Recurrence Properties of Autoregressive Processes With Super-Heavy Tailed Innovations

Assaf Zeevi*
Columbia University

Peter W. Glynn†
Stanford University

Submitted: December 17, 2002 Revised: November 2003

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 behavior. 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.

Short Title: AR processes with super-heavy tails

Keywords: autoregressive process, heavy-tails, Markov processes, recurrence, stochastic stability, transience.

2000 AMS 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 modeling and analysis, the focus has been on linear models, namely, autoregressive and moving average models [see, e.g., Brockwell and Davis (1991)]. In signal processing applications, AR models have been

*Graduate School of Business, e-mail: assaf@gsb.columbia.edu
†Management Science and Engineering, e-mail: glynn@stanford.edu

used extensively in filter design, modeling noise processes and quantization effects, and spectral analysis [see, e.g., Porat (1994)]. 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 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 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 $|\alpha| < 1$, which we shall henceforth refer to as the "usual stability condition."

If, in addition to the usual stability condition, we assume that $\mathbb{E}\log(1 + |Z_1|) < \infty$, then X is a positive recurrent Markov chain [see Athreya and Pantula (1986, 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 one assumes the usual stability condition (i.e., $|\alpha| < 1$), and $\mathbb{E}\log(1 + |Z_1|) = \infty$?

The above question essentially concerns the behavior of AR processes that are subject to innovations that assume "very large" values. If Z satisfies the condition $\mathbb{E}\log(1 + |Z_1|) = \infty$, 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 follows

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

Note that in the super-heavy tailed regime, $\max\{Z_k : 1 \leq k \leq n\}$ 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 sub-class of super-heavy tailed distributions. We say that a non-negative random variable Z has *log-Pareto tails* if $\log(1+Z)$

is distributed Pareto with parameters (p, β), i.e., for some $p > 0$ and $\beta > 0$

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

when $z \geq 0$. 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, $\mathbb{E}\log(1+|Z_1|) < \infty$, 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 behavior 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 $\mathbb{E}[\log(1+|Z|)]^p = \infty$ for some $p < 1$, then 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 appears explicitly in Athreya and Pantula (1986). Vervaat (1979) investigates stability of linear stochastic recursions under more general dependence conditions [see also Borovkov (1998)]. Recurrence classification for a class of non-linear generalizations of AR processes is investigated in continuous time by Brockwell, Resnick and Tweedie (1982) and in discrete time by Rai, Glynn and Glynn (2002). 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 are discussed more generally in Borovkov (1998) and Meyn and Tweedie (1993).

2 Main Results

Let us first define more rigorously what we mean by recurrence properties. Consider a Markov chain $X = (X_n :\geq 0)$ on a state space $E \subseteq \mathbb{R}^d$. Let $\mathcal{E} = \mathcal{B}(E)$ denote the associated Borel sigma-field over E. Let $\mathbb{P}_x$ denote the underlying probability measure conditional on $X_0 = x$, and let $\mathbb{E}_x[\cdot] := \mathbb{E}[\cdot|X_0 = x]$. For any $A \in \mathcal{E}$, put $T_A := \inf\{n \geq 1 : X_n \in A\}$ to be the hitting time of A. Let $\|\cdot\|$ denote the usual Euclidean norm, where $|\cdot|$ denotes 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 a an E-valued Markov chain X.

Definition 1 ***(Recurrence)***

i.) X is said to be *recurrent* if there exists a compact set A such that

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

and *transient* otherwise.

ii.) X is said to be *positive recurrent* if there exists a compact set A such that

$$\mathbb{E}_x T_A < \infty, \quad \text{for all } x \in E$$

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

The definition of 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 Meyn and Tweedie (1993)]. An alternative definition of recurrence uses the notion of *potential*. In particular, let

$$\eta_A = \sum_{n=1}^{\infty} \mathbb{I}_{\{X_n \in A\}} \tag{4}$$

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

$$U(x, A) := \mathbb{E}_x[\eta_A] = \sum_{n=1}^{\infty} \mathbb{P}_x(X_n \in A) \tag{5}$$

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 $\mathbb{E}\log(1 + |Z_1|) < \infty$, 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 $\mathbb{E}\log(1+|Z_1|) < \infty$, in which case X admits a unique stationary distribution, π, and $\mathbb{P}_x(X_n \in \cdot) \Rightarrow \pi(\cdot)$, as $n \to \infty$.*

Proposition 1 indicates that the stability condition $|\alpha| < 1$ together with 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 Athreya and Pantula (1986) and necessity for the case where Z is nonnegative is given in Brockwell et al. (1982) and Rai et al. (2002).

