Stability of Cyclic Threshold and Threshold-like Autoregressive Time Series Models

Thomas R. Boucher, Daren B.H. Cline

Plymouth State University, Texas A&M University

Abstract: We investigate the stability, in terms of V-uniform ergodicity or transience, of *cyclic* threshold autoregressive time series models. These models cycle through one of a number of collections of subregions of the state space when the process is large. Our results can be applied in cases where the model has multiple cycles and/or affine thresholds. The bounds on the parameter space are sharper than those in previous results, and are easily verified. We extend these results to autoregressive nonlinear time series that can be approximated well by a threshold model (*threshold-like*).

Key words and phrases: ergodicity; Markov chain; nonlinear autoregressive time series; nonlinear time series; threshold autoregressive time series.

1. Introduction

1.1 Background

The threshold autoregressive model of order p, delay $d \le p$, and s regimes (TAR$(p; d; s)$) is the piecewise linear autoregression

$$Y_t = \begin{cases} \phi_1^{(1)} Y_{t-1} + \cdots + \phi_{p_1}^{(1)} Y_{t-p_1} + \xi_t, & & Y_{t-d} \le r_1 \\ \phi_1^{(i)} Y_{t-1} + \cdots + \phi_{p_i}^{(i)} Y_{t-p_i} + \xi_t, & i = 2, \ldots, s-1 & r_{i-1} < Y_{t-d} \le r_i, \\ \phi_1^{(s)} Y_{t-1} + \cdots + \phi_{p_s}^{(s)} Y_{t-p_s} + \xi_t, & & Y_{t-d} > r_{s-1} \end{cases} \tag{1.1}$$

where $\phi_1^{(i)}, \ldots, \phi_{p_i}^{(i)}$ are the autoregression coefficients, $p_i \le p$, $i = 1, \ldots, s$, the $\{\xi_t\}$ are mean zero i.i.d. random variables, and the constants r_i, $i = 1, \ldots, s-1$, are called the *thresholds* of the process. We consider more general thresholds of the form $a_i z_{t-d} + b_i = 0$, where $z_{t-d} = (y_{t-1}, \ldots, y_{t-d})'$, and a_i, b_i are vectors in $\mathbb{R}^p$, $i = 1, \ldots, s-1$. These generalized thresholds allow for a richer variety of behaviors.

Stability refers to the set of parameter values $\{\phi_j^{(i)}\}$, $j = 1, \ldots, p_i$, $i = 1, \ldots, s$ that allow $\{Y_t\}$ to reach a stationary distribution. There exist stability conditions explicitly for, or that can be applied to, TAR$(p; d; s)$ models. However, much of the existing research applies to specific models and is not generally applicable. Other results are general, but are either difficult to verify or are so strong that they unnecessarily restrict the parameter space. We attempt to remedy these defects by providing easily verified conditions that are applicable to a wide class of models, and that allow the recovery of more, if not all, of the stable parameter space.

1.2 Markov chain embedding

The TAR$(p; d; s)$ model from (1.1) can be embedded in the following general state Markov chain on $\mathbb{R}^p$:

$$X_t = \sum_{i=1}^{l} A_i I(X_{t-1} \in R_i) X_{t-1} + \nu_t, \tag{1.2}$$

with $X_t = (Y_t, Y_{t-1}, \ldots, Y_{t-p+1})'$, $\nu_t = \xi_t(1, 0, \ldots, 0)'$, and the space $\mathbb{R}^p$ is partitioned into l regions $R_i, i = 1, \ldots, l$, with the R_i determined by, but not always corresponding exactly to, the $s-1$ thresholds $a_i Z_{t-d} + b_i = 0$, $i = 1, \ldots, s-1$. The A_i are called the *companion matrices* and are given by

$$A_i = \begin{pmatrix} \phi_1^{(i)} & \phi_2^{(i)} & \ldots & \phi_{p_i-1}^{(i)} & \phi_{p_i}^{(i)} \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & \vdots & \ldots & 1 & 0 \end{pmatrix}, \quad i = 1, \ldots, l.$$

Since the regions do not always correspond to the thresholds, a particular companion matrix may apply to more than one region.

Embedding the time series in the Markov chain allows one to take advantage of existing Markov chain stability results. These results are often expressed as Foster-Lyapunov drift criteria. Authors such as Cline and Pu (2001a), Meyn and Tweedie (1993), and Tjøstheim (1990), among others, have followed this strategy.

1.3 Cyclic Models

Cyclic threshold autoregressive time series models exhibit asymptotic cyclic behavior, i.e., the Markov chain $\{X_t\}$ tends to cycle through one of a number of collections of subregions of the state space when the process is large. Tong and Lim (1980) originally introduced threshold autoregressive models to account for, among other things, cyclic phenomena in time series data.

Define the *deterministic skeleton* $\{x_t\}$ of the Markov chain $\{X_t\}$ to be the process with the additive errors removed, i.e.,

$$x_t = \sum_{i=1}^{l} A_i I(x_{t-1} \in R_i) x_{t-1}. \tag{1.3}$$

A k*-cycle* for the deterministic skeleton is a collection $\mathbb{C} = \{R^{(1)}, \ldots, R^{(k)}\}$ of k regions with corresponding companion matrices $\{A^{(1)}, \ldots, A^{(k)}\}$, so that $x \in R^{(i)}$ implies $A^{(i)}x \in R^{(i+1) \mod (k)}$. The general cyclic case has $1 \leq m < \infty$ cycles $\mathbb{C}_1, \ldots, \mathbb{C}_m$, each of length k_i, $i = 1, \ldots, m$. A cyclic TAR$(p; d; s)$ occurs where the skeleton has one or more cycles, all points are mapped by the skeleton to a cycle by a uniformly finite time, and the behavior of $\{X_t\}$ mirrors that of $\{x_t\}$ when $\{X_t\}$ is large. This entails certain assumptions on the behavior of the skeleton and the error distribution. We formalize this in our assumptions in Section 3. Tjøstheim (1990) dealt briefly with the case of a single-cycle cyclic TAR(p, d, s) model, and we have borrowed and modified some of his notation.

In dealing with stability of Markov chains, Meyn and Tweedie (1993), Nummelin (1984), and Tjøstheim (1990) each employed a k-step method which takes advantage of the fact that stability of the one-step chain $\{X_t\}$ and the k-step chain $\{X_{tk}\}$ are equivalent for a finite positive integer k, though this equivalence has not been demonstrated when V-uniform ergodicity is the criteria. Stability conditions for $\{X_t\}$ are then determined through analysis of $\{X_{tk}\}$. Cyclic TAR(p, d, s) models are particularly amenable to this approach with $k = \prod k_i$; these models need not shrink in expectation at every transition, just at each pass through the cycle.

In summary, we develop stability criteria for cyclic TAR(p, d, s) models by embedding the time series $\{Y_t\}$ in the Markov chain $\{X_t\}$. Due to cyclicity we consider the k-step chain $\{X_{tk}\}$. Stability of $\{X_{tk}\}$ is defined as V-uniform ergodicity; stability conditions for $\{X_{tk}\}$ are based on the stability of its skeleton $\{x_{tk}\}$ and stability of the error distribution. The sufficiency of these conditions is demonstrated through the use of Foster-Lyapunov drift criteria. We then employ an original result which equates the V-uniform ergodicity of $\{X_{tk}\}$ with that of $\{X_t\}$. Stability of $\{Y_t\}$ follows.

