

PIGGYBACKING THRESHOLD PROCESSES WITH A FINITE STATE MARKOV CHAIN

THOMAS R. BOUCHER

Department of Mathematics, Plymouth State University, Plymouth, NH 03264-1595, USA tboucher1@mail.plymouth.edu

DAREN B. H. CLINE

Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA dcline@stat.tamu.edu

Received 10 October 2008

The state-space representations of certain nonlinear autoregressive time series are general state Markov chains. The transitions of a general state Markov chain among regions in its state-space can be modeled with the transitions among states of a finite state Markov chain. Stability of the time series is then informed by the stationary distributions of the finite state Markov chain. This approach generalizes some previous results.

Keywords: Ergodicity; Markov chain; nonlinear autoregressive time series; nonlinear time series; threshold autoregressive time series.

1. Introduction

Threshold processes are some of the simplest nonlinear time series, yet they can possess very complicated behavior. Though much effort has been spent in delineating the stable and unstable parameter spaces for these processes, their behavior is fully understood for only the very simplest models. Conditions sufficient for large classes of models are often too restrictive when applied in specific cases, and it is in the more specialized cases that the utility of these time series models in explaining features found in data is fully realized.

The threshold autoregressive (TAR) process of order p and l regimes is the piecewise linear autoregression

$$Y_{t} = \phi_{0}^{(i)} + \phi_{1}^{(i)} Y_{t-1} + \dots + \phi_{p_{i}}^{(i)} Y_{t-p_{i}} + \sigma_{i} e_{t},$$

if $(Y_{t-1}, \dots, Y_{t-p})' \in R_{i}, \quad i = 1, \dots, l,$ (1)

where the state-space $\mathcal{X} := \mathbb{R}^p$ is partitioned into regions $R_i, i = 1, \ldots, l$, the boundaries of which are called the *thresholds* of the process. The autoregression coefficients are $\phi_0^{(i)}, \ldots, \phi_{p_i}^{(i)}$, with $p_i \leq p$, and the σ_i are positive scalars, $i = 1, \ldots, l$. The $\{e_t\}$ are mean zero i.i.d. random variables with $Var(e_t) = 1$. We assume throughout that the error distribution possesses an everywhere positive density which is lower semi-continuous.

The TAR process in (1) has the state-space representation $X_t = (Y_t, Y_{t-1}, \ldots, Y_{t-p+1})'$, where

$$X_t := \sum_{i=1}^{l} (\phi_0^{(i)} + A_i X_{t-1} + \nu_t^{(i)}) I_{X_{t-1} \in R_i},$$
(2)

with $\nu_t^{(i)} = \sigma_i(e_t, 0, \dots, 0)'$. The A_i are called the *companion matrices* and are defined using the autoregression coefficients.

We can extend (2) a bit and define as *TAR-like* any process of the form

$$X'_{t} = \sum_{i=1}^{l} (A_{i}X'_{t-1} + g(X'_{t-1}) + \nu_{t}^{(i)})I_{X'_{t-1} \in R_{i}},$$
(3)

where $g(\cdot)$ is locally bounded, measurable, and $\limsup_{\|x\|\to\infty} \|g(x)\|/\|x\| = 0$.

One method for analyzing the behavior of a time series is *piggybacking* [3], which involves modeling the behavior of the time series with that of a simpler embedded process. Traditionally, the embedded process has been a deterministic system known as the *homogeneous skeleton* ([2, 6-8]) of the state-space process

$$x_t := \sum_{i=1}^l A_i x_{t-1} I_{x_{t-1} \in R_i}.$$
 (4)

Conditions for stability of the time series are then "piggybacked" upon those for the embedded process, i.e. they are derived through analysis of this more tractable embedded process.

Example 1. Consider the TAR process analyzed by Petrucelli and Woolford [5]

$$Y_t = \begin{cases} a_1 Y_{t-1} + e_t, & \text{if } Y_{t-1} \ge 0, \\ a_2 Y_{t-1} + e_t, & \text{if } Y_{t-1} < 0. \end{cases}$$

Suppose $a_1 < 0$, $a_2 < 0$ and $E|e_t|^r < \infty$ for some r > 0. The skeleton is $y_t = a_1y_{t-1}I_{y_{t-1}\geq 0} + a_2y_{t-1}I_{y_{t-1}<0}$. Let $R_1 := \{y : y < 0\}$ and $R_2 := \{y : y \ge 0\}$, and let $R_i \to R_j$ indicate all y in R_i which are mapped by the skeleton to R_j . Then for $y \neq 0$ we have $R_1 \to R_2$, $R_2 \to R_1$. By picking $Y_{t-1} = y$ with |y| large enough, the transition probabilities of Y_t between regions R_1 and R_2 can be bounded arbitrarily closely to zero or one for all such y. Thus, the skeleton $\{y_t\}$ accurately models the behavior of $\{Y_t\}$ for large values of the processes, making appropriate the piggyback method using the skeleton $\{y_t\}$ as the embedded process. Stability of the time series is then inferred from stability of the skeleton, resulting in the well-known constraint on the parameters $a_1a_2 < 1$.

However, the skeleton-piggyback requires that the skeleton and the time series share similar growth behaviors, in the sense that the probabilities of the transitions among regions in the state-space for both the time series and the skeleton converge as the two grow larger in magnitude. This means that the time series is essentially a well-behaved dynamical system when it is large in magnitude.

Example 2. Now consider the second order TAR process

$$Y_{t} = \begin{cases} a_{0} + a_{1}Y_{t-1} + a_{2}Y_{t-2} + \sigma_{1}e_{t}, & \text{if } Y_{t-1} \ge b_{1}Y_{t-2}, \\ b_{0} + b_{1}Y_{t-1} + \sigma_{2}e_{t}, & \text{if } Y_{t-1} < b_{1}Y_{t-2}. \end{cases}$$
(5)

Assume $E|e_t|^r < \infty$ for some r > 0 and the parameters satisfy

$$a_1 > 0, \quad a_2 > 0, \quad b_1 < 0, \quad a_1 b_1 + a_2 < 0.$$
 (6)

The state vector for the time series is $X_t = (Y_t, Y_{t-1})'$. The state-space representation is given in (2) with order p = 2, number of regimes l = 2, the regions given by

$$R_1 = \{(y_1, y_2)' : y_1 \ge b_1 y_2\}, \quad R_2 = \{(y_1, y_2)' : y_1 < b_1 y_2\},\$$

and companion matrices

$$A_1 = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} b_1 & 0 \\ 1 & 0 \end{pmatrix}.$$

The skeleton $\{x_t\} = \{(y_t, y_{t-1})'\}$, defined in (4), maps all $x \in R_2$ onto the threshold $y_1 = b_1 y_2$. However, $\{X_t\}$ can move to either side of the threshold according to the probability distribution of e_t . When the errors are taken into account, therefore, the transition probabilities between R_1 and R_2 cannot be bounded arbitrarily closely to zero or one no matter how large the process is. Thus, the skeleton-piggyback will not work here.

