p = \infty$ for some $p < 1$. Put $W := \lfloor \log(1+Z_1) \rfloor$, where $\lfloor \cdot \rfloor$ is the integer-part function, and note that $\mathrm{E}[W] = \infty$. We now construct a distribution for W such that its mean is infinite, yet the integral in (6) diverges when $\delta = 1$. This will imply that the AR chain is null recurrent by Lemma 1 and Proposition 1.

We will specify two nonnegative sequences, t_n which increases to ∞ as $n \to \infty$, and p_n which decreases to zero as $n \to \infty$. We then specify the distribution of W through its tail values as follows:

$$\mathrm{P}(W > y) = p_{n+1} \quad \text{for } y \in [t_n, t_{n+1}).$$

Note that in order to guarantee the divergence of $\mathrm{E}[W]$, we simply impose the constraint that $p_{n+1}(t_{n+1} - t_n) = 1/(n+1)$ for each $n \geq 1$. The construction now proceeds recursively. For brevity, set

$$\psi(t) := \exp\biggl\{-2\int_0^t \mathrm{P}(W > y)\,\mathrm{d}y\biggr\}.$$

Suppose that, when $t = t_n$, we have $\psi(t_n) = a_n$. Then, for $t > t_n$,

$$\psi(t) = a_n \exp\biggl\{-2\int_{t_n}^t \mathrm{P}(W > y)\,\mathrm{d}y\biggr\}.$$

But now use the fact that $\mathrm{P}(W > y) = p_{n+1}$ for $y \in [t_n, t_{n+1})$ to get that

$$\psi(t) = a_n \exp\{-2p_{n+1}(t - t_n)\}.$$

Thus, we can integrate $\psi(\cdot)$ over $[t_n, t_{n+1})$ to give

$$\int_{t_n}^{t_{n+1}} \psi(t)\,\mathrm{d}t = \frac{a_n}{2p_{n+1}}(1 - \exp\{-2p_{n+1}(t_{n+1} - t_n)\}).$$

First, note that by construction $p_{n+1}(t_{n+1} - t_n) = 1/(n+1)$; thus, as $n \to \infty$,

$$1 - \exp\{-2p_{n+1}(t_{n+1} - t_n)\} \sim \frac{2}{n},$$

where we say that $x_n \sim y_n$ if $x_n/y_n \to 1$ as $n \to \infty$. Thus,

$$\int_{t_n}^{t_{n+1}} \psi(t)\,\mathrm{d}t \sim \frac{a_n}{np_{n+1}}.$$

To make the whole integral diverge we set $p_{n+1} = \min\{a_n, p_n, 1/n\}$, so that $a_n/p_{n+1} \geq 1$. (We make this choice of p_{n+1} to comply with the constraint $p_n \searrow 0$ as $n \to \infty$.) Consequently, $\int_0^\infty \psi(t)\,\mathrm{d}t$ diverges and the chain must therefore be null recurrent in spite of the fact that $\mathrm{E}(\log(1+Z_1))^{1/2} = \infty$.

The above example demonstrates that we cannot expect in general to distinguish between null-recurrence and transient behaviour of AR processes based on simple moment conditions. This observation also emphasizes the surprising 'sharpness' of Proposition 1, namely that the log-moment condition is necessary and sufficient to determine positive recurrence. The reader may wonder what the intuition behind the counterexample described above is. The construction gives rise to a distribution with support on a sequence t_n that is increasing very rapidly with probabilities that are decaying very rapidly. In particular, the behaviour of the maximum of the Z_i will not exhibit 'smooth' monotonic increasing behaviour; rather, it will tend to be constant on long stretches of time, and then jump. This is due to the recursive construction of the distribution above, namely we must go farther and farther 'out into the tail' to encounter the next mass point. Since it is extremely unlikely that up to time n the Z_i take on values that are in the far tail (the quantile being much greater than $1 - 1/n$), the maximum over time blocks of length n will tend to be almost constant, and finally will jump only when the time that elapsed is long enough so that the next support point in the tail is more likely to give rise to a value of Z. Thus, the Z_i are effectively bounded over long time intervals, and therefore the contraction due to $\alpha < 1$ will imply that the chain returns to a compact set before the next jump in the value of the Z is encountered. This, in turn, suggests that the chain ought to exhibit recurrent rather than transient behaviour.