The paper is organized as follows: lemmas containing the drift conditions for V-uniform ergodicity and transience, and establishing the equivalence of the V-uniform ergodicity of $\{X_t\}$ and $\{X_{tk}\}$, are in Section 2. Section 3 contains the stability results, some examples, and an extension of the stability results to more general nonlinear autoregressive time series. A brief discussion is in Section 4, and proofs are in Section 5.

2. General State Markov Chains

We provide some definitions here, referring the reader to Meyn and Tweedie (1993) for these and additional explanation. In the following we assume $\{X_t\}$ is a Markov chain.

2.1 V-uniform ergodicity.

Following Meyn and Tweedie (1993), define for a function $V \geq 1$ the *V-norm distance* between two kernels P_1 and P_2 as

$$|||\, P_1 - P_2 \,|||_V := \sup_x \sup_{|g| \leq V} \frac{|P_1(x,g) - P_2(x,g)|}{V(x)}, \tag{2.1}$$

where $P(x,g) := \int g(y)P(x,dy)$ for a kernel P and a measurable function g. Let P and P^n denote, respectively, the one-step and n-step transition kernels of $\{X_t\}$; if the stationary distribution of $\{X_t\}$ exists, denote it by π and define the kernel $\pi(x,A) = \pi(A)$ for all x and sets A. *V-uniform ergodicity* of $\{X_t\}$ is equivalent to geometric convergence of P to π in the V-norm (Meyn and Tweedie (1993), Theorem 16.0.1), i.e., $|||\, P^n - \pi \,|||_V \leq Rr^{-n}$, $R < \infty$, $r > 1$, for any integer n.

Note that $V \equiv 1$ implies both uniform ergodicity and geometric ergodicity. In addition to convergence of the n-step transition probabilities, V-uniform ergodicity implies convergence of moments: if $g(x) = x^p$ is bounded by the test function $V(x)$, then $P^n(x,g)$ converges to $\pi(g)$ in the V-norm. Meyn and Tweedie ((1993), Theorem 17.0.1) established asymptotic results for V-uniformly ergodic Markov chains with obvious implications for large sample inference about nonlinear time series. They also made connections between V-uniform ergodicity and mixing properties (see Liebscher (2005) for a more recent paper in a similar vein).

2.2 Equivalence of the V-uniform ergodicity of $\{X_t\}$ and $\{X_{tk}\}$.

Suppose k is a positive integer. If the Markov chain $\{X_t\}$ with transition kernel P is V-uniformly ergodic then so is $\{X_{tk}\}$, since $\{X_{tk}\}$ has transition kernel P^k and, restricting ourselves to positive integers n so that n/k is an integer,

$$||| P^n - \pi |||_V \leq Rr^{-n} \Longrightarrow ||| (P^k)^{n/k} - \pi |||_V \leq R(r^k)^{-n/k}, \;\; R < \infty, \;\; r^k > 1.$$

Thus V-uniform ergodicity of $\{X_t\}$ and $\{X_{tk}\}$ are equivalent if V-uniform ergodicity of $\{X_{tk}\}$ in turn implies V-uniform ergodicity of $\{X_t\}$. This is precisely what the following tells us.

Lemma 1. *Suppose $\{X_t\}$ is a Markov chain having transition kernel P with $||| P |||_V < \infty$. Assume for a positive integer $k < \infty$ and function $V \geq 1$ that $\{X_{tk}\}$ is V-uniformly ergodic. Then $\{X_t\}$ is V-uniformly ergodic.*

Lemma 1 allows us to restrict our search for stability conditions for $\{X_t\}$ to a search for stability conditions for $\{X_{tk}\}$.

2.2 Drift conditions for V-uniform ergodicity and transience.

Let $E_x[\cdot] := E_x[\cdot|X_0 = x]$ and $P_x(\cdot) := P_x(\cdot|X_0 = x)$. We state brief definitions here, reminding the reader of the text by Meyn and Tweedie (1993) as a reference. These are generalized definitions of standard concepts in the theory of countable state Markov chains.

A general state Markov chain $\{X_t\}$ with state-space $\mathbb{X}$ is called *ψ-irreducible* if there exists a measure ψ on the Borel sets of $\mathbb{X}$ such that whenever $\psi(A) > 0$, we have $P_x(\tau_A < \infty) > 0$ for all $x \in \mathbb{X}$, where τ_A is the first return time to the set A. A ψ-irreducible Markov chain is called *aperiodic* if there is no collection of disjoint subsets of the state-space, other than the state-space itself, that the chain cycles through with probability one.

Assuming aperiodicity and ψ-irreducibility of $\{X_t\}$, the various types of ergodicity of Markov chains can be demonstrated through the use of Foster-Lyapunov drift criteria. Lemma 2 generalizes drift conditions given by Meyn and Tweedie for V-uniform ergodicity ((1993), Theorem 16.0.1), and for transience (Tweedie (1976), Theorem 11.3).

Lemma 2. *Assume* $\{X_t\}$ *is a* ψ*-irreducible, aperiodic general state Markov chain on* $\mathbb{R}^p$.

1. *Suppose* $V(\cdot) \geq 1$ *is a measurable function, unbounded and locally bounded. If for some integer* $0 < k < \infty$ *and all* $M < \infty$
$$\limsup_{V(x)\to\infty} \frac{E_x[V(X_k)]}{V(x)} < 1, \qquad \sup_{V(x)\leq M} E_x[V(X_k)] < \infty, \qquad \sup_x \frac{E_x[V(X_1)]}{V(x)} < \infty \tag{2.2}$$
and the level sets $C_M^{(V)} := \{x : V(x) \leq M\}$ *are petite, then* $\{X_t\}$ *is* V*-uniformly ergodic.*

2. *Suppose* V *is a measurable, nonnegative, unbounded function. Suppose that* $\psi(\{x : V(x) > M\}) > 0$ *for all* $M < \infty$*. If for some integer* $0 < k < \infty$
$$\limsup_{V(x)\to\infty} E_x\left[\frac{V(x)}{V(X_k)}\right] < 1, \tag{2.3}$$
then $\{X_t\}$ *is transient.*

By (2.1), for any measurable $V \geq 1$ and measurable function g
$$|||\, P \,|||_V := \sup_x \sup_{|g|\leq V} \frac{|Pg|}{V(x)} = \sup_x \sup_{|g|\leq V} \frac{|E_x[g(X_1)]|}{V(x)} \leq \sup_x \frac{E_x[V(X_1)]}{V(x)},$$
implying that if $\{X_{tk}\}$ is V-uniformly ergodic and $\sup_x E_x[V(X_1)]/V(x) < \infty$, then $|||\, P \,|||_V < \infty$, an assumption of Lemma 1. The process $\{X_t\}$ cannot have unbounded explosions and be stable.

An alternative formulation of Lemma 2.1 above would assume $\{X_t\}$ is a T-chain. Theory concerning petite sets and T-chains may be found in Meyn and Tweedie (1993). Aperiodicity and ψ-irreducibility are easy consequences of the time series having an error distribution which is continuous and everywhere positive; the relationship between level sets being petite and $\{X_t\}$ being a T-chain is another consequence of this (Cline and Pu (1998)).

The function V referred to in Lemma 2 is called a *test function.* Applying drift-criteria to demonstrate the various forms of ergodicity requires constructing test functions of the process that satisfy these criteria. This is usually done on a case by case basis that can be tedious, and produces results limited in scope. We derive a 'catch-all' test function which applies in a variety of situations, saving the time and effort involved in test function construction.