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

$$X_{n+1} = AX_n + Z_{n+1} \tag{6}$$

where $A \in \mathbb{R}^{d \times d}$ is nonsingular, and X_n and Z_n takes values in $\mathbb{R}^d$. Assume that the innovation process $Z = (Z_n : n \geq 1)$ is comprised 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 follows

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

Let $\|v\|$ denote the Euclidean norm of a vector in $v \in \mathbb{R}^d$. The following is an analogue of Proposition 1.

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

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

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

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

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

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 $\mathbb{E}\log(1+|Z_1|) < \infty$ and $\mathcal{P}(z) = \det(A - zI)$ has all its roots strictly inside the unit disk 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, e.g., in Athreya and Pantula (1986, 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 behavior that AR processes can exhibit in this super-heavy tailed setting. To this end, we will restrict attention to non-negative innovation processes and $0 < \alpha < 1$ 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 $A = [0, c]$, T_A, and the potential function, $U(x, A)$, for the process started at $X_0 = x$. Proposition 4 asserts that the notion of recurrence articulated in Definition 1 is equivalent to non-finiteness of the potential function.

Proposition 3 *Let X be an AR(1) process with nonnegative innovations and $\alpha \in (0,1)$. Then, setting $A = [0,c]$ we have that for all $x \geq 0$*

$$U(x,A) = \infty \quad \textit{if and only if} \quad \mathbb{P}_x(T_A < \infty) = 1 \ . \tag{7}$$

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. Duflo (1997) and the recent survey by Diaconis and Freedman (1999)]. 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

$$S_n = \alpha^n X_0 + \sum_{i=1}^{n} \alpha^{i-1} Z_i \ . \tag{8}$$

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 behavior 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 limit 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 in Diaconis and Freedman (1999).]

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

$$\tau(c) = \inf\{n \geq 1 : S_n > c\} \ . \tag{9}$$

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 $c > 0$, $\mathbb{P}_0(\tau(c) = \infty) > 0$;

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

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

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 $\beta \log(1/\alpha) < 1$ then 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 $\mathbb{E}(\log(1+|Z_1|)) < \infty$ 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 $\sup\{p : \mathbb{E}(\log(1+|Z_1|))^p < \infty\} < 1$ 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 $\mathbb{E}(\log(1+|Z_1|))^p = \infty$ 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 non-negative innovations and $\alpha \in (0,1)$. If for some $\delta > 0$*

$$\int_0^\infty \exp\left\{-(1+\delta)\int_0^t \mathbb{P}\left(\log(1+Z_1) > y\right)\right\} dt = \infty \ ,$$

then X is recurrent.

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

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

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

Note that in order to guarantee the divergence of $\mathbb{E}W$, 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

$$\psi(t) := \exp\left\{-2\int_0^t \mathbb{P}(W > y)dy\right\}.$$

Suppose that at $t = t_n$, we have $\psi(t_n) = a_n$. We then have for $t > t_n$, that

$$\psi(t) = a_n \exp\left\{-2\int_{t_n}^t \mathbb{P}(W > y)dy\right\} .$$

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

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

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

$$\int_{t_n}^{t_{n+1}} \psi(t)dt = \frac{a_n}{2p_{n+1}} \left(1 - \exp\{-2p_{n+1}(t_{n+1} - t_n)\}\right)$$

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

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

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

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

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 set p_{n+1} as above to comply with the constraint $p_n \searrow$ as $n \to \infty$.) Consequently, $\int_0^\infty \psi(t)dt$ diverges and the chain must therefore be null recurrent in spite of the fact that $\mathbb{E}[\log(1 + Z_1)]^{1/2} = \infty$.