While it turns out that $\mathrm{E}(\log(1 + Z_1))^p = \infty$ for some $p < 1$ does not imply transience, we can guarantee the latter if we strengthen the moment condition as follows.

Proposition 5. *If* $\alpha \in (0, 1)$ *and the innovations are such that*

$$\liminf_{z\to\infty} (\log z)^p \mathrm{P}(Z_1 > z) \geq c$$

for $p < 1$ *and some finite positive constant c, then the chain associated with the* $AR(1)$ *process is transient.*

3. Proofs

3.1. Proof of Proposition 1

Sufficiency of these conditions for positive recurrence and the existence and uniqueness of a stationary distribution, as well as the weak convergence asserted in the proposition, were proved by Athreya and Pantula [1, Proposition 1]. To prove necessity, we use convergence properties of random series. Theorem 2.7(ii) of Doob [7, p. 115] asserts that if W_i is a sequence of independent random variables, then $S_n = \sum_{i=1}^n W_i \Rightarrow S$ if and only if $S_n \to S$ almost surely. Observe that

$$X_n = \alpha^n X_0 + \sum_{i=0}^{n-1} \alpha^i Z_{n-i};$$

thus, X_n is equal in distribution to the 'backward process' $S_n = \alpha^n X_0 + \sum_{i=1}^n \alpha^{i-1} Z_i$. Consider first the necessity that $|\alpha| < 1$; in particular, suppose that $|\alpha| \geq 1$. Then, for any $\varepsilon > 0$, $\mathrm{P}(|\alpha^i Z_i| > \varepsilon) \geq \mathrm{P}(|Z_1| \geq \varepsilon)$ for all i. Thus, $\sum_i \mathrm{P}(|\alpha^i Z_i| \geq \varepsilon) < \infty$ for all $\varepsilon > 0$ if and only if $Z_i = 0$ for all i. By Kolmogorov's three-series theorem (see e.g. [2, Theorem 22.8, p. 290]) this is necessary for the almost-sure convergence of S_n and thus (by [7, Theorem 2.7, p. 115]) for the weak convergence of S_n. Consequently it is also necessary for the weak convergence of X_n. The only matter left is the necessity of the log-moment assumption. To this end, note that $\mathrm{E}\log(1 + |Z_1|) = \infty$ implies that

$$\sum_{n=1}^{\infty} \mathrm{P}(|Z_n| \geq \alpha^{-n}) = \infty,$$

and thus by the second Borel–Cantelli lemma we have that $Z_n \geq \alpha^{-n}$ for infinitely many n, almost surely. Consequently, S_n is almost surely not convergent and, therefore, by the aforementioned result on random series, X_n cannot admit a stationary distribution; in particular, it is not positive recurrent. This concludes the proof.

3.2. Proof of Proposition 2

The necessity of the stability condition $\rho(A) < 1$ follows as in the proof of Proposition 1. The remainder of the proof is split into two parts.

3.2.1. *Sufficiency.* To argue sufficiency note that

$$X_n = A^n X_0 + \sum_{i=0}^{n-1} A^i Z_{n-i} \tag{7}$$

and the first term on the right-hand side converges to zero almost surely as $n \to \infty$, since $\rho(A) < 1$ implies that $\|A^m\| < 1$ for some $m < \infty$ (indeed $\|A^n\|^{1/n} \to \rho(A)$ as $n \to \infty$). Here $\|A\| := \sup\{\|Ay\| : \|y\| = 1\}$. Thus, it suffices to consider the case where $X_0 = 0$. Then the second term on the right-hand side of (7) equals in distribution $S_n := \sum_{i=1}^{n} A^{i-1} Z_i$ (the backward iterated chain). Now, the log-moment condition ensures, by the Borel–Cantelli lemma, that $\|Z_n\| \le \exp(cn)$ for all $c > 0$ and sufficiently large n, almost surely. Thus, since $\rho(A) < 1$,

$$\sum_{i=1}^{\infty} \|A^{i-1} Z_i\| < \infty$$

almost surely, and therefore S_n converges almost surely to $S_\infty = \sum_{i=1}^{\infty} A^{i-1} Z_i$. The latter defines the (unique) stationary distribution for X, and also establishes that $X_n \Rightarrow X_\infty$, where X_∞ has the same distribution as S_∞.