The stability and transience conditions in Lemma 2 nearly partition the parameter space as is to be expected by the recurrence/transience dichotomy of Markov chains. One might expect a drift condition sufficient for transience would be that $\limsup_{V(x)\to\infty} E_x[V(X_k)/V(x)] > 1$, but note the intermediate condition $\liminf_{V(x)\to\infty} E_x[V(X_k)/V(x)] \geq 1$ would be sufficient for non-positivity, while transience is a 'stronger' form of instability than non-positivity and thus would require a more unstable drift condition, leaving us to require the stronger condition in (2.2): $\limsup_{V(x)\to\infty} E_x\big[V(x)/V(X_k)\big] < 1$.

3. Stability Results and Examples

3.1 Cyclic TAR($p; d; s$) models

Intuitively, when $\{X_t\}$ is large the errors $\{\xi_t\}$ become negligible and the stochastic process $\{X_t\}$ behaves like the deterministic system $\{x_t\}$. Suppose the dynamics of $\{x_t\}$ keep it away from the thresholds. When large, $\{X_t\}$ will therefore tend to avoid the thresholds. Thus $\{X_t\}$ will have a stability criterion analogous to that of $\{x_t\}$, i.e., that $\{X_t\}$ shrinks in expectation with each pass through the cycle. Demonstrating transience is a worst-case affair; if just one of the cycles were such that $\{x_t\}$ grew with each pass through the cycle, then $\{X_t\}$ would be transient.

Formally, suppose $\{X_t\}$ is as defined in (1.2), with state space $\mathbb{R}^p$ partitioned into regions R_i, each region with companion matrix A_i, $i = 1, \ldots, l$, $l < \infty$. Suppose there exist $m < \infty$ cycles $\mathbb{C}_1, \ldots, \mathbb{C}_m$ for the skeleton $\{x_t\}$ defined as in (1.3), each cycle $\mathbb{C}_i$ of finite length k_i, $i = 1, \ldots, m$. Suppose each cycle $\mathbb{C}_i$ consists of regions $R_1^{(i)}, \ldots, R_{k_i}^{(i)}$ with corresponding companion matrices $A_1^{(i)}, \ldots, A_{k_i}^{(i)}$, $i = 1, \ldots, m$.

For an integer $n < \infty$ and a sequence $i_1, \ldots, i_n$ of elements from $\{1, \ldots, l\}$, define an *n-path* to be a sequence of n regions $R_{i_1}, \ldots, R_{i_n}$, with companion matrices $A_{i_1}, \ldots, A_{i_n}$, that the skeleton $\{x_t\}$ of the process moves through, i.e., $x_{t-1} = x \in R_{i_j}$ implies $x_t = A_{i_j}x \in R_{i_{j+1}}$, $j = 1, \ldots, n-1$. When x begins in a cycle, the n-path consists of the regions in the cycle; when x does not begin in a cycle the n-path consists of some other collection of regions.

Let $\| \cdot \|$ denote the Euclidean norm throughout and let $B_\delta(x) := \{y : \| y - x \| < \delta\}$.

Assumption 1. (A1) *There exist $M_1 < \infty$ and, associated with each region R_j, $j = 1, \ldots, l$, a positive, strictly increasing function $g_j(\cdot)$ with $g(x) \to \infty$ as $x \to \infty$, so that for an arbitrary n-path $R_{i_1}, \ldots, R_{i_n}$, if $\| x \| > M_1$,*

$$x \in R_{i_j} \Rightarrow B_{g_{i_j}(\|A_{i_j}x\|)}(A_{i_j}x) \subset R_{i_{j+1}}, \quad j = 1, \ldots, n-1. \tag{3.1}$$

Assume also that each g_{i_j} satisfies a triangle inequality, $g_{i_j}(x+y) \le g_{i_j}(x) + g_{i_j}(y)$.

Assumption 2. (A2) *Assume $\{x_t\}$ is such that there exist $M_2, n^* < \infty$ so that for each $x_0 = x \notin \cup_i \mathbb{C}_i$ with $\| x \| > M_2$, there is an integer $d = d(x) \le n^*$ implying if $\| x_d \| > M_2$ then $x_d \in \cup_i \mathbb{C}_i$.*

(A1) has it that all $x \in \mathbb{R}^p$ with $\| x \|$ large are mapped by the skeleton away from the thresholds. (A2) is the condition that all large $x \in \mathbb{R}^p$ not in a cycle that remain large are mapped by the skeleton to a cycle within a uniformly finite time. In applications, satisfying these assumptions may require subdividing one or more regions; suppose the partition of the state space $R_1, \ldots, R_l$ is defined both by the thresholds and any necessary further subdivision of these regions.

Since the skeleton is piecewise linear, vectors are mapped to vectors and so, for example, if the rotation of the map is toward the interior of the next region, the assumptions are satisfied. Many of the classic examples in the literature (see Chan et al (1985), Chen and Tsay (1991), Guo and Petruccelli (1991), Kunitomo (2001) and Petruccelli and Woolford (1984) for instance) satisfy these assumptions, and their results follow immediately from ours.

Example 1: Petruccelli and Woolford (1984) considered the SETAR(2;1;1) model

$$X_t = \phi_1 I(X_{t-1} > 0) X_{t-1} + \phi_2 I(X_{t-1} \le 0) X_{t-1} + \xi_t.$$

Assume $\phi_1 \ne 0, \phi_2 \ne 0$. Note $\{X_t\}$ is a Markov chain. The skeleton of $\{X_t\}$ is given by $x_t = \phi_1 I(x_{t-1} > 0) x_{t-1} + \phi_2 I(x_{t-1} \le 0) x_{t-1}$. Denote the regions $R_1 = \{x : x > 0\}$, $R_2 = \{x : x \le 0\}$. Let $R_i \to R_j$ denote $x_{t-1} \in R_i$ implies $x_t \in R_j$, $i, j = 1, 2$.

Depending on the values of $\mathrm{sgn}(\phi_1)$ and $\mathrm{sgn}(\phi_2)$, the cycles could be $R_1 \to R_1$ and/or $R_2 \to R_2$, or $R_1 \to R_2 \to R_1$. These cycles are reached by any $x_0 = x \in \mathbb{R}$ within one transition ($n^* = 1$ for any $M_2 \geq 0$), satisfying (A2). Since $\phi_1 \neq 0$, $\phi_2 \neq 0$, (A1) is satisfied for any $M_1 \geq 0$ with $g_1(x) = |\phi_1 x|$, $g_2(x) = |\phi_2 x|$.

When one or more of the companion matrices are not of full rank, we can have x large in magnitude 'mapped small' by the skeleton. This does not prove to be a problem since stability concerns the behavior of large values of the process. If all of the A_i are of full rank, then large x are all mapped large by the skeleton. Moreover, note that we need the cyclicity assumption to hold only for large x.

We apply these points in the next example. The model is reminiscent of the ASTAR models discussed by Lewis and Stevens (1991), fit using adaptive regression splines.