The above example demonstrates that we cannot expect in general to distinguish between null recurrence and transient behavior 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 is the intuition behind the counter example described above. 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 behavior of the maximum of the Z_i's will not exhibit "smooth" monotonic increasing behavior, 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, one 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's take on values that are in the far tail (quantile $\gg 1 - 1/n$), 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's 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's is encountered. This, in turn suggests that the chain ought to exhibit recurrent rather than transient behavior.

While it turns out that $\mathbb{E}[\log(1+Z_1)]^p = \infty$ 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*

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

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

3 Proofs

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, is proved in Athreya and Pantula (1986, Proposition 1). To prove necessity, we use convergence properties of random series. Theorem 2.7 (ii) in Doob (1953, p. 115) asserts that if W_i is a sequence of independent r.v.'s, then $S_n = \sum_{i=1}^n W_i \Rightarrow S$, if and only if, $S_n \to S$ almost surely. To this end, 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 of $|\alpha| < 1$, in particular, suppose that $|\alpha| \geq 1$. Then, for any $\epsilon > 0$ we have that $\mathbb{P}(|\alpha^i Z_i| > \epsilon) \geq \mathbb{P}(|Z_1| \geq \epsilon)$, for all i. Thus, $\sum_i \mathbb{P}(|\alpha^i Z_i| \geq \epsilon) < \infty$ for all $\epsilon > 0$ if and only if $Z_i = 0$ for all i. By Kolmogorov's three series theorem [see, e.g., Billingsley (1995, Theorem 22.8, p. 290)] this is necessary for the almost sure convergence of S_n and thus [by Doob (1953, 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 $\mathbb{E}\log(1+|Z_1|) = \infty$ implies that

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

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. $\blacksquare$

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.

Sufficiency: To argue sufficiency note that

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

and the first term on the right hand side converges to zero almost surely as $n \to \infty$, since $\rho(A) < 1$ implies $\|A^m\| < 1$ 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 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 $\rho(A) < 1$ we have that

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

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_∞.

Necessity: To argue necessity note that we may apply Theorem 2.7 (ii) in Doob (1953, 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

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

and consequently $\|A^n Z_n\| \ge \|Z_n\| \|(A^n)^{-1}\|^{-1} \ge \|Z_n\| \|A^n\|$. Now, if $\mathbb{E} \log(1 + \|Z_1\|) = \infty$ 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 limsup is infinite. Since the aforementioned component of S_n does not converge almost surely, it follows [by Doob (1953, 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. $\blacksquare$

Proof of Proposition 3. We need to prove the equivalence between the behavior of the potential and probability of the return time to compact intervals $T_A = \inf\{n \ge 1 : X_n \in A\}$.

Suppose first that $\mathbb{P}_x(T_A < \infty) = 1$. Consider the case where Z_1 has bounded support, where we may have $\mathbb{P}_x(T_A = 1) = 1$ for the particular choice of x. If this holds for all $x \in A$ then clearly the chain never leaves A and the proof follows, otherwise there exists $x' > x$ in A such that $\mathbb{P}_{x'}(T_A = 1) < 1$. Thus, assume without loss of generality that $\mathbb{P}_x(T_A = 1) < 1$. Conditioning on the chain at time 1, gives the expression

$$\mathbb{P}_x(T_A < \infty) = \int_{y \in A} P(x, dy) + \int_{y \in A^c} \mathbb{P}_y(T_A < \infty) P(x, dy)$$

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

$$p := \mathbb{P}_0(T_A < \infty) < 1.$$

Hence, a "geometric trials" argument establishes that

$$\mathbb{P}_0(\eta_A = n) \leq p^n,$$

and consequently $U(0, A) = \sum_{n=1}^{\infty} \mathbb{P}_0(\eta_A = n) < \infty$. By stochastic monotonicity of X we have that $U(x, A) \leq U(0, A)$ for all $x \geq 0$ and therefore the expected occupation measure is sigma-finite. This concludes the proof. ■