3.2.2. *Necessity.* To argue necessity note that we may apply Theorem 2.7(ii) of Doob [7, p. 115] component by component since each coordinate of S_n is expressed as the sum of i.i.d. random variables. Therefore, the ith component converges weakly if and only if it converges almost surely. Now, observe that

$$\|Z_n\| = \|(A^n)^{-1} A^n Z_n\| \le \|(A^n)^{-1}\| \|A^n Z_n\|,$$

and consequently $\|A^n Z_n\| \ge \|Z_n\| \|(A^n)^{-1}\|^{-1} \ge \|Z_n\| \|A^n\|$. Now, if $\mathrm{E} \log(1 + \|Z_1\|) = \infty$, then $\|Z_n\| \ge \exp(c_1 n)$ for all $c_1 > 0$ infinitely often, almost surely, by the second Borel–Cantelli lemma. Thus, since $\|A^n\| \ge c_2 \rho^n$ for some $c_2 > 0$ and $\rho \in (0, 1)$, it follows that $\limsup_{n\to\infty} \|A^n Z_n\| = \infty$ almost surely. There will therefore be at least one component of the vector $A^n Z_n$ (since there are only finitely many components) on which the above lim sup is infinite. Since the aforementioned component of S_n does not converge almost surely, it follows (by [7, Theorem 2.7, p. 115]) that S_n does not converge in distribution. Thus, it must be that X_n does not converge in distribution and consequently X does not admit a stationary distribution. Therefore, the chain cannot be positive recurrent, and the proof is complete.

3.3. Proof of Proposition 3

We need to prove the equivalence of the behaviour of the potential and the probability of the return time to compact intervals $T_{[0,c]} = \inf\{n \ge 1 : X_n \in [0, c]\}$. Suppose first that $\mathrm{P}_x(T_{[0,c]} < \infty) = 1$. Consider the case where Z_1 has bounded support, where we may have $\mathrm{P}_x(T_{[0,c]} = 1) = 1$ for the particular choice of x. If this holds for all $x \in [0, c]$, then clearly the chain never leaves $[0, c]$ and the proof follows, otherwise there exists an $x' > x$ in $[0, c]$ such that $\mathrm{P}_{x'}(T_{[0,c]} = 1) < 1$. Thus, assume without loss of generality that $\mathrm{P}_x(T_{[0,c]} = 1) < 1$. For any set A, let $P(x, A) := \mathrm{P}_x(X_1 \in A)$. Conditioning on the chain at time 1 gives the expression

$$\mathrm{P}_x(T_{[0,c]} < \infty) = \int_{y \in [0,c]} P(x, \mathrm{d}y) + \int_{y \in [0,c]^{\mathrm{c}}} \mathrm{P}_y(T_{[0,c]} < \infty) P(x, \mathrm{d}y).$$

Thus, by the assumption that $\mathrm{P}_x(T_{[0,c]} < \infty) = 1$, there must exist values of y in $[0, c]^{\mathrm{c}}$ such that the probability of hitting $[0, c]$ in finite time is 1. But, for each such y, stochastic monotonicity guarantees that $\mathrm{P}_z(T_{[0,c]} < \infty) = 1$ for all $z \le y$. Thus, it follows that the return probability to $[0, c]$ must be 1 for all initial states $x \in [0, y]$, and repeating the argument above it follows that the hitting time of $[0, c]$ is finite, almost surely, for all initial states. Hence, the chain returns to $[0, c]$ infinitely often and, therefore, $U(x, [0, c]) = \infty$ for all x. Conversely, suppose that there exists an $x \in [0, c]$ for which $\mathrm{P}_x(T_{[0,c]} < \infty) < 1$. Then, from the above argument, it follows that