Example 2 suggests the piggyback should be extended through the use of a stochastic, rather than a deterministic, embedded process. In the next section we apply the piggyback method more generally using a finite state Markov chain as the embedded process, yielding stability conditions for a wider class of processes than is achieved using the skeleton-piggyback. Examples 1 and 2 are continued in the next section. They are admittedly simple and are intended to be illustrative; more substantial examples are dealt with in Sec. 3.

2. Results

2.1. Piggybacking with a finite state chain

Since the state-space of $\{X_t\}$ for a TAR or TAR-like process can be partitioned into a finite number of regions, the transitions of $\{X_t\}$ among these regions can be modeled more generally with the transitions of a finite state Markov chain, rather than those of a deterministic system. In the case of Example 1, note a (trivial) finite state chain $\{J_t\}$ on state-space $\{1, 2\}$ with transition probability matrix

$$P := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

will approximate the transition probabilities of $\{X_t\}$ between R_1 and R_2 to within an arbitrary degree of accuracy when $\{X_t\}$ is large. Further extension of the piggyback method uses nontrivial finite state chains as models for the transitions among regions. In fact, the full utility of the finite state chain-piggyback is in accounting for cases with multiple and/or nontrivial stationary distributions for the finite state chain. Some examples of this more complex situation are in Sec. 3.

First, the details. The finite state chain-piggyback makes use of a finite state Markov chain $\{J_t\}$ on a collection of states $\{1, \ldots, l\}$ corresponding to a collection of regions $\{R_1, \ldots, R_l\}$ that forms a partition of the state-space of $\{X_t\}$.

Assumption 1 (A1). Suppose $\{J_t\}$ is a finite state Markov chain on the states $S = \{1, \ldots, l\}$. Decompose the state space $S = (\bigcup_{u=1}^k S_u) \cup T$, for some finite integer k, where each S_u is irreducible and recurrent and T is the set of all transient states. Let $G := \bigcup_{u=1}^k S_u$. Let $\pi^{(u)}$ be the stationary distribution for S_u , so that $\pi_j^{(u)} > 0$ for $j \in S_u$ and $\pi_j^{(u)} = 0$ for $j \notin S_u$.

Note that $\{J_t\}$ is not necessarily ergodic since it is not assumed to be irreducible, but since $\{J_t\}$ is a finite state chain, it must have at least one collection of recurrent states, and therefore at least one stationary distribution is always guaranteed to exist. Next, sufficient conditions under which the finite state chain-piggyback will succeed. Let P_i , P_x denote probabilities conditioned upon initial states $i \in \{1, \ldots, l\}$ and $x \in \mathcal{X}$, respectively, and let $\|\cdot\|$ be the Euclidean norm.

Assumption 2 (A2). Suppose $\{X_t\}$ is a general state Markov chain on state-space $\mathcal{X} := \mathbb{R}^p$ which can be partitioned into regions $\{R_1, \ldots, R_l\}$, and these regions can be grouped by their indices according to the sets G and T in (A1), in the sense that

$$\max_{\substack{i,j\in G\\ \|x\|\to\infty}} \limsup_{\substack{x\in R_i\\ \|x\|\to\infty}} |P_x(X_1\in R_j) - P_i(J_1=j)| = 0,$$
(7)

and given $\epsilon > 0$ there exists $t^* < \infty$ so that for $t \ge t^*$

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} P_x \left(X_t \in \bigcup_{k \in T} R_k \right) < \epsilon.$$
(8)

In words, for regions in the state-space of $\{X_t\}$ corresponding to the recurrent states of $\{J_t\}$, the transition probabilities of $\{J_t\}$ from state *i* to state *j* must be within an arbitrarily small amount of the "transition" probabilities of $\{X_t\}$ from region R_i to region R_j when $\{X_t\}$ is large enough. This is condition (7). Since $\{J_t\}$ is a finite state chain, the collection of transient states is uniformly transient, and the similar condition (8) on the regions in the state-space of $\{X_t\}$ corresponding to the transient states of $\{J_t\}$ is required.

A word about application: verifying (A2) can require an artful partitioning of the state-space. The regimes specified by the original model definition (4) often need to be refined, that is, partitioned further so that (A2) holds. For example, as the behavior of the skeleton can depend on the signs of the variables, the axes become de facto thresholds. Additionally, regions that cannot be given precise transition probabilities, but ultimately are transient, may be separated from artificial thresholds. Each time another threshold is created, however, the transition probabilities must be considered anew, especially for situations where the skeleton either hits or is attracted to a threshold.

Example 2 (Cont'd). Returning to the process (5) with parameters satisfying (6), we first refine R_1 and R_2 by splitting each into two regions:

$$R_1^+ = \{(y_1, y_2)' : y_1 \ge \max(0, b_1 y_2)\}, \quad R_1^- = \{(y_1, y_2)' : b_1 y_2 \le y_1 < 0\},\$$

$$R_2^+ = \{(y_1, y_2)' : 0 < y_1 < b_1 y_2\}, \quad R_2^- = \{(y_1, y_2)' : y_1 < \min(0, b_1 y_2)\}.$$

Essentially, we are making the vertical axis into another threshold. See Fig. 1.

The skeleton maps $x = (y_1, y_2) \in R_1^+$ into the first quadrant; in fact,

$$a_1y_1 + a_2y_2 \ge -\frac{a_1b_1 + a_2}{\sqrt{1 + b_1^2}} ||x|| > 0.$$

Thus, as $||x|| \to \infty$,

$$P_x(X_1 \notin R_1^+) \le P\left(A_0 + \frac{|a_1b_1 + a_2|}{\sqrt{1 + b_1^2}} \|x\| + \sigma_1 e_1 \le 0\right) = o(\|x\|^{-r}),$$



Fig. 1. Regions of the state-space in Example 2.

by Markov's inequality and the fact that $E|e_1|^r < \infty$. For the finite state chain, then $R_1^+ \to R_1^+$ w.p. 1. On the other hand, R_1^- is mapped into any of the other three regions (but not itself) with probabilities $\alpha_1^+(x), \alpha_2^+(x), \alpha_2^-(x)$, respectively, that converge as $||x|| \to \infty$ and $x/||x|| \to \theta$ but not uniformly in θ .

The skeleton maps $x = (y_1, y_2) \in R_2^+$ onto the threshold in the second quadrant. If $b_0 + \sigma_2 e_1 < 0$, then X_1 will be below the threshold and $X_1 \in R_2^-$. If $0 \le b_0 + \sigma_2 e_1 < -(a_1y_1 + a_2y_2)$ then $X_1 \in R_1^-$, and if $b_0 + \sigma_2 e_1 \ge -(a_1y_1 + a_2y_2)$, which has probability approaching zero as $||x|| \to \infty$, then $X_1 \in R_1^+$. For $x \in R_2^-$, a similar transition is possible but x is mapped to the threshold in the fourth quadrant and the choice, asymptotically, is between R_2^+ and R_1^+ .