Example 2: Consider the TAR(2;2;3) model

$$Y_t = \begin{cases} \phi_1^{(1)} Y_{t-1} + \phi_2^{(1)} Y_{t-2} + \xi_t, & Y_{t-1} \geq c,\ Y_{t-2} \geq 0, \\ \phi_1^{(2)} Y_{t-1} + \xi_t, & Y_{t-1} < c, \\ \phi_1^{(3)} Y_{t-1} + \xi_t, & Y_{t-1} \geq c,\ Y_{t-2} < 0. \end{cases}$$

Suppose $\xi_t \sim N(0, \sigma^2)$, $\phi_1^{(1)} > 0$, $\phi_2^{(1)} > 0$, $\phi_1^{(2)} < 0$, $\phi_1^{(3)} < 0$, $c > 0$. Let $X_t = (Y_t, Y_{t-1})'$, $x_t = (y_t, y_{t-1})'$, $\nu_t = (\xi_t, 0)'$ and

$$A_1 = \begin{pmatrix} \phi_1^{(1)} & \phi_2^{(1)} \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \phi_1^{(2)} & 0 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} \phi_1^{(3)} & 0 \\ 1 & 0 \end{pmatrix}.$$

Define regions $R_1 = \{(y_1, y_2) : y_1 \geq c, y_2 \geq 0\}$, $R_2 = \{(y_1, y_2) : y_1 < c\}$, $R_3 = \{(y_1, y_2) : y_1 \geq c, y_2 < 0\}$. The general state Markov chain $\{X_t\}$ on $\mathbb{R}^2$ and its skeleton $\{x_t\}$ are

$$X_t = A_1 X_{t-1} I(X_{t-1} \in R_1) + A_2 X_{t-1} I(X_{t-1} \in R_2) + A_3 X_{t-1} I(X_{t-1} \in R_3) + \nu_t,$$
$$x_t = A_1 x_{t-1} I(x_{t-1} \in R_1) + A_2 x_{t-1} I(x_{t-1} \in R_2) + A_3 x_{t-1} I(x_{t-1} \in R_3).$$

Let $R_i \to R_j$ denote $x_{t-1} \in R_i$ implies $x_t \in R_j$. When the magnitude is large the dynamics for the skeleton are $R_1 \to R_1$, $R_2 \to R_2$, $R_2 \to R_3$, $R_3 \to R_2$. The region R_2 feeds into two regions: when $|y_{t-1}|$ is small, points in R_2 are mapped back to R_2; when $|y_{t-1}|$ is large, points in R_2 are mapped to R_3. Suppose

$x_0 = x \in R_2$. For $M_2 < \infty$,

$$\| x_1 \| = \| A_2 x \| = |y_{t-1}| \times \| (\phi_1^{(2)}, 1)' \| > M_2 \Leftrightarrow |y_{t-1}| > M_2 / ([\phi_1^{(2)}]^2 + 1).$$

Update the definition of regions by defining $R_2' = \{x : \| A_2 x \| \leq M_2\}$ and $R_2 = \{x : \| A_2 x \| > M_2\}$. Picking M_2 large enough leaves us with $R_1 \to R_1$ and $R_2 \to R_3 \to R_2$ as the cycles of interest, satisfying (A2) with $n^* = 1$. The points in R_2' likewise satisfy (A2) since, though they do not reach a cycle, they do not grow large in magnitude.

Points $x \in R_2$ are mapped to the ray $(\phi_1^{(2)}, 1)'$, while points $x \in R_3$ are mapped to the ray $(\phi_1^{(3)}, 1)'$. The matrix A_1 is full-rank. Clearly then an $M_1 < \infty$ and functions $g_1(\cdot), g_2(\cdot), g_3(\cdot)$ exist so that (A1) is satisfied.

We turn to the stability result. For a square matrix A, let $\rho(A)$ denote the spectral radius of A.

Theorem 1. *Suppose $\{X_t\}$ is as defined in* (1.2)*, with state space $\mathbb{R}^p$ partitioned into regions R_i and each region with companion matrix A_i, $i = 1, \ldots, l$. Suppose (A1), (A2) are satisfied, $E|\xi_t|^r < \infty$ for some $r > 0$, and ξ_t has a continuous density everywhere positive. Then*

1. *$\{X_t\}$ is V-uniformly ergodic if*
$$\max_{i \in \{1,\ldots,m\}} \rho(\prod_{j=1}^{k_i} A_j^{(i)}) < 1;$$
2. *$\{X_t\}$ is transient if*
$$\max_{i \in \{1,\ldots,m\}} \rho(\prod_{j=1}^{k_i} A_j^{(i)}) > 1.$$

Assumptions (A1) and (A2) on the skeleton, when combined with the appropriate stability condition on the skeleton and a 'stability' condition on the errors, give us sufficient conditions for V-uniform ergodicity of $\{X_t\}$. When (A1) and (A2) are combined with an instability condition on the skeleton and a 'stability' condition on the errors, we have sufficient conditions for transience of $\{X_t\}$.

Previous authors have employed the strategy of analyzing the dynamics of the skeleton to find the cycles and combining the stability condition on these cycles with one on the error distribution to construct a test function which satisfies the appropriate drift criterion. Chan et al. (1985), Kunitomo (2001), Lim (1992) and Petruccelli and Woolford (1984), among others, all used this logic in dealing with specific models. We emphasize again that our results allow one to proceed directly from analysis of the skeleton to stability without having to devise a test function specific to a given model.

Example 1 (cont.): With an appropriate condition on the error distribution, we have Petruccelli and Woolford's (1984) sufficient condition for ergodicity (geometric in their case, V-uniform in ours); that $\phi_1 < 1$, $\phi_2 < 1$, $\phi_1\phi_2 < 1$, and the sufficient condition for transience that $\phi_1 > 1$ or $\phi_2 > 1$ or $\phi_1\phi_2 > 1$.

Theorem 1 applies beyond existing results, and is able to handle more elaborate models whose stability properties are not detailed in the literature.

Example 2 (cont.): Supposing $\xi_t \sim N(0, \sigma^2)$, the assumptions behind Theorem 1 are satisfied and we have $\{X_t\}$ is V-uniformly ergodic if $\phi_1^{(1)} > 0$, $\phi_2^{(1)} > 0$, $\phi_1^{(2)} < 0$, $\phi_1^{(3)} < 0$ if

$$\rho(A_1) < 1 \Leftrightarrow \phi_1^{(1)} + \phi_2^{(1)} < 1, \quad \rho(A_2 A_3) < 1 \Leftrightarrow \phi_1^{(2)}\phi_1^{(3)} < 1;$$

$\{X_t\}$ is transient if either

$$\rho(A_1) > 1 \Leftrightarrow \phi_1^{(1)} + \phi_2^{(1)} > 1, \quad \text{or} \quad \rho(\mathrm{A}_2\mathrm{A}_3) > 1 \Leftrightarrow \phi_1^{(2)}\phi_1^{(3)} > 1.$$

Note the ergodic parameter spaces in the examples are unbounded, contrary to what we would expect through analogy with linear AR(p) time series. Both examples reflect the fact that commonly stated global conditions for stability can unnecessarily restrict the parameter space. Focusing on only the relevant cyclic behavior gives sharper bounds, leading to a wider variety of better performing models. The gain can be tremendous; as in the examples, other threshold autoregressive models have been shown to have unbounded parameter spaces.

3.2 Nonlinear autoregressive models

The nonlinear autoregressive model of order p (NLAR(p)) with additive noise ξ_t is

$$Y_t = f(Y_{t-1}, \ldots, Y_{t-p}) + \xi_t, \tag{3.3}$$

where $\{\xi_t\}$ is a mean zero i.i.d. sequence and f is a nonlinear function. Consider the general state Markov chain $\{X_t\}$ with $X_t = (Y_t, \ldots, Y_{t-p+1})'$. By Chan (in Tong (1990), Prop. A1.7), if f is bounded on bounded sets and ξ_t has a pdf that is everywhere continuous and positive, then $\{X_t\}$ is aperiodic and ψ-irreducible, ψ being Lebesgue measure.