$$p := \mathrm{P}_0(T_{[0,c]} < \infty) < 1.$$

Hence, a 'geometric trials' argument establishes that

$$\mathrm{P}_0(\eta_{[0,c]} = n) \le p^n,$$

and consequently $U(0, [0, c]) = \sum_{n=1}^{\infty} \mathrm{P}_0(\eta_{[0,c]} = n) < \infty$. By the stochastic monotonicity of X, $U(x, [0, c]) \le U(0, [0, c])$ for all $x \ge 0$ and, therefore, the expected occupation measure is sigma-finite. This concludes the proof.

3.4. Proof of Proposition 4

Note that because $X_n = \alpha^n X_0 + \sum_{i=0}^{n-1} \alpha^i Z_{n-i}$ and $\alpha \in (0, 1)$, it follows that the recurrence properties of X do not depend on its initial value $X_0 = x$. It suffices therefore to consider the case $X_0 = 0$. Fix $c > 0$. Now, using the definition of the backward process given in (4) and the first exit time of that process from $[0, c]$ given in (5), it follows that

$$\mathrm{E}_0[\eta_{[0,c]}] = \mathrm{E}_0[\tau(c)] - 1.$$

That is, the expected value of the occupation measure of a set $[0, c]$ is equal to the expected value of the first exit time of that set by the backward iterated process, minus 1. It therefore follows that $\mathrm{E}_0[\tau(c)] = U(0, [0, c]) + 1$. Case (iii) in the proposition now follows because $U(0, [0, c]) < \infty$ defines transience and, conversely, $U(0, [0, c]) = \infty$ defines recurrence by Proposition 3. Now, the key to determining null recurrence and distinguishing this behaviour from positive recurrence hinges on $\mathrm{P}_0(\tau(c) < \infty)$. To this end, let $E_n = \{S_n \in [0, c]\}$, where $S_n = \sum_{i=1}^{n} \alpha^{i-1} Z_i$ is the backward process started at zero. Since S_n is P_0-almost surely nondecreasing, we have that $\{\bigcap_n E_n\} = \{\tau(c) = \infty\}$. Thus,

$$\begin{aligned}
\mathrm{P}_0(\tau(c) = \infty) &= \mathrm{P}_0\Big(\bigcap_n E_n\Big) \\
&= \lim_{n \to \infty} \mathrm{P}_0(S_n \in [0, c]) \\
&= \lim_{n \to \infty} \mathrm{P}_0(X_n \in [0, c]),
\end{aligned}$$

since S_n is equal in distribution to X_n for all n. Now, if $\mathrm{P}_0(\tau(c) = \infty) > 0$, it follows that $\lim_{n\to\infty} \mathrm{P}_0(X_n \in [0, c]) > 0$, which implies that X cannot be null recurrent since in that case the above limit must be zero for any interval $[0, c]$. Conversely, if X is positive recurrent, then, by Proposition 1, $\mathrm{P}_0(X_n \in [0, c]) \to \mathrm{P}_\pi(X_1 \in [0, c])$ as $n \to \infty$, and the latter must be positive for some $c > 0$. This concludes the proof.

3.5. Proof of Theorem 1

Our proof uses the necessary and sufficient conditions given in Propositions 1 and 3.

(i) $p > 1$. This follows since the log-moment condition of Proposition 1 is satisfied.