Let $\gamma_2 = P(b_0 + \sigma_2 e_t < 0)$. With a slight abuse of notation, order the finite state space of $\{J_t\}$ as $\{R_1^+, R_1^-, R_2^+, R_2^-\}$, and we now have the transition matrix

$$P = \begin{pmatrix} 1 & 0 & 0 & 0\\ \alpha_1^+(x) & 0 & \alpha_2^+(x) & \alpha_2^-(x)\\ 0 & 1 - \gamma_2 & 0 & \gamma_2\\ 1 - \gamma_2 & 0 & \gamma_2 & 0 \end{pmatrix}.$$
 (9)

Examination of P reveals that R_1^+ is the only recurrent state; thus the stationary distribution for J_t is $\pi = (1, 0, 0, 0)'$. Moreover, R_1^- is escaped immediately each time it is hit. Set $\bar{\gamma} = \max(\gamma_2, 1-\gamma_2) < 1$ and so the verification of (A2) is completed by observing

$$\limsup_{\substack{x \notin R_1^+ \\ \|x\| \to \infty}} P_x(X_t \notin R_1^+) \le \bar{\gamma}^{t/2-1}, \quad \text{for all } t \ge 1.$$

2.2. Stability using the finite state chain-piggyback

For the vector norm $\|\cdot\|$ on the state-space and a companion matrix A_i define

$$\rho_{i,I} := \liminf_{\substack{x \in R_i \\ \|x\| \to \infty}} \frac{\|A_i x\|}{\|x\|}, \quad \rho_{i,S} := \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} \frac{\|A_i x\|}{\|x\|}.$$
 (10)

The stability results in Theorem 1 rely upon $\rho_{i,I}$ and $\rho_{i,S}$, rather than the operator norm, since stability is a question of the behavior of the process when the process is large. In many cases the three coincide. Naturally, it is assumed throughout that each $||A_i||$ is finite, and clearly $\rho_{i,I} \leq \rho_{i,S} \leq ||A_i||$ for each *i*.

Heuristically, when the process $\{X_t\}$ is large, the expected log-change in $\{X_t\}$ is (at worst) approximately $E[\sum \log(\rho_{i,S})I_{X_{t-1}\in R_i}]$. Under (A2), the transitions of $\{X_t\}$ among regions are similar to those of $\{J_t\}$ among states. Since $\{J_t\}$ has a finite number of states, then $E[\log(\rho_{J_t,z})]$ will converge to $\sum_u q_u(i)E_{\pi^{(u)}}[\log(\rho_{J_t,z})]$, for z = I or z = S where $q_u(i)$ is the probability that J_t ends up in S_u given that $J_0 = i$ and $\pi^{(u)}$ is the corresponding stationary distribution. Then $\sum \pi_i \log(\rho_{i,S}) < 0$ will guarantee stability of $\{X_t\}$, while $\sum \pi_i \log(\rho_{i,I}) > 0$ will guarantee the transience of $\{X_t\}$. We define stability in terms of the V-uniform ergodicity of the chain. The

various types of ergodicity or the transience of Markov chains are demonstrated by constructing a *test function* of the chain that satisfies an appropriate drift condition. Full details are in Meyn and Tweedie [4], and a brief summary precedes the proofs in Sec. 5.

Theorem 1. Suppose (A1) and (A2) hold. Consider $\{X_t\}$ as in (3). Assume there exists r > 0 for which $E|e_t|^r < \infty$.

(i) Suppose for every stationary distribution $\pi_j^{(u)}$ of $\{J_t\}, u \in \{1, \ldots, k\}$, that

$$\sum_{i=1}^{l} \pi_i^{(u)} \log(\rho_{i,S}) < 0.$$
(11)

Then there exists a bounded function $\lambda(x)$ such that $\{X_t\}$ is V-uniformly ergodic with test function $V(x) = 1 + \lambda(x) ||x||^s$, s < r.

(ii) Suppose for some stationary distribution $\pi_j^{(u)}$ of $\{J_t\}, u \in \{1, \ldots, k\}$, that

$$\sum_{i=1}^{l} \pi_i^{(u)} \log(\rho_{i,I}) > 0, \tag{12}$$

and that $\limsup_{\substack{\|x\|\to\infty\\x\in Q}} \|x\|^r P_x(X_1 \notin Q) = 0$, where $Q = \bigcup_{j\in S_u} R_j$. Then there exists a bounded function $\lambda(x)$ such that $\{X_t\}$ is transient with test function $V(x) = 1 + \lambda(x) \|x\|^s$, s < r.

Example 2 (Cont'd). It was noted that R_1^+ is the only recurrent state. Since the regions are all cones, $||A_i||$, $\rho_{i,I}$, and $\rho_{i,S}$ all coincide for all *i* where the eigenvector corresponding to the maximal eigenvalue of A_i is contained in region R_i . Thus, by Theorem 1, under the conditions $||A_1|| < 1$, or equivalently $a_1 + a_2 < 1$, and $E|e_t|^r < \infty$ for some r > 0, it holds that $\{X_t\}$ is *V*-uniformly ergodic, with *V* as in Theorem 1. Conversely, $\{X_t\}$ is transient if $a_1 + a_2 > 1$ since the error distribution satisfies the additional condition $||x||^r P_x(X_1 \notin R_1^+) \to 0$ as previously shown.

Interestingly, for stability the assumptions $b_1 < 0$, $a_1b_1 + a_2 < 0$ and $b_1^2 > a_1b_1 + a_2$ place no upper bounds on $|b_1|$ if $b_1 < 0$. In particular, $|b_1| < 1$, which would be the condition generalizing from the linear case, is far too restrictive. In fact, there is a positive probability of temporary "explosions" if $|b_1| \gg 1$. Capturing such behavior would be one of the real advantages of this model and is lost if relying on conditions analogous to the linear case.

As Example 2 shows, a simpler formulation of Theorem 1, in terms of $||A_i||$, has the advantage that the regions can be defined somewhat crudely. This may make it possible to provide a sufficient, if not sharp, condition for ergodicity without a detailed analysis of the dynamics. On the other hand, one can often improve the stability condition by successively refining the regions and using Theorem 1 as expressed here to analyze these more complex dynamics. Examples of this will be delayed until Sec. 3.