Tong (1990) suggests finding stability conditions for (3.3) through local linearization. Describe the transitions of $\{X_t\}$ by the mapping $\Phi(\cdot)$, where

$$\Phi(x_{t-1}) = \Phi(y_{t-1}, \ldots, y_{t-p}) = (f(y_{t-1}, \ldots, y_{t-p}), y_{t-1}, \ldots, y_{t-p+1})'.$$

Supposing f is continuous and everywhere differentiable, we can use a Taylor expansion to approximate $\{X_t\}$ by $X_t = J(x)X_{t-1} + c(x) + \nu_t$ around the point $x = (y_1, \ldots, y_p)'$, where

$$J(x) = \begin{pmatrix} \frac{\partial}{\partial y_1} f(x) & \cdots & \frac{\partial}{\partial y_{p-1}} f(x) & \frac{\partial}{\partial y_p} f(x) \\ 1 & \cdots & 0 & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}$$

is the Jacobian of $\Phi(\cdot)$ evaluated at x, $\nu_t = \xi_t(1, 0, \ldots, 0)$ and $c(x) = \Phi(x) - J(x)x$ plus the remainder of the Taylor expansion. With appropriate conditions on $c(x)$ and the errors, the condition for stability would then be $\limsup_{\|x\|\to\infty} \rho(J(x)) < 1$.

We loosen the restriction that $f(\cdot)$ be continuous and differentiable (or even Lipschitz) and instead require $\{X_t\}$ be asymptotically *threshold-like*, i.e., $\Phi(x) = \sum_{i=1}^{l} A_i I(x \in R_i)x + h(x)$, with appropriate conditions on h for large x. We then apply the work done for cyclic TAR($p; d; s$) models, allowing us to weaken the global condition $\limsup_{\|x\|\to\infty} \rho(J(x)) < 1$ to a condition on the relevant cyclic behavior, so that the process need not shrink in expectation at every transition as $\limsup_{\|x\|\to\infty} \rho(J(x)) < 1$ implies, just at every pass through the cycle, giving sharper bounds on the parameter space.

Theorem 2. *Assume there exists* $M_3 < \infty$ *so that for* $X_{t-1} = x$ *with* $\| x \| > M_3$ *we can embed the NLAR(p) model* (3.3) *in the Markov chain*

$$X_t = \sum_{i=1}^{l} A_i I(X_{t-1} \in R_i) X_{t-1} + h(X_{t-1}) + \nu_t, \tag{3.4}$$

where $x_t = \sum_{i=1}^{l} A_i I(x_{t-1} \in R_i) x_{t-1}$ *satisfies the conditions of Theorem 1,* $h(x) = O(\| x \|^{\epsilon})$, $0 < \epsilon < 1$, $E|\xi_t|^r < \infty$ *for some* $r > 0$, *and* ξ_t *has a continuous density everywhere positive. Then*

1. $\{X_t\}$ *is* V*-uniformly ergodic if*

$$\max_{i \in \{1,\ldots,m\}} \rho\Big(\prod_{j=1}^{k_i} A_j^{(i)}\Big) < 1; \tag{3.5}$$

2. $\{X_t\}$ *is transient if*

$$\max_{i \in \{1,\ldots,m\}} \rho\Big(\prod_{j=1}^{k_i} A_j^{(i)}\Big) > 1. \tag{3.6}$$

Example 3: Consider the threshold-like EXPAR model (Tong (1990)) which generalizes the TAR(2;2;3) model in Example 2:

$$Y_t = \begin{cases} (\alpha_1^{(1)} + \beta_1^{(1)} e^{-Y_{t-1}^2}) Y_{t-1} + (\alpha_2^{(1)} + \beta_2^{(1)} e^{-Y_{t-2}^2}) Y_{t-2} + \xi_t, & Y_{t-1} \geq c,\ Y_{t-2} \geq 0, \\ (\alpha_1^{(2)} + \beta_1^{(2)} e^{-Y_{t-1}^2}) Y_{t-1} + \xi_t, & Y_{t-1} < c, \\ (\alpha_1^{(3)} + \beta_1^{(3)} e^{-Y_{t-1}^2}) Y_{t-1} + \xi_t, & Y_{t-1} \geq c,\ Y_{t-2} < 0. \end{cases}$$

Let $X_t = (Y_t, Y_{t-1})'$, $R_1 = \{(y_1, y_2) : y_1 \geq c, y_2 \geq 0\}$, $R_2 = \{(y_1, y_2) : y_1 < c\}$, $R_3 = \{(y_1, y_2) : y_1 \geq c, y_2 < 0\}$. Then

$$X_t = \sum A_i I(X_{t-1} \in R_i) X_{t-1} + h(X_{t-1}) + \nu_t,$$

where

$$A_1 = \begin{pmatrix} \alpha_1^{(1)} & \alpha_2^{(1)} \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \alpha_1^{(2)} & 0 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} \alpha_1^{(3)} & 0 \\ 1 & 0 \end{pmatrix},$$

$$h(X_{t-1}) = \begin{cases} (\beta_1^{(1)} e^{-Y_{t-1}^2} Y_{t-1} + \beta_2^{(1)} e^{-Y_{t-1}^2} Y_{t-2}, 0)', & X_{t-1} \in R_1 \\ (\beta_1^{(2)} e^{-Y_{t-1}^2} Y_{t-1}, 0)', & X_{t-1} \in R_2 \\ (\beta_1^{(3)} e^{-Y_{t-1}^2} Y_{t-1}, 0)', & X_{t-1} \in R_3. \end{cases}$$

Note that $h(x)$ is bounded. By this and reasoning similar to Example 2, if $E|\xi_t|^r < \infty$ then $\{X_t\}$ is V-uniformly ergodic in the case where $\alpha_1^{(1)} > 0$, $\alpha_2^{(1)} > 0$, $\alpha_1^{(2)} < 0$ and $\alpha_1^{(3)} < 0$ if $\rho(A_1) < 1 \Leftrightarrow \alpha_1^{(1)} + \alpha_2^{(1)} < 1$ and $\rho(A_2A_3) < 1 \Leftrightarrow \alpha_1^{(2)}\alpha_1^{(3)} < 1$, while $\{X_t\}$ is transient if either $\rho(A_1) > 1 \Leftrightarrow \alpha_1^{(1)} + \alpha_2^{(1)} > 1$ or $\rho(A_2A_3) > 1 \Leftrightarrow \alpha_1^{(2)}\alpha_1^{(3)} > 1$.

4.1 Discussion

When $\max_{i\in\{1,\ldots,m\}} \rho\big(\prod_{j=1}^{k_i} A_j^{(i)}\big) = 1$, the processes $\{X_t\}$ and $\{Y_t\}$ are most likely null recurrent, though this is an open question. If intercept terms are included, however, the process can be ergodic or transient as well. The error distribution can also have an effect. See, for example, the TAR(1) models considered in Cline and Pu (1999).

Note assumptions (A1) and (A2) exclude chaotic dynamics of the skeleton. There can be no arbitrarily long cycles or dense orbits before reaching a cycle—unless the process stays small in magnitude. Interestingly, this points out the need for uniformity in the behavior of the skeleton in order to determine stability. Chaotic dynamical systems with bounded attractors have ergodic distributions of their own. Processes with a chaotic skeleton perhaps have a limiting distribution, where it exists, that is a mixture of the chaotic ergodic and the error distributions.