(ii) $p < 1$. Fix $c > 0$ and $x \in [0, c]$. Since the potential is given in this context by

$$U(x, [0, c]) = \sum_{n=1}^{\infty} \mathrm{P}_x(X_n \le c),$$

the key is to bound the rate at which the probabilities on the right-hand side decay to zero. To this end, note that $U(x, [0, c]) \le U(0, [0, c])$ for all $x > 0$. Thus, we may take $x = 0$. Observe that we have the following inclusion of events in the underlying sigma-field:

$$\{X_n \le c\} \subseteq \{Z_{n-i} \le c\alpha^{-i},\ i = 0, \dots, n-1\} \tag{8}$$

for all $c \in \mathbb{R}_+$. To see why this holds, note that, if $Z_{n-i} \ge c\alpha^{-i}$ for some i, then $X_n = \sum_{i=1}^{n} \alpha^i Z_{n-i} \ge c$. Thus,

$$\begin{aligned} \mathrm{P}_x(X_n \le c) &\le \mathrm{P}(Z_i \le c\alpha^{-i},\ i = 1, \dots, n) \\ &= \prod_{i=1}^{n} [1 - \mathrm{P}(Z_1 > c\alpha^{-i})] \\ &= \exp\left\{\sum_{i=1}^{n} \log[1 - \mathrm{P}(Z_1 > c\alpha^{-i})]\right\} \\ &\le \exp\left\{-\sum_{i=1}^{n} \mathrm{P}(Z_1 > c\alpha^{-i})\right\}, \end{aligned} \tag{9}$$

where the first step follows from the set inclusion (8), the second step uses independence of the innovations, and the last step uses the inequality $\log(1 - y) \le -y$. Now, using the assumption on the distribution of Z_1, we have

$$\mathrm{P}(Z_1 > c\alpha^{-i}) = (1 + \beta \log(1 + c\alpha^{-i}))^{-p} \ge C_1 \frac{1}{i^p}$$

for some finite positive constant C_1 and for sufficiently large i. Thus,

$$\sum_{i=1}^{n} \mathrm{P}(Z_1 > c\alpha^{-i}) \ge C_2(1 + n^{1-p})$$

for sufficiently large n, and therefore

$$U(x,[0,c]) = \sum_{n=1}^{\infty} \mathrm{P}_x(X_n \le c) < \infty.$$

Since c was arbitrary, the chain is transient.

(iii) $p = 1$. Fix $c > 0$ and $x \in [0, c]$. There are two subcases to deal with.

Case 1. $\beta \log(1/\alpha) \ge 1$. We use the following set inclusion:

$$\{X_n \le c\} \supseteq \left\{ Z_{n-i} \le \frac{c}{2}\frac{b}{(i+1)^2}\alpha^{-i},\ i = 0, \dots, n-1 \right\},$$

which holds for all $n > \max\{\lfloor \log(2x/c)/|\log\alpha| \rfloor, 1\}$. Here $b := (\sum_{i=1}^{\infty} i^{-2})^{-1} = 6/\pi^2$. Note that the lower bound on n ensures that $x\alpha^n \le c/2$. Observe that, if $Z_{n-i} \le (c/2)(b/(i+1)^2)\alpha^{-i}$ for $i = 0, \dots, n-1$, then

$$X_n \le \frac{c}{2} + \sum_{i=0}^{n-1} \alpha^i Z_{n-i} \le \frac{c}{2}\left(1 + b\sum_{i=1}^{n}\frac{1}{i^2}\right) \le c.$$

Similar to the derivation in (ii), we now have that

$$\begin{aligned}
\mathrm{P}_x(X_n \le c) &\ge \mathrm{P}\left(Z_i \le \frac{c}{2}\frac{b}{i^2}\alpha^{-i},\ i = 1, \dots, n \right) \\
&= \exp\left\{ \sum_{i=1}^{n} \log\left[1 - \mathrm{P}\left(Z_1 > \frac{c}{2}\frac{b}{i^2}\alpha^{-i} \right) \right] \right\} \\
&\ge C_1 \exp\left\{ -\sum_{i=1}^{n} \mathrm{P}\left(Z_1 > \frac{c}{2}\frac{b}{i^2}\alpha^{-i} \right) - \sum_{i=1}^{n} \left(\mathrm{P}(Z_1 > \frac{c}{2}\frac{b}{i^2}\alpha^{-i}) \right)^2 \right\} \\
&=: C_1 \exp\{-R_n - Q_n\}, \qquad (10)
\end{aligned}$$