2.3. TAR-like processes

Neither (A1) nor (A2) require the time series to be a TAR process, and Theorem 1 permits some generalization. Piggybacking with a finite state Markov chain can be extended to TAR-like processes

$$X'_{t} = \sum_{i=1}^{l} (A_{i}X'_{t-1} + g(X'_{t-1}) + \nu_{t}^{(i)})I_{X'_{t-1} \in R_{i}},$$

via an iterated piggyback using a TAR skeleton

$$X_t = \sum_{i=1}^{l} (\phi_0^{(i)} + A_i X_{t-1} + \nu_t^{(i)}) I_{X_{t-1} \in R_i}.$$

If the TAR skeleton satisfies (A2) and the conditions of Theorem 1, then $\{X'_t\}$ can be piggybacked with $\{X_t\}$ and $\{X_t\}$ can be piggybacked with the finite state Markov chain $\{J_t\}$. Note that Theorem 1 could also be expressed more generally as there being a measurable set B such that the assumptions hold on the set B with

$$\limsup_{\substack{\|x\|\to\infty\\x\in B}} \|g(x)\|/\|x\| = 0, \quad \limsup_{\|x\|\to\infty} P_x(X_1 \notin B) = 0,$$

but we restrict our attention to the case $B = \mathcal{X}$.

Certain smooth threshold autoregressive (STAR) processes and the amplitudedependent exponential autoregressive (EXPAR) processes [7] are TAR-like time series.

Example 2 (Cont'd). Consider the TAR-like EXPAR/TAR hybrid [7] process which generalizes Example 2

$$Y_{t} = \begin{cases} (a_{1} + c_{1}e^{-[Y_{t-1}^{2} + Y_{t-2}^{2}]})Y_{t-1} \\ + (a_{2} + c_{2}e^{-[Y_{t-1}^{2} + Y_{t-2}^{2}]})Y_{t-2} + \sigma_{1}e_{t}, & \text{if } Y_{t-2} \ge \frac{1}{b_{1}}Y_{t-1} \\ (b_{1} + d_{1}e^{-[Y_{t-1}^{2} + Y_{t-2}^{2}]})Y_{t-1} + \sigma_{2}e_{t}, & \text{if } Y_{t-2} < \frac{1}{b_{1}}Y_{t-1}. \end{cases}$$

This has clear representations in the forms $\{X'_t\}$ as in (3) and $\{X_t\}$ as in (2), with $X_t = (Y_t, Y_{t-1})'$. Suppose $b_1 < 0$, $a_1, a_2 > 0$, $a_1b_1 + a_2 < 0$ and $b_1^2 > a_1b_1 + a_2$. Example 2 showed $\{X_t\}$ satisfies the assumptions of Theorem 1. Thus, so does (3). The conditions for stability follow from those found in Example 2, that is, $a_1 + a_2 < 1$ and no restrictions on $|b_1|$ if $b_1 < 0$. Conversely, $\{X_t\}$ is transient if $a_1 + a_2 > 1$ since $||x||^r P_x(X_1 \notin Q) \to 0$ by the argument previously mentioned.

3. Applications

The full power of the finite state chain piggyback as compared to the skeleton piggyback becomes obvious in cases where the finite state Markov chain piggyback will not have a trivial stationary distribution. As previously mentioned, partitioning the state-space so that (A2) is satisfied may require some cleverness.

Example 3. Consider

$$Y_{t} = \begin{cases} a_{0} + a_{1}Y_{t-1} + a_{2}Y_{t-2} + \sigma_{1}e_{t} & \text{if } Y_{t-1} \leq c_{1}Y_{t-2}, \ Y_{t-2} > 0, \\ b_{0} + b_{1}Y_{t-1} + b_{2}Y_{t-2} + \sigma_{2}e_{t} & \text{if } Y_{t-1} > c_{1}|Y_{t-2}|, \\ c_{0} + c_{1}|Y_{t-1}| + \sigma_{3}e_{t} & \text{if } Y_{t-1} \leq -c_{1}Y_{t-2}, \ Y_{t-2} \leq 0. \end{cases}$$

The state-space for $X_t = (Y_t, Y_{t-1})$ is obviously \mathbf{R}^2 . The relevant thresholds are both axes and the rays containing $(c_1, -1)$, $(c_1, 1)$ and $(-\frac{b_2}{b_1}, 1)$, resulting in seven basic regions.

We consider the subset of the parameter space for which

$$a_1 > 0$$
, $b_1 < 0$, $c_1 > 0$, $a_2 < -a_1c_1$, $-b_1c_1 < b_2 < c_1^2 - b_1c_1$

Note $a_2 < 0$ and $b_2 > 0$. Then

$$w_{1} < 0, \ w_{2} > 0 \Rightarrow a_{1}w_{1} + a_{2}w_{2} < 0,$$

$$0 < w_{1} < c_{1}w_{2}, \ w_{2} > 0 \Rightarrow a_{1}w_{1} + a_{2}w_{2} < (a_{1}c_{1} + a_{2})w_{2} < 0,$$

$$w_{1} > c_{1}|w_{2}|, \ w_{1} > -\frac{b_{2}}{b_{1}}w_{2} \Rightarrow b_{1}w_{1} + b_{2}w_{2} < 0,$$

$$c_{1}w_{2} < w_{1} < -\frac{b_{2}}{b_{1}}w_{2}, \ w_{2} > 0 \Rightarrow 0 < b_{1}w_{1} + b_{2}w_{2} < \left(b_{1} + \frac{b_{2}}{c_{1}}\right)w_{1} < c_{1}w_{1}$$

From this it may be seen that the simpler version of Theorem 1 alluded to after Example 2 applies to the seven regions mentioned above.

Here, however, we will apply Theorem 1 in its full generality by refining the regions optimally. With high probability (as $||x|| \to \infty$) the third quadrant in \mathbb{R}^2 is reached with four steps or fewer from anywhere. From the third quadrant, the process (when large) leads immediately to a narrow cone containing the ray $(c_1, -1)$. Depending on which side of that ray the process falls (which is a consequence of the random error), it will successively hit a sequence of other narrow cones until it returns to the third quadrant once again.

We therefore set up a scheme of regions as follows. Let $n > c_1^{-1}$ be arbitrarily large. Define

$$R_{n,1} = \left\{ (w_1, w_2) : 0 < \left(-c_1 + \frac{1}{n} \right) w_2 < w_1 < -c_1 w_2 \right\}$$

and

$$R_{n,2} = \left\{ (w_1, w_2) : 0 < -c_1 w_2 \le w_1 < \left(-c_1 - \frac{1}{n} \right) w_2 \right\}.$$