4.2 Acknowledgements

The authors wish to thank the anonymous referees for their helpful suggestions.

5. Proofs

Proof of Lemma 1.

Proof. Since $||| P |||_V < \infty$, there exists $Q < \infty$ such that $||| P |||_V \leq Q$. Denote the common invariant distribution of $\{X_{tk}\}$ and $\{X_t\}$ by π. Note $\{X_{tk}\}$ has transition kernel P^k. Since $\{X_{tk}\}$ is V-uniformly ergodic,

$$||| P^{kn} - \pi |||_V = ||| (P^k)^n - \pi |||_V \leq Rr^{-n} = Rr^{-\frac{nk}{k}} = R(r^{1/k})^{-nk} = Rr_*^{-nk}, \quad (5.1)$$

where $r_* = r^{1/k} > 1$. Then for integers n and $1 \leq j < k$, k fixed, using (5.1) and the invariance of π,

$$||| P^{kn+j} - \pi |||_V = ||| P^j(P^{kn} - \pi) |||_V \leq (||| P |||_V)^j \, ||| P^{kn} - \pi |||_V \leq R' r_*^{-(kn+j)},$$

where $R' = \max(Q, Q^k)Rr_*^k$. Since each integer $n' = nk+j$ for some n, $1 \leq j < k$, this implies $\|| P^{n'} - \pi \||_V \leq R' r_*^{-n'}$, $r^* > 1$, $R' < \infty$, which by Meyn and Tweedie ((1993), Theorem 16.0.1) is true if and only if $\{X_t\}$ is V-uniformly ergodic. □

Proof of Lemma 2.

Proof. (1) By (2.2), and since V is unbounded, there exists $M < \infty$ so that $V(x) > M$ implies $E_x[V(X_k)]/V(x)] < 1$. Since $C_M^{(V)} := \{x : V(x) \leq M\}$ is petite, $V(\cdot) \geq 1$ and locally bounded, there exist $K < \infty$, $\beta > 0$ so that $E_x[V(V_k)] - V(x) \leq -\beta V(x) + KI_C(x)$; thus by Meyn and Tweedie ((1993), Theorem 16.0.1), $\{X_{tk}\}$ is V-uniformly ergodic. Since $\sup_x E_x[V(X_1)/V(x)] < \infty$ then $\|| P \||_V < \infty$ and, by Lemma 1, $\{X_t\}$ is V-uniformly ergodic.
(2) Transience follows immediately from Cline and Pu (2001a), Lemma 4.1) and Meyn and Tweedie ((1993), Theorem 8.2.6). □

Recall we defined an *n-path* to be a sequence of n regions $R_{i_1}, \ldots, R_{i_n}$, with companion matrices $A_{i_1}, \ldots, A_{i_n}$, that the skeleton of the process moves through, i.e., $x_{t-1} = x \in R_{i_j}$ implies $x_t = A_{i_j} x \in R_{i_{j+1}}$, $j = 1, \ldots, n-1$. The next lemma demonstrates any initial $X_0 = x$ large enough in magnitude will remain large and in the n-path with a high probability. We make use of this lemma several times in the proofs of Theorems 1 and 2. Let $P_x(\cdot) := P(\cdot | X_0 = x)$ and recall that $\| \cdot \|$ denotes the Euclidean norm.

Lemma 3. *Suppose* $\{X_t\}$ *as at* (1.2), *(A1) holds, and* $E|\xi_t|^r < \infty$ *for some* $r > 0$. *For an integer n assume, beginning at* $x_0 = x$, *the skeleton* $\{x_t\}$ *follows the n-path* $R_{i_1}, \ldots, R_{i_n}$. *Let* $q(x) := [\min(\min_l \{g_{i_l}(\| \prod_{j=1}^n A_{i_j} x \|)\}, \| \prod_{j=1}^n A_{i_j} x \|)]^{-r(n-1)}$. *Then*

$$P_x \left(\cup_{k=1}^{n-1} [(X_k \notin R_{i_{k+1}}) \cup (\| X_k \| \leq M)] \right) = O(q(x)). \tag{5.2}$$

Proof. Let C_0 be the empty set and $C_k = \cap_{j=1}^k [(X_j \in R_{i_{j+1}}) \cap (\| X_j \| > M)]$ for $k = 1, \ldots, n$. Then

$$P_x(C_n) = P_x(C_1) \prod_{k=2}^n P_x \left((X_k \in R_{i_{k+1}}) \cap (\| X_k \| > M) \middle| C_{k-1} \right) \tag{5.3}$$

We can write $X_k = \prod_{j=1}^k A_{i_j} x + \sum_{l=1}^{k-1} (\prod_{m=1}^l A_{i_m}) \xi_l$, where the process remains in the n-path until time k.

Note that $\| \prod_{j=1}^{k} A_{i_j} x \| \geq \| \prod_{j=1}^{n} A_{i_j} x \| / \prod_{j=k+1}^{n} \| A_{i_j} \|$, for $k = 1, \ldots, n$, so defining $D_1 := [\max_{k \in \{1,\ldots,n\}} \prod_{j=k+1}^{n} \| A_{i_j} \|]^{-1}$, we have for $k = 1, \ldots, n$ that $\| \prod_{j=1}^{k} A_{i_j} x \| \geq D_1 \| \prod_{j=1}^{n} A_{i_j} x \|$.

Get functions $g_{i_j}(\cdot)$ and $M = M_1$ from (A1). Pick $D_2 \geq [\max_j \| A_{i_j} \|]^n$. Suppose w.l.o.g. $D_2 > 1$. Then since the g_{i_j} satisfy a triangle inequality, we have for $k = 1, \ldots, n$ by Markov's Inequality,

$$\begin{aligned} P_x(X_k \notin R_{i_{k+1}} | C_{k-1}) &\leq P\Big(|\xi_l| + D_2 \sum_{l=1}^{k-1} g_{i_l}(|\xi_l|) > g_{i_k}(\prod_{j=1}^{k} \| A_{i_k} x \|) \Big) \\ &\leq \frac{(1+D_2) \max_{j \in \{1,\ldots,n\}} g_{i_j}(E|\xi_k|^r)}{[g_{i_k}(\| \prod_{j=1}^{k} A_{i_j} x \|)]^r} \\ &\leq \frac{(1+D_2) \max_{j \in \{1,\ldots,n\}} g_{i_j}(E|\xi_k|^r)}{[g_{i_k}(D_1 \| \prod_{j=1}^{n} A_{i_j} x \|)]^r}, \end{aligned}$$

$$\begin{aligned} P_x(\| X_k \| \leq M | C_{k-1}) &\leq P\Big(|\xi_k| + D_2 \sum_{l=1}^{k-1} |\xi_l| \geq \| \prod_{j=1}^{k} A_{i_k} x \| - M \Big) \\ &\leq \frac{(1+D_2) E|\xi_k|^r}{(D_1 \| \prod_{j=1}^{n} A_{i_j} x \| - M)^r}. \end{aligned}$$

Then by complementation for $k = 1, \ldots, n$,

$$\begin{aligned} & P_x\Big((X_k \in R_{i_{k+1}}) \cap (\| X_k \| > M) \Big| C_{k-1} \Big) \\ \geq\ & 1 - 2 \max([g_{i_k}(D_1 \| \prod_{j=1}^{n} A_{i_j} x \|)]^{-r}, (D_1 \| \prod_{j=1}^{n} A_{i_j} x \| - M)^{-r}). \end{aligned} \tag{5.4}$$

Using (5.3) and (5.4),

$$\begin{aligned} & P_x \left(\cap_{k=1}^{n-1} [(X_k \in R_{i_{k+1}}) \cap (\| X_k \| > M)] \right) \\ \geq\ & \prod_{k=1}^{n-1} [1 - 2 \max([\max_l g_{i_l}(D_1 \| \prod_{j=1}^{n} A_{i_j} x \|)]^{-r}, (D_1 \| \prod_{j=1}^{n} A_{i_j} x \| - M)^{-r}) \\ =\ & [1 - 2 \max([\max_l g_{i_l}(D_1 \| \prod_{j=1}^{n} A_{i_j} x \|)]^{-r}, (D_1 \| \prod_{j=1}^{n} A_{i_j} x \| - M)^{-r})]^{n-1}. \end{aligned}$$

The conclusion follows. □

The following result is partially due to Ciarlet (1982). The proof is ours.