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 $c > 0$ and $A = [0, c]$. Now, using the definition of the backward process given in (8) and the first exit time of that process from A given in (9) it follows that

$$\mathbb{E}_0 \eta_A = \mathbb{E}_0 \tau(c) - 1\ .$$

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

Thus,

$$\begin{aligned}\mathbb{P}_0(\tau(c)=\infty) &= \mathbb{P}_0(\cap_n E_n) \\ &= \lim_{n\to\infty} \mathbb{P}_0(S_n \in A) \\ &= \lim_{n\to\infty} \mathbb{P}_0(X_n \in A)\ ,\end{aligned}$$

since S_n is equal in distribution to X_n for all n. Now, if $\mathbb{P}_0(\tau(c)=\infty)>0$ then it follows that $\lim_{n\to\infty}\mathbb{P}_0(X_n \in A)>0$ 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 we have that $\mathbb{P}_0(X_n \in A)\to \mathbb{P}_\pi(X_1\in A)$ as $n\to\infty$, and the latter must be positive for some $c>0$. This concludes the proof. ∎

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

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

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

$$U(x,A)=\sum_{n=1}^{\infty}\mathbb{P}_x(X_n\le c)$$

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

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

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,

$$\begin{aligned}\mathbb{P}_x(X_n\le c) &\le \mathbb{P}(Z_i\le c\alpha^{-i},\ i=1,\dots,n) \\ &= \prod_{i=1}^{n}\left[1-\mathbb{P}(Z_1>c\alpha^{-i})\right] \\ &= \exp\left\{\sum_{i=1}^{n}\log\left[1-\mathbb{P}(Z_1>c\alpha^{-i})\right]\right\} \\ &\le \exp\left\{-\sum_{i=1}^{n}\mathbb{P}(Z_1>c\alpha^{-i})\right\}, \qquad (10)\end{aligned}$$

where the first step follows from the above set inclusion, 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

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

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

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

for sufficiently large n, and therefore

$$U(x, A) = \sum_{n=1}^{\infty} \mathbb{P}_x(X_n \leq c) < \infty \ .$$

Since c was arbitrary, the chain is transient.

Case iii.) $\boldsymbol{p = 1}$: Fix $c > 0$, let $A = [0, c]$ and fix $x \in A$. There are two subcases to deal with.

1. $\boldsymbol{\beta \log(1/\alpha) \geq 1}$: We use the following set inclusion

$$\{X_n \leq c\} \supseteq \{Z_{n-i} \leq (c/2)(b/(i+1)^2)\alpha^{-i}, \ i = 0, \ldots, n-1\}, \tag{11}$$

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 $x\alpha^n \leq c/2$. Observe that if $Z_{n-i} \leq (c/2)(b/(i+1)^2)\alpha^{-i}$ for $i = 0, \ldots, n-1$ then,

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

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

$$\begin{aligned}
\mathbb{P}_x(X_n \leq c) &\geq \mathbb{P}(Z_i \leq (c/2)(b/i^2)\alpha^{-i}, \ i = 1, \ldots, n) \\
&= \exp\left\{ \sum_{i=1}^{n} \log\left[1 - \mathbb{P}(Z_1 > (c/2)(b/i^2)\alpha^{-i})\right] \right\} \\
&\geq C_1 \exp\left\{ -\sum_{i=1}^{n} \mathbb{P}(Z_1 > (c/2)(b/i^2)\alpha^{-i}) - \sum_{i=1}^{n} \left(\mathbb{P}(Z_1 > (c/2)(b/i^2)\alpha^{-i})\right)^2 \right\} \\
&=: C_1 \exp\{-R_n - Q_n\} \ ,
\end{aligned} \tag{12}$$

where we have used the inequality $\log(1-y) \geq -y - y^2$ which holds for all $0 \leq y \leq 1/2$, and where C_1 is a finite positive constant. Using the assumption on the distribution of Z_1, we have

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

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

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

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

$$\int_a^n \frac{1}{x - k\log x} dx = \log(x - k\log x)|_a^n + \int_a^n \frac{k/x}{x - k\log x} dx .$$

(We 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

$$\int_a^\infty \frac{k/x}{x - k\log x} dx < \infty .$$

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

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

which holds for all n sufficiently large and finite positive constants C_2, C_3. Combining the above with (12) we conclude that $\sum_n \mathbb{P}_x(X_n \leq c)$ diverges when $\beta\log(1/\alpha) \geq 1$.