Note that points in the third quadrant map into $R_{n,1} \cup R_{n,2}$ with probability approaching 1, uniformly as $||x|| \to \infty$. Now define additional regions

$$\begin{split} R_{n,3} &= \left\{ (w_1, w_2) : 0 < c_1 w_2 \le w_1 < \left(c_1 + \frac{1}{n}\right) w_2 \right\}, \\ R_{n,4} &= \left\{ (w_1, w_2) : 0 < \left(c_1 - \frac{1}{n}\right) w_2 \le w_1 < c_1 w_2 \right\}, \\ R_{n,5} &= \left\{ (b_1 w_1 + b_2 w_2, w_1) : (w_1, w_2) \in R_{n,3} \right\}, \\ R_{n,6} &= \left\{ (b_1 w_1 + b_2 w_2, w_1) : (w_1, w_2) \in R_{n,2} \right\}, \\ R_{n,7} &= \left\{ (a_1 w_1 + a_2 w_2, w_1) : (w_1, w_2) \in R_{n,4} \right\}, \\ R_{n,8} &= \left\{ (a_1 w_1 + a_2 w_2, w_1) : (w_1, w_2) \in R_{n,5} \right\}, \\ R_{n,9} &= \left\{ (a_1 w_1 + a_2 w_2, w_1) : (w_1, w_2) \in R_{n,6} \right\}, \\ R_{n,10} &= \left\{ (a_1 w_1 + a_2 w_2, w_1) : (w_1, w_2) \in R_{n,7} \right\}, \\ R_{n,11} &= \left\{ (a_1 w_1 + a_2 w_2, w_1) : (w_1, w_2) \in R_{n,8} \right\}. \end{split}$$

See Fig. 2. The corresponding companion matrices are

$$A_{1} = \begin{pmatrix} c_{1} & 0\\ 1 & 0 \end{pmatrix},$$

$$A_{2} = A_{3} = \begin{pmatrix} b_{1} & b_{2}\\ 1 & 0 \end{pmatrix},$$

$$A_{4} = A_{5} = A_{6} = A_{7} = A_{8} = \begin{pmatrix} a_{1} & a_{2}\\ 1 & 0 \end{pmatrix},$$

$$A_{9} = A_{10} = A_{11} = \begin{pmatrix} -c_{1} & 0\\ 1 & 0 \end{pmatrix}.$$

Everything else is designated $R_{n,0}$, which is transient (and has varying companion matrices).

Let $\gamma = P(c_0 + \sigma_3 e_1 < 0)$. This gives us 11 recurrent regions (assuming *n* is sufficiently large) with the dynamics

$$\begin{split} R_{n,1} &\to \begin{cases} R_{n,3} & \text{w.p. } 1 - \gamma, \\ R_{n,4} & \text{w.p. } \gamma, \end{cases} \\ R_{n,2} &\to R_{n,6} \to R_{n,9}, \\ R_{n,3} \to R_{n,5} \to R_{n,8} \to R_{n,11}, \\ R_{n,4} \to R_{n,7} \to R_{n,10}, \end{cases} \\ R_{n,j} &\to \begin{cases} R_{n,1} & \text{w.p. } \gamma, \\ R_{n,2} & \text{w.p. } 1 - \gamma, \end{cases} j = 9, 10, 11. \end{split}$$



Fig. 2. Regions of the state-space in Example 3.

The path of the finite state chain consists of iid excursions leaving either $R_{n,1}$ or $R_{n,2}$, namely

$$\begin{aligned} R_{n,1} &\to R_{n,3} \to R_{n,5} \to R_{n,8} \to R_{n,11} \quad \text{w.p. } \gamma(1-\gamma), \\ R_{n,1} &\to R_{n,4} \to R_{n,7} \to R_{n,10} \quad \text{w.p. } \gamma^2, \\ R_{n,2} \to R_{n,6} \to R_{n,9} \quad \text{w.p. } 1-\gamma. \end{aligned}$$

Each of these occurs independently of previous excursions, and is followed by another. The expected excursion length is $5\gamma(1-\gamma)+4\gamma^2+3(1-\gamma)=(1+\gamma)(3-\gamma)$. By considering the expected time spent in each state during an excursion, we find that the stationary distribution is

$$\frac{\left[\gamma \ 1-\gamma \ \gamma(1-\gamma) \ \gamma^2 \ \gamma(1-\gamma) \ 1-\gamma \ \gamma^2 \ \gamma(1-\gamma) \ 1-\gamma \ \gamma^2 \ \gamma(1-\gamma)\right]}{(1+\gamma)(3-\gamma)}$$

As $n \to \infty$, $R_{n,1}$ and $R_{n,2}$ each shrink to the ray containing $(c_1, -1)$. Thus, each excursion may be associated with a limiting growth factor which is the product of growth factors for its steps. Letting

$$\theta = \left(\frac{c_1}{\sqrt{1+c_1^2}}, \frac{-1}{\sqrt{1+c_1^2}}\right)$$

and optimizing with $n \to \infty$, we find that the critical constant is

$$\rho = ((\|A_{11}A_8A_5A_3A_1\theta\|)^{\gamma(1-\gamma)}(\|A_{10}A_7A_4A_1\theta\|)^{1-\gamma}(\|A_9A_4A_2\theta\|)^{\gamma^2})^{1/(1+\gamma)(3-\gamma)}$$

= $((\|A_9A_4^2A_2A_1\theta\|)^{\gamma(1-\gamma)}(\|A_9A_4^2A_1\theta\|)^{1-\gamma}(\|A_9A_4A_2\theta\|)^{\gamma^2})^{1/(1+\gamma)(3-\gamma)}.$

That is, if $\rho < 1$ then $\{X_t\}$ is ergodic while since $||x||^r P_x(X_1 \notin \bigcup_{i=1}^{11} R_{n,i}) \to 0$, if $\rho > 1$ then $\{X_t\}$ is transient.

Classification is much more subtle in case $\rho = 1$. In this case X_t can be ergodic, null recurrent or transient. Precise classification will depend on the intercepts a_0, b_0, c_0 and the variances $\sigma_1^2, \sigma_2^2, \sigma_3^2$.

Example 4. Consider the TAR(3) model

$$Z_{t} = \begin{cases} a_{0} + a_{1}Y_{t-1} + a_{2}Y_{t-2} + \sigma_{1}e_{t} & \text{if } Y_{t-1} \ge a_{1}Y_{t-2} + a_{2}Y_{t-3}, \\ b_{0} + b_{1}Y_{t-1} + b_{2}Y_{t-2} + b_{3}Y_{t-3} + \sigma_{2}e_{t} & \text{if } Y_{t-1} < a_{1}Y_{t-2} + a_{2}Y_{t-3}, \end{cases}$$
$$Y_{t} = |Z_{t}|.$$

The state vector is $X_t = (Y_t, Y_{t-1}, Y_{t-2})$ and the state space is $(0, \infty)^3$. Suppose

$$a_1 > 0$$
, $a_2 > 0$, $b_1a_1 + b_2 > a_1^2 + a_2$, $b_1a_2 + b_3 > a_1a_2$.