Lemma 4. *If a matrix A has $\rho(A) < 1$, then there exists a matrix norm $\| \cdot \|_{mat}$ induced by a vector norm $\| \cdot \|_{vec}$ and a constant $0 < \lambda < 1$ such that*

$$\| Ax \|_{vec} \leq \| A \|_{mat} \| x \|_{vec} \leq \lambda \| x \|_{vec}, \ \forall x. \tag{5.5}$$

Also, with $\| \cdot \|$ the Euclidean norm, $\| x \| \to \infty$ implies $\| x \|_{vec} \to \infty$.

Proof. It is a well-known fact (Martelli (1992), Lemma 4.2.1, for example) that $\rho(A) < 1$ implies the existence of an integer $c < \infty$ and a vector norm

$$\| x \|_{vec} := \| x \| + \| Ax \| + \cdots + \| A^{c-1}x \|, \tag{5.6}$$

with the matrix norm $\| \cdot \|_{mat}$ induced by $\| \cdot \|_{vec}$ having $\| A \|_{mat} < 1$. Also, $\| A \|_{mat} < 1$ implies the existence of a constant $0 < \lambda < 1$ with $\lambda \geq \| A \|_{mat}$. Then for all x by a norm inequality, $\| Ax \|_{vec} \leq \| A \|_{mat} \| x \|_{vec} \leq \lambda \| x \|_{vec}$. From (5.6) and the non-negativity of $\| \cdot \|$ we have that $\| x \| \to \infty$ implies $\| x \|_{vec} \to \infty$. □

Lemma 5. *Suppose $\{X_t\}$ as at (1.2) and the conditions of Theorem 1 (1.) hold. Let $\| \cdot \|_i$ denote the vector norm for cycle $\mathbb{C}_i$, $i = 1, \ldots, m$, implied by Lemma 4, let $\| \cdot \|$ denote the Euclidean norm, and define*

$$V'(x) := \sum_{i=1}^{m} \| x \|_i^{r/2} I\{x \in \mathbb{C}_i\} + \| x \|^{r/2} I\{x \notin \cup_{i=1}^{m} \mathbb{C}_i\}.$$

Then the level sets $C_{M_1}^{E[V']} = \{x : E_x[V'(X_n)] \leq M_1\}$ are petite for any integer n and each $M_1 < \infty$.

Proof. By the definition of $\| \cdot \|_i$ in (5.6), clearly $E_x[\| X_n \|^{r/2}] \leq E_x[V'(X_n)]$. Since all matrix norms are finite, using norm inequalities and (5.6) again, there exists $C < \infty$ with $E_x[V'(X_n)] \leq C E_x[\| X_n \|^{r/2}]$. Thus $E_x[V'(X_n)] \to \infty$ iff $E_x[\| X_n \|] \to \infty$. From this and $E|\xi_t|^r < \infty$, we can clearly find $M_1, M_2 < \infty$ so that

$$\inf_{\{x:\, E_x[V'(X_n)] \leq M_1\}} P_x(\| X_n \| \leq M_2) > 0.$$

By Meyn and Tweedie ((1993), Prop. 5.5.4) this implies $\{x : E_x[V'(X_n)] \leq M_1\}$ is petite for the sampled chain $\{X_{nt}\}$, and thus is petite for $\{X_t\}$. □

In the following proofs of Theorems 1 and 2 the strategy is simple: find a test function V so that the appropriate drift condition, whether (2.2) for V-uniform ergodicity or (2.3) for transience, is satisfied. The bulk of the proofs involves creating the test function V by piecing it together from consideration of the dynamics of the skeleton and the assumptions on the Markov chain in which the time series is embedded.

Proof of Theorem 1.

Proof. (1) We need a test function $V(\cdot)$ for $\{X_t\}$ so that (2.2) is satisfied. By the assumptions and Lemma 4, there exists $\lambda = \max(\lambda_1, \ldots, \lambda_m)$ and for each cycle $\mathbb{C}_i$ there exist vector norms $\| \cdot \|_i$ with $x \in \mathbb{C}_i$ satisfying

$$0 < \rho(\prod_{j=1}^{k_i} A_j^{(i)}) < \lambda < 1, \quad \| \prod_{j=1}^{k_i} A_j^{(i)} x \|_i < \lambda \| x \|_i, \quad i = 1, \ldots, m. \tag{5.7}$$

Get r from $E|\xi_t|^r < \infty$. Suppose $r < 1$ and define

$$V'(x) := \sum_{i=1}^{m} \| x \|_i^{r/2} I\{x \in \mathbb{C}_i\} + \| x \|^{r/2} I\{x \notin \cup_{i=1}^{m} \mathbb{C}_i\}.$$

Get n^* and let $k := \prod k_i$, $M = \max(M_1, M2)$ from the assumptions. Suppose w.l.o.g. that $x_{n^*} \in R_1^{(1)}$, and $x_0 = x$ follows the n^*-path $R_{i_1}, \ldots, R_{i_n^*}$ before entering the cycle.

Let $I_k = I\{[\cap_{j=1}^{k}(X_{n^*+j} \in R_{j+1}^{(1)})] \cap [\cap_{j=1}^{k}(\| X_{n^*+j} \| > M)]\}$ and denote $E_{X_{n^*}}(\cdot) := E(\cdot|X_{n^*})$. Then if $\| X_{n^*} \| > M$, $X_{n^*} \in R_1^{(1)}$, from (5.7) and by $E|\xi_t|^r < \infty$ there exists $K_1 < \infty$ so that

$$E_{X_{n^*}}[V'(X_{n^*+k})I_k] < \lambda^{k/k_1} V'(X_{n^*}) + K_1. \tag{5.8}$$

Let I_k^c denote the indicator of the complement of the argument of I_k. Suppose w.l.o.g. that $\min_j[g_j(y)] \leq y$ for large y. Then by (3.1), applying Lemma 3 with $n = k$, conditioning on X_{n^*}, and applying Cauchy-Schwarz there exist $K_2, K_3 < \infty$ so that $\| X_{n^*} \| > M$ and $X_{n^*} \in R_1^{(1)}$ implies

$$E_{X_{n^*}}[V'(X_{n^*+k})I_k^c] < (K_2 V'(X_{n^*}) + K_3) O(\| \prod_{j=1}^{k} A_j^{(1)} X_{n^*} \|^{-(k-1)r/2}). \tag{5.9}$$