2. $\boldsymbol{\beta\log(1/\alpha) < 1}$: In this case we have that

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

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

$$\sum_{i=1}^n \mathbb{P}(Z_1 > c\alpha^{-i}) \geq (\beta\log(1/\alpha) + \epsilon)^{-1} \log n,$$

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

$$\sum_{n=1}^\infty \exp\Big((\beta\log(1/\alpha) + \epsilon)^{-1} \log n\Big) < \infty ,$$

and using the bound derived in (10) we have that

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

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. ∎

Proof of Lemma 1. Fix $c > 0$, let $A = [0, c]$ and fix $x \in A$. We use the same set inclusion used in (11) in the proof of Theorem 1. In what follows we fix $\delta > 0$ and let C_i be finite positive constants that may depend on δ. Then, using (12) we have that

$$\mathbb{P}_x(X_n \leq c) \geq C_1 \exp\left\{-(1+\delta) \sum_{i=1}^n \mathbb{P}(Z_1 > (c/2)(b/i^2)\alpha^{-i})\right\} .$$

The RHS of the above can be lower bounded in turn to yield

$$\begin{aligned}\mathbb{P}_x(X_n \le c) &\ge C_1 \exp\left\{-(1+\delta)\sum_{i=1}^{n} \mathbb{P}(\log(1+Z_1) > \log(cb/2) - 2\log i + i\log(1/\alpha))\right\} \\ &\ge C_2 \exp\left\{-(1+\delta)\sum_{i=1}^{n} \mathbb{P}(\log(1+Z_1) > i)\right\} .\end{aligned}$$

Since $U(x,A) := \sum_n \mathbb{P}_x(X_n \le c)$, we have that X is recurrent if the sum over n of lower bound above diverges. Note that $\varphi(n) = \sum_{i=1}^{n} \mathbb{P}(\log(1+Z_1) > i)$ is monotone non-decreasing in n and thus $\psi(n) = \exp(-\varphi(n))$ is monotone non-increasing in n. We therefore have that the convergence and divergence of the sum over n of the RHS above is equivalent to that of its "integral version" given in the lemma. In particular,

$$U(x,A) = \infty \quad \text{if} \quad \int_0^\infty \exp\left\{-(1+\delta)\int_0^t \mathbb{P}(\log(1+Z_1) > y)\right\} dt = \infty \ .$$

Finally, by assumption there exists a choice $\delta^* > 0$ for which the aforementioned integral diverges. Since A and $x \in A$ were arbitrary, the chain is recurrent. ■

Proof of Proposition 5. The proof follows straightforwardly from the proof of case ii.) in Theorem 1. ■

4 Concluding Remarks

The two main messages in the paper are: (1) 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, (2) one 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 one can classify the exact nature of the Markov chain associated with the AR process. Further study is needed to determine what are 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

Athreya, K. and Pantula, S. (1986), Mixing properties of Harris chains and autoregressive processes, *J. Appl. Probab.* **23**, 880–892.

Borovkov, A. A. (1998), *Ergodicity and Stability of Stochastic Processes.*, Wiley & Sons.

Billingsley, P. (1995), *Probability and Measure*, Wiley.

Brockwell, P. and Davis, R. (1991), *Time Series: Theory and Methods*, Springer.

Brockwell, P., Resnick, S. and Tweedie, R. (1982), Storage processes with general release rule and additive inputs, *Adv. in Appl. Probab.* **12**, 392–433.

Diaconis, P. and Freedman, D. (1999), Iterated random functions, *SIAM Review* **41**, 45–67.

Doob, J. (1953), *Stochastic Processes*, Wiley & Sons.

Duflo, M. (1997), *Random Iterative Models*, Springer Verlag.

Meyn, S. and Tweedie, R. (1993), *Markov Chains and Stochastic Stability*, Springer.

Porat, B. (1994), *Digital Processing of Random Signals*, Pretice Hall.

Rai, S., Glynn, J.E. and Glynn, P.W. (2002), Recurrence classification for a family of non-linear storage models, Technical report, Dept. of Management Science and Engineering, Stanford Univeristy.

Vervaat, W. (1979), On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables, *Adv. in Appl. Probab.* **11**, 750–783.