The two regions are

$$R_1 = \{ (w_1, w_2, w_3) : w_1 \ge a_1 w_2 + a_2 w_3, w_2 > 0, w_3 > 0 \},\$$

$$R_2 = \{ (w_1, w_2, w_3) : 0 < w_1 < a_1 w_2 + a_2 w_3, w_2 > 0, w_3 > 0 \}.$$

Since a_1 and a_2 are positive and, for positive w_2, w_3 ,

$$b_1(a_1w_2 + a_2w_3) + b_2w_2 + b_3w_3 > a_1(a_1w_2 + a_2w_3) + a_2w_2,$$

we see that $Y_t = Z_t$ with high probability if the process is large. So, for the sake of studying stability, the dynamics will assume Y_t and Z_t are the same. (What follows is applicable even if we had defined $Y_t = Z_t$, but it would cover only one part of the dynamics needed to study stability.)

The companion matrices are

$$A_1 = \begin{pmatrix} a_1 & a_2 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} b_1 & b_2 & b_3\\ 1 & 0 & 0\\ 0 & 1 & 0 \end{pmatrix}.$$

Let $\lambda_1 = \frac{a_1 + \sqrt{a_1^2 + 4a_2}}{2}$ which is the positive (and maximal) eigenvalue of A_1 and is associated with eigenvector $\theta_1 = \begin{pmatrix} \lambda_1 \\ 1 \\ 1/\lambda_1 \end{pmatrix}$. We make the additional assumption that θ_1 is not an eigenvector for A_2 .

Note that (when ||x|| is large) points in R_1 map to a narrow cone containing the semi-plane U defined by

$$U = \{ (w_1, w_2, w_3) : w_1 = a_1 w_2 + a_2 w_3, w_2 > 0, w_3 > 0 \}.$$

Anything in R_2 maps back into R_1 , due to the parameter assumptions, Moreover, R_1 maps into R_2 with probability $\gamma = P(a_0 + \sigma_1 e_1 < 0)$ and into R_1 with probability $1 - \gamma$. We now define refined regions

$$R_{n,1} = \left\{ (w_1, w_2, w_3) : 0 \le w_1 - a_1 w_2 + a_2 w_3 < \frac{1}{n} (w_2 + w_3), w_2 > 0, w_3 > 0 \right\},$$

$$R_{n,2} = \left\{ (w_1, w_2, w_3) : -\frac{1}{n} (w_2 + w_3) < w_1 - a_1 w_2 + a_2 w_3 < 0, w_2 > 0, w_3 > 0 \right\},$$

$$R_{n,3} = \left\{ (b_1 w_1 + b_2 w_2, b_3 w_3, w_1, w_2) : (w_1, w_2, w_3) \in R_{n,2} \right\}.$$

Asymptotically, as $||x|| \to \infty$, we thus have the dynamics

$$\begin{split} R_{n,1} &\to \begin{cases} R_{n,1} & \text{w.p. } 1 - \gamma, \\ R_{n,2} & \text{w.p. } \gamma, \end{cases} \\ R_{n,2} &\to R_{n,3}, \end{cases} \\ R_{n,3} &\to \begin{cases} R_{n,1} & \text{w.p. } 1 - \gamma, \\ R_{n,2} & \text{w.p. } \gamma. \end{cases} \end{split}$$

The stationary distribution for the 3-state chain is $\frac{1}{1+\gamma}(1-\gamma,\gamma,\gamma)$. Now observe that

$$\lambda_1 = \sup_{x \in U} \frac{\|A_1x\|}{\|x\|}$$
 and $\lambda_2 = \sup_{x \in U} \frac{\|A_1A_2x\|}{\|x\|}$,

are eigenvalues of A_1 and A_1A_2 , respectively (since both A_1 and A_1A_2 have eigenvectors in U). Applying Theorem 1, it easily follows that

$$\rho_{i,S} = (\lambda_1^{1-\gamma} \lambda_2^{\gamma})^{1/(1+\gamma)}$$

and a sufficient condition for ergodicity is $\rho_{i,S} < 1$.

By assumption, $\lambda_1 \neq \lambda_2$. Therefore, we conjecture that the condition above is not sharp. It is, however, apparent that a deeper analysis would be required to improve it substantially. At best one must consider the random process on all of Uwith the dynamics

$$x \to \begin{cases} A_1 x & \text{w.p. } 1 - \gamma, \\ A_1 A_2 x & \text{w.p. } \gamma. \end{cases}$$

4. Discussion

In cases where some of the regions that do not allow approximation by a finite state chain as in (7) are "recurrent" in the sense that (8) fails, the finite state chain approximation is of no use. This could happen when an eigenvector of a companion matrix lies near to and parallel to a threshold, for example, or when the action of a companion matrix in a "recurrent" region rotates points towards a threshold.

The TAR and TAR-like processes benefit from the fact that the errors become inconsequential as the process grows large. Some threshold processes do not possess these properties, TAR-GARCH processes being an example. The results in this paper are not general enough to handle these cases.

5. Proofs

We direct the interested reader to Meyn and Tweedie [4] for details and definitions. Considering the state-space/general state Markov chain representation of the time series allows the use of well-known techniques for determining stability of the process (see [4]). One common technique is to demonstrate that the process satisfies a drift criterion for stability. This usually entails constructing a test function designed expressly for the purpose.

These drift criteria all depend upon the Markov chain being ψ -irreducible and aperiodic. The TAR process leads to a ψ -irreducible, aperiodic Markov chain under reasonable conditions, such as if the error distribution possesses an everywhere positive density which is lower semi-continuous [9]. In this case the maximal irreducibility measure ψ is absolutely continuous with respect to Lebesgue measure, and compact sets are petite.

There are different drift criteria for the different forms of Markov chain stability. From Lemma 2(i) in [1], $\{X_t\}$ is V-uniformly ergodic if $\{X_t\}$ is a ψ -irreducible, aperiodic general state Markov chain, $V \ge 1$ is an unbounded, locally bounded and measurable test function, if for some integer $k \ge 1$ and all $M < \infty$

$$\limsup_{V(x)\to\infty} \frac{E_x[V(X_k)]}{V(x)} < 1, \quad \sup_{V(x)\le M} E_x[V(X_k)] < \infty, \quad \sup_x \frac{E_x[V(X_1)]}{V(x)} < \infty, \quad (13)$$

and if the sublevel sets $C_M^{(V)} := \{x : V(x) \leq M\}$ are petite. Under certain conditions, pre-compact sets are petite, so if the topology of the state-space is determined by the norm $\|\cdot\|$ and $V \to \infty$ as $\|x\| \to \infty$, then the sublevel sets $\{x : V(x) \leq c\}$ are petite.