where we have used the inequality $\log(1-y) \ge -y - y^2$, which holds whenever $0 \le y \le \frac{1}{2}$, and where C_1 is a finite positive constant. Using the assumption on the distribution of Z_1, we have that

$$\begin{aligned}
\mathrm{P}\left(Z_1 > \frac{c}{2}\frac{b}{i^2}\alpha^{-i} \right) &\le \mathrm{P}\left(\log(1+Z_1) > \log\left(\frac{cb}{2}\right) - 2\log i + i\log\left(\frac{1}{\alpha}\right) \right) \\
&\le \frac{1}{i\beta\log(1/\alpha) - 2\beta\log i},
\end{aligned}$$

assuming without loss of generality that c is chosen such that $c > 2/b$. This implies that

$$Q_n < \sum_{i=1}^{\infty} \left(\mathrm{P}\left(Z_1 > \frac{c}{2}\frac{b}{i^2}\alpha^{-i} \right) \right)^2 < \infty.$$

To evaluate the magnitude of R_n we use the following relation, which is readily verified:

$$\int_a^n \frac{1}{x - k\log x}\,\mathrm{d}x = \log(x - k\log x)|_a^n + \int_a^n \frac{k/x}{x - k\log x}\,\mathrm{d}x. \tag{11}$$

(The authors are grateful to the referee for suggesting this approach.) Here $a > 1$ is such that $x - k\log x > 0$ for all $x \geq a$. Note that

$$\int_a^\infty \frac{k/x}{x - k\log x}\,\mathrm{d}x < \infty.$$

Setting $k := 2\beta(\beta\log(1/\alpha))^{-1}$ we can use the integral relation (11) to give

$$\begin{aligned} R_n &\leq C_2 + \frac{1}{\beta\log(1/\alpha)}\int_a^n \frac{1}{x - k\log x}\mathrm{d}x \\ &\leq C_3 + \frac{1}{\beta\log(1/\alpha)}\log n, \end{aligned}$$

which holds for all n sufficiently large and some finite positive constants C_2, C_3. Combining the above with (10) we conclude that $\sum_n \mathrm{P}_x(X_n \leq c)$ diverges when $\beta\log(1/\alpha) \geq 1$.

Case 2. $\beta\log(1/\alpha) < 1$. In this case we have that

$$\begin{aligned} \mathrm{P}(Z_1 > c\alpha^{-i}) &= \frac{1}{1 + \beta\log(1 + c\alpha^{-i})} \\ &\geq \frac{1}{C_4 + i\beta\log(1/\alpha)} \end{aligned}$$

for some $C_4 > 0$ and for i sufficiently large. It then follows that

$$\sum_{i=1}^n \mathrm{P}(Z_1 > c\alpha^{-i}) \geq \left(\beta\log\left(\frac{1}{\alpha}\right) + \varepsilon\right)^{-1}\log n$$

for n sufficiently large and $\varepsilon \in (0, 1 - \beta\log(1/\alpha))$. Thus,

$$\sum_{n=1}^\infty \exp\left(\left(\beta\log\left(\frac{1}{\alpha}\right) + \varepsilon\right)^{-1}\log n\right) < \infty,$$

and using the bound derived in (9) we have that

$$\sum_{n=1}^\infty \mathrm{P}_x(X_n \leq c) < \infty.$$

Thus, if $\beta\log(1/\alpha) < 1$, the chain is transient while, if $\beta\log(1/\alpha) \geq 1$, it is null recurrent (by Proposition 1 it cannot be positive recurrent). This concludes the proof.