Similarly, demonstrating instability of Markov chains is equivalent to constructing a test function of the chain which satisfies a drift criterion for transience. From Lemma 2(ii) in [1], $\{X_t\}$ is transient if $\{X_t\}$ is a ψ -irreducible, aperiodic general state Markov chain, $V \ge 0$ is an unbounded function with $\psi(\{x : V(x) > M\}) > 0$ for all $M < \infty$ and if for some positive integer k

$$\limsup_{V(x)\to\infty} E_x \left[\frac{V(x)}{V(X_k)} \right] < 1.$$
(14)

To prove V-uniform ergodicity or transience we construct a test function which satisfies the appropriate drift criterion (13) or (14). In the interest of readability, we first present two lemmas. This first lemma demonstrates the expected log change in $\{X_t\}$ when piggybacked upon a stationary distribution of $\{J_t\}$ is negative in the case of stability and positive in the case of transience, demonstrates that convergence of the expected log change, conditioned upon an arbitrary initial value x, to these expected values occurs in a finite time. This is used to derive a function h_n which tracks the log-drift of $\{X_t\}$ among regions when moving according to the stationary distribution of $\{J_t\}$ and which satisfies conditions similar to the drift conditions for stability and transience. The test functions for $\{X_t\}$ will then be piggybacked upon h_n . Recall the definitions of $\rho_{i,I}$, $\rho_{i,S}$ in Eq. (10). **Lemma 1.** Suppose (A1) holds. Define $h(i, z) := \log(\rho_{i,z})$ and $h_n(i, z) := \sum_{t=0}^{n-1} \frac{n-t}{n} E_i[h(J_t, z)]$, for z = I, S.

(i) If (11) in Theorem 1 holds, then there exists $n_* < \infty$ such that for $n \ge n_*$

$$\max_{1 \le i \le l} E_i[h_n(J_1, S)] + h(i, S) - h_n(i, S) < 0.$$
(15)

(ii) If (12) in Theorem 1 holds, then there exists $n_* < \infty$ such that for $n \ge n_*$

$$\max_{i \in S_{u_*}} h_n(i, I) - E_i[h_n(J_1, I)] - h(i, I) < 0.$$
(16)

Proof. By item,

(i) The assumptions imply $\pi^{(u)}h(S) := \sum_j \pi_j^{(u)}h(j,S) < 0$ for each $u \in \{1,\ldots,k\}$. If $i \in S_u$ for some $u \in \{1,\ldots,k\}$, then $E_i[h(J_t,S)] \to \pi^{(u)}h(S)$. Since the number of states is finite, the transient states T are uniformly transient, and since h is bounded, there exists $n_* < \infty$ so that $n \ge n_*$ implies $\frac{1}{n} \sum_{t=1}^n E_i[h(J_t,S)] < 0$, for all i in $\{1,\ldots,l\}$, and thus that

$$E_{i}[h_{n}(J_{1},S)] + h(i,S) - h_{n}(i,S)$$

$$= E_{i}\left(\sum_{t=0}^{n} \frac{n-t+1}{n} E_{J_{1}}[h(J_{t},S)] - \sum_{t=0}^{n-1} \frac{n-t}{n} E_{i}[h(J_{t},S)]\right)$$

$$= \frac{1}{n} \sum_{t=0}^{n} E_{i}[h(J_{t},S)] < 0.$$

(ii) Suppose (12) in Theorem 1 holds for $u = u_*$. By reasoning similar to the proof of item (i) above, with (12) implying $\pi^{(u_*)}h(I) := \sum_j \pi_j^{(u_*)}h(j,I) > 0$, then (16) follows.

This next lemma introduces a counterpart h'_n to the function h_n introduced in Lemma 1, a function that tracks the log-drift among regions for $\{X_t\}$. It is demonstrated that the expectation of h'_n will be arbitrarily close to the expectation of the function h_n , averaged over a sufficiently long time, when the process $\{X_t\}$ is large. Let $h'(x, z) = \sum_{j=1}^{l} h(j, z) I_{x \in R_j}$, and let $h'_n(x, z) = \sum_{t=0}^{n-1} \frac{n-t}{n} E_x[h'(X_t, z)]$, for $n < \infty$.

Lemma 2. Suppose (A1), (A2) and either (11) or (12) hold. Then for arbitrary $\gamma > 0$, there exists $n = n(\gamma)$ with

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} |E_x[h'_n(X_1, z) - h'_n(x, z)] - E_i[h_n(J_1, z) - h_n(i, z)]| < \gamma,$$

$$z \in \{I, S\}. \quad (17)$$

Proof. Let $N = \max_j |h(j, S)|$. Given $\gamma > 0$ pick $\epsilon > 0$ so that $\epsilon < \gamma/4N$. The transient states T being uniformly transient implies there exists a $t^{**} < \infty$ with

 $P_i(J_t \in T) < \epsilon$ for $i \in T$ and $t \ge t^{**}$. Using t^* from (8) set $t' = \max(t^*, t^{**})$. Get n_* from Lemma 1 and pick $n \ge \max(n_*, t')$ so that $2N((t'-1)/n) + 2\epsilon) < \gamma$.

Applying the Markov property, and since $x \in R_i$ implies h'(x, z) = h(i, z)

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} |E_x[h'_n(X_1, z) - h'_n(x, z)] - E_i[h_n(J_1, z) - h_n(i, z)]|$$

$$\le \frac{N}{n} \sum_{t=1}^n \max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} |P_x(X_t \in R_j) - P_i(J_t = j)|.$$
(18)

Now for $t \ge t'$, by (8)

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} \max_{j \in T} |P_x(X_t \in R_j) - P_i(J_t = j)| < 2\epsilon.$$
(19)

Let $B_G(t, M) = \{ [X_t \in G] \cap [||X_t|| > M] \}$. Then from (7) given $\epsilon > 0$ there exists $M < \infty$ so that

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} \max_{j \in G} |P_x(X_t \in R_j) - P_i(J_t = j)|$$

$$\le \epsilon + \max_{\substack{1 \le i \le l \\ \|x\| \to \infty}} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} P_x([B_G(t - 1, M)]^C).$$
(20)

Since the number of regions/states is finite, suppose w.l.o.g. that there exists $\eta > 0$ so that $||A_i||^S > \eta$ for $i \in \{1, \ldots, l\}$. It then follows that $\limsup_{\|x\|\to\infty} P_x(\|X_t\| \le M) = 0$. Then from this and (8), for t > t'

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} P_x([B_G(t-1,M)]^C) < \epsilon.$$
(21)

The conclusion follows from (18)–(21), and $1/n \sum_{t=1}^{t'-1} |E_x[h'_n(X_t, z)] - E_i[h_n(J_t, z)]| \le 2N(t'-1)/n.$

What remains is to use the knowledge that the functions h'_n, h_n are close in expectation to build the test functions for $\{X_t\}$ which satisfy the drift criteria for ergodicity and transience.