3.6. Proof of Lemma 1

Fix $c > 0$ and $x \in [0, c]$. We use the same set inclusion (11) used in the proof of Theorem 1. In what follows we fix $\delta > 0$ and let the C_i be finite positive constants that may depend on δ. Then, using (10), we have that

$$\mathrm{P}_x(X_n \le c) \ge C_1 \exp\biggl\{-(1+\delta)\sum_{i=1}^{n} \mathrm{P}\biggl(Z_1 > \frac{c}{2}\frac{b}{i^2}\alpha^{-i}\biggr)\biggr\}. \tag{12}$$

The right-hand side of (12) can be bounded below in turn to yield that

$$\begin{aligned}\mathrm{P}_x(X_n \le c) &\ge C_1 \exp\biggl\{-(1+\delta)\sum_{i=1}^{n} \mathrm{P}\biggl(\log(1+Z_1) > \log\biggl(\frac{cb}{2}\biggr) - 2\log i + i\log\biggl(\frac{1}{\alpha}\biggr)\biggr)\biggr\} \\ &\ge C_2 \exp\biggl\{-(1+\delta)\sum_{i=1}^{n} \mathrm{P}(\log(1+Z_1) > i)\biggr\}. \end{aligned} \tag{13}$$

Since $U(x, [0, c]) := \sum_n \mathrm{P}_x(X_n \le c)$, it follows that X is recurrent if the sum over n of the above lower bound diverges. Note that $\varphi(n) = \sum_{i=1}^{n} \mathrm{P}(\log(1+Z_1) > i)$ is monotone nondecreasing in n and thus $\exp(-\varphi(n))$ is monotone nonincreasing in n. It therefore follows that the convergence and divergence of the sum over n of the right-hand side of (13) is equivalent to that of its 'integral version' given in the lemma. In particular,

$$U(x, [0, c]) = \infty \quad \text{if} \int_0^\infty \exp\biggl\{-(1+\delta)\int_0^t \mathrm{P}(\log(1+Z_1) > y)\biggr\}\,\mathrm{d}t = \infty. \tag{14}$$

Finally, by assumption, there exists a choice of $\delta > 0$ for which the integral in (14) diverges. Since $[0, c]$ and $x \in [0, c]$ were arbitrary, the chain is recurrent.

3.7. Proof of Proposition 5

The proof follows straightforwardly from the proof of (ii) in Theorem 1.

4. Concluding remarks

The two main messages in the paper are: (i) AR processes that are 'stable' in terms of the roots of their characteristic polynomial can still exhibit transient and even null-recurrent dynamics that are due to the nature of the innovation process; and (ii) we cannot determine whether an AR process is null recurrent or transient based on a simple moment condition on the innovations. Recall that the log-moment condition is necessary and sufficient for positive recurrence in both the scalar and vector cases. In contrast, it seems that transience and null recurrence are determined by finer properties of the distribution of the innovations. For example, under a more precise tail-decay assumption, we can classify the exact nature of the Markov chain associated with the AR process. Further study is needed to determine the necessary and sufficient conditions on the innovations that support a complete recurrence-classification theory.

Acknowledgement

The authors wish to thank an anonymous referee for his/her careful reading and helpful suggestions. These comments helped clarify and strengthen various statements that were contained in the previous version of the paper.

References

[1] Athreya, K. and Pantula, S. (1986). Mixing properties of Harris chains and autoregressive processes. *J. Appl. Prob.* **23,** 880–892.
[2] Billingsley, P. (1995). *Probability and Measure*, 3rd edn. John Wiley, New York.
[3] Borovkov, A. A. (1998). *Ergodicity and Stability of Stochastic Processes.* John Wiley, Chichester.
[4] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer, New York.
[5] Brockwell, P. J., Resnick, S. I. and Tweedie, R. L. (1982). Storage processes with general release rule and additive inputs. *Adv. Appl. Prob.* **14,** 392–433.
[6] Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Rev.* **41,** 45–76.
[7] Doob, J. L. (1953). *Stochastic Processes*. John Wiley, New York.
[8] Duflo, M. (1997). *Random Iterative Models*. Springer, Berlin.
[9] Meyn, S. P. and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
[10] Porat, B. (1994). *Digital Processing of Random Signals*. Prentice Hall, Englewood Cliffs, NJ.
[11] Rai, S., Glynn, J. E. and Glynn, P. W. (2002). Recurrence classification for a family of non-linear storage models. Tech. Rep., Department of Management Science and Engineering, Stanford University. Available at http://www.stanford.edu/dept/MSandE/faculty/glynn/.
[12] Vervaat, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Prob.* **11,** 750–783.