Proof of Theorem 1. For an integer n and scalar s define $H_n(s, x, z) := \exp[sh'_n(x, z)]$. By item,

(i) Since the number of states is finite, from (15) there exists $\gamma > 0$ so that $E_i[h_n(J_1, S)] + h(i, S) - h_n(i, S) < -\gamma$ for each *i*, for $n \ge n_*$. Combining this with (17) implies $E_x[h'_n(X_1, S)] - h'_n(x, S) + h(i, S) < 0$ for each $x \in R_i$ and each *i*, and $n \ge \max(n(\gamma), n_*)$. Now apply the fact that for y > 0, $\frac{y^s - 1}{s} \to \log(y)$, as

 $s \to 0$. This limit is only locally uniform in y, but since the random variables are bounded, for $n \ge \max(n(\gamma), n_*)$, for s small

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} E_x \left(\frac{H_n(s, X_1, S)(\rho_{i,S})^s}{H_n(s, x, S)} \right) < 1.$$
(22)

Since $E|e_t|^r < \infty$ and $\limsup_{\|x\|\to\infty} \|g(x)\|/\|x\| = 0$, then for $s < \min(r/2, 1)$

$$\max_{1 \le i \le l} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} E_x \left(\frac{\|X_1\|^s}{(\rho_{i,S})^s \|x\|^s} \right) = 1.$$
(23)

For these choices of $n \ge \max(n(\gamma), n_*)$ and $s < \min(r/2, 1)$ small enough so that (22) holds, let $V(x) = 1 + [H_n(s, x, S) ||x||^s]^{1/2}$. Then $V \ge 1$ is norm-like, measurable, bounded on compact sets and thus locally bounded, and by (22), (23) and Cauchy–Schwarz,

$$\limsup_{V(x)\to\infty} E_x\left(\frac{V(X_1)}{V(x)}\right) < 1.$$

Since $E|e_t|^r < \infty$ and each of the $||A_i||$ are finite, then $E_x[V(X_1)]$ is bounded for each $M < \infty$ on $\{x : V(x) \le M\}$ and also $\sup_{x \ne 0} E_x[V(X_1)]/V(x) < \infty$, while $E|e_t|^r < \infty$ implies $E_{x=0}[V(X_1)]/V(0) < \infty$. Thus $\{X_t\}$ is V-uniformly ergodic by (13) with k = 1.

(ii) Suppose (12) holds for $u = u_*$. Arguments like those leading to (22) will show by (16) from Lemma 1 that for small s, for $n \ge n_*$ there exists $\epsilon_1 > 0$ so that

$$\max_{i \in S_{u_*}} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} E_x \left(\frac{H_n(s, x, I)}{H_n(s, X_1, I)(\rho_{i,I})^s} \right) < 1 - \epsilon_1.$$
(24)

Pick $\epsilon_2 > 0$ so that $\max_{i \in S_{u_*}} (1 - \epsilon_1) [\rho_{i,I} / (\rho_{i,I} - \epsilon_2)]^s < 1$. Let $C = \max_{i \in S_u} \sigma_i$. Now $||g(x)|| + C|e_1| \le \epsilon_2 ||x||$ implies $||X_1|| \ge (\rho_{i,I} - \epsilon_2) ||x||$. Since H_n is bounded away from zero, choose m_1, m_2 such that $m_1 \le H_n(s, x, I)^{-1} \le m_2$ for all s, x. Also assume ||x|| is so large that $||g(x)|| / ||x|| \le \epsilon_2/2$. Then

$$\frac{H_n(s,x,I)^{-1} + \|x\|^s}{H_n(s,X_1,I)^{-1} + \|X_1\|^s} \le \frac{m_2 + \|x\|^s}{m_1 + (\rho_{i,I} - \epsilon_2)^s \|x\|^s} I_{\{\|g(x)\| + C|e_1| \le \epsilon_2 \|x\|\}} + \frac{m_2 + (2C|e_1|/\epsilon_2)^s}{m_1} I_{\{\|g(x)\| + C|e_1| > \epsilon_2 \|x\|\}},$$

so that since $E|e_1|^s < \infty$ for s < r, by Markov's inequality

$$\max_{i \in S_{u_*}} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} E_x \left(\frac{(H_n(s, x, I)^{-1} + \|x\|^s)(\rho_{i,I})^s}{H_n(s, X_1, I)^{-1} + \|X_1\|^s} \right) \le \left(\frac{\rho_{i,I}}{\rho_{i,I} - \epsilon_2} \right)^s.$$
(25)

Recall $Q := \bigcup_{i \in S_{u_*}} R_i$ and observe that since H_n is bounded, for any x

$$\frac{H_n(s, x, I)^{-1} + \|x\|^s}{H_n(s, X_1, I)^{-1} + \|X_1\|^s I_{X_1 \in Q}} \leq m_1^{-1} [H_n(s, x, I)^{-1} + \|x\|^s] I_{X_1 \notin Q} + \frac{H_n(s, x, I)^{-1} + \|x\|^s}{[H_n(s, X_1, I)^{-1} + \|X_1\|^s]} I_{X_1 \in Q}.$$
(26)

For the chosen n and s < r let $V(x) = (1 + ||x||^s H_n(s, x, I) I_{x \in Q})^{1/2}$. Then by (24)–(26), Cauchy–Schwarz, and since by assumption $\limsup_{\substack{\|x\|\to\infty\\x\in Q}} x$

$$\|x\|^r P_x(X_1 \notin Q) = 0,$$

$$\lim_{V(x) \to \infty} E_x\left(\frac{V(x)}{V(X_1)}\right) = \max_{i \in S_{u_*}} \limsup_{\substack{x \in R_i \\ \|x\| \to \infty}} E_x\left(\frac{V(x)}{V(X_1)}\right) < 1$$
(27)

so that (14) is satisfied with k = 1 and therefore $\{X_t\}$ is transient.

References

- T. R. Boucher and D. B. H. Cline, Stability of cyclic threshold autoregressive time series models, *Statist Sinica* 17 (2007) 43–62.
- K. S. Chan and H. Tong, On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations, Adv. Appl. Probab. 17 (1985) 666–678.
- D. B. H. Cline and H. Pu, Stability and the Lyapunov exponent of threshold AR-ARCH models, Ann. Appl. Probab. 14 (2004) 1920–1949.
- 4. S. P. Meyn and R. L. Tweedie, Markov Chains and Stochastic Stability (1993).
- J. D. Petruccelli and S. W. Woolford, A threshold AR(1) model, J. Appl. Probab. 21 (1984) 270–286.
- H. Tong, Threshold Models in Nonlinear Time Series Analysis, Lecture Notes in Statistics, Vol. 21 (Springer-Verlag, 1983).
- H. Tong, Non-Linear Time Series: A Dynamical System Approach (Oxford Univ. Press, 1990).
- H. Tong and K. S. Lim, Threshold autoregression, limit cycles and cyclical data (with discussion), J. Roy. Statist. Soc. (Ser. B) 42 (1980) 245–292.
- D. B. H. Cline and H. Pu, Verifying irreducibility and continuity of a nonlinear time series, *Stat. Probab. Lett.* 40 (1998) 139–148.