From (5.8) and (5.9), $X_{n^*} \in R_1^{(1)}$ with $\| X_{n^*} \| > M$ implies

$$\begin{aligned} & E_{X_{n^*}}[V^{'}(X_{n^*+k})] \\ < \quad & \lambda^{k/k_1} V^{'}(X_{n^*}) + K_1 + (K_2 V^{'}(X_{n^*}) + K_3) O(\| \prod_{j=1}^{k} A_j^{(1)} X_{n^*} \|^{-(k-1)r/2}). \end{aligned} \tag{5.10}$$

By (3.1), applying Lemma 3 with $n = n^* - 1$, conditioning on $X_0 = x$, and applying Cauchy-Schwarz, we have similar to (5.10) that if $\| X_{n^*} \| < M$ or $X_{n^*} \notin R_1^{(1)}$,

$$E_{X_{n^*}}[V^{'}(X_{n^*+k})] < (K_2 V^{'}(X_{n^*}) + K_3) O(\| \prod_{j=1}^{n^*} A_{i_j} x \|^{-(n^*-1)r/2}). \tag{5.11}$$

Note $O(\| \prod_{j=1}^{k} A_j^{(1)} X_{n^*} \|^{-(k-1)r/2}) = O_p(\| \prod_{j=1}^{n^*} A_{i_j} x \|^{-(k-1)r/2})$. Note also from the definition of $V^{'}$ and the triangle inequalities that there exists $C < \infty$ such that $E_x[\| X_{n^*} \|^{r/2}] \leq E_x[V^{'}(X_{n^*})] \leq C E_x[\| X_{n^*} \|^{r/2}]$. Combining this with Jensen's Inequality applied to $E_x[\| X_{n^*} \|^{r/2}]$ and $C E_x[\| X_{n^*} \|^{r/2}]$, implies we have $E_x[V^{'}(X_{n^*})] \to \infty$ iff $\| \prod_{j=1}^{n^*} A_{i_j} x \|$ does. Iterating expectations implies $E_x(E_{X_{n^*}}[V^{'}(X_{n^*+k})]) = E_x[V^{'}(X_{n^*+k})] = E_x(E_{X_k}[V^{'}(X_{n^*+k})])$. From this, (5.10) and (5.11),

$$\limsup_{E_x[V^{'}(X_{n^*})] \to \infty} \frac{E_x(E_{X_k}[V^{'}(X_{n^*+k})])}{E_x[V^{'}(X_{n^*})]} = \lambda^{k/k_1} < 1. \tag{5.12}$$

Let $V(x) := 1 + E_x[V^{'}(X_{n^*})]$. Note $V \geq 1$ is locally bounded, unbounded and measurable and, by (5.12), V satisfies (2.2). Applying Lemma 5 with $n = n^*$, it follows the level sets $C_{M_1}^V = \{x : V(x) \leq M_1\}$ are petite; thus by Lemma 2 the process $\{X_t\}$ is V-uniformly ergodic.

(2) We need a test function $V(\cdot)$ for $\{X_t\}$ so that (2.3) is satisfied. Suppose w.l.o.g. $\mathbb{C}_1$ has $\rho := \rho(\prod_{j=1}^{k_1} A_j^{(1)}) > 1$ and $X_0 = x \in R_1^{(1)} \in \mathbb{C}_1$. Let e_1 denote a normalized eigenvector corresponding to the eigenvalue λ_1 with $\rho = |\lambda_1| > 1$, and let $a_1(x)$ denote the projection of x onto e_1.

Let $I_{k_1} = I\{[\cap_{j=1}^{k_1}(X_j \in R_{j+1}^{(i)})] \cap [\cap_{j=1}^{k_1}(\| X_j \| > M)]\}$ and $I_{k_1}^C$ denote the indicator of the complement of the argument of I_{k_1}. Get r from the assumptions and pick $q < r/2$. Define $\phi(x) = \min[g_j(x)]$. The assumptions on the g_j imply that $\phi(\cdot)$ is strictly increasing. By Cauchy-Schwarz,

$$\begin{aligned} & E_x\Big[\frac{1+[\phi(|a_1(x)|)]^q}{1+[\phi(|a_1(X_{k_1})|)]^q}\Big] \\ \leq \quad & E_x\Big[\frac{1+[\phi(|a_1(x)|)]^q}{1+[\phi(|a_1(X_{k_1})|)]^q} I_{k_1}\Big] + (1+[\phi(|a_1(x)|)]^q)[E_x(I_{k_1}^C)]^{1/2}. \end{aligned} \tag{5.13}$$

Note that $E_x(I_{k_1}^C) = O([\phi(\| \prod_{j=1}^{k_1} A_j^{(1)} x \|)]^{-(k_1-1)r}) = O([\phi(|a_1(x)|)]^{-(k_1-1)r})$ follows from $E|\xi_t|^r < \infty$, the definition of $\phi(\cdot)$, and Lemma 3 with $n = k_1$. Then since $q < r/2$, we have $|a_1(x)| \to \infty$ implies $(1+[\phi(|a_1(x)|)]^q)[E_x(I_{k_1}^C)]^{1/2} \to 0$.

Let $I_g := I\{\phi(|a_1(X_{k_1})|) \leq \phi(|a_1(x)|)\}$. Then

$$E_x\Big[\frac{1+[\phi(|a_1(x)|)]^q}{1+[\phi(|a_1(X_{k_1})|)]^q} I_{k_1}\Big] < 1 + (1+[\phi(|a_1(x)|)]^q)[E_x(I_{k_1} I_g)]^{1/2}. \tag{5.14}$$

Now $E_x(I_{k_1} I_g) < P_x(I_g = 1 | I_{k_1} = 1)$ and, as $|a_1(x)| \to \infty$,

$$P_x(I_g = 1 | I_{k_1} = 1) = O([\phi(|a_1(x)|)]^{-r}). \tag{5.15}$$

Define the test function $V(x) = (1+[\phi(|a_1(x)|)]^q)$. Then by (5.13) through (5.15)

$$\limsup_{V(x)\to\infty} E_x\Big[\frac{V(x)}{V(X_{k_1})}\Big] = \limsup_{|a_1(x)|\to\infty} E_x\Big[\frac{(1+[\phi(|a_1(x)|)]^q)}{(1+[\phi(|a_1(X_{k_1})|)]^q)}\Big] \quad < \quad 1. \tag{5.16}$$

By assumption $\{X_t\}$ is aperiodic and ψ-irreducible, with ψ being Lebesgue measure. Clearly $\psi(\{x : V(x) > M\}) > 0$ for all $M < \infty$. By Lemma 2, $\{X_t\}$ is transient. □

In the proof of transience in Theorem 1 we concentrated on what happens to the process projected along the eigenvector corresponding to the eigenvalue of maximum modulus. When this is larger than one, the process grows in this direction and is therefore transient. This is the logic used by such previous authors as Kunitomo (2001), Petruccelli and Woolford (1984) and Tjøstheim (1990) in their proofs of transience for simpler models.

Proof of Theorem 2.

Proof. (1) Define $k = \prod k_i$ and $V'(x) := \sum_{i=1}^m \| x \|_i^{r/2} I\{x \in \mathbb{C}_i\} + \| x \|^{r/2}$

$I\{x \notin \cup_{i=1}^{m} \mathbb{C}_i\}$. Then there exist $K_1, \ldots, K_5 < \infty$, $0 < \lambda < 1$ so that

$$\begin{aligned} & E_x[V^{'}(X_{n^*+k})|X_{n^*}] \\ < \quad & \lambda^{k/k_1} V^{'}(X_{n^*}) + K_1 + (K_2 V^{'}(X_{n^*}) + K_3) O(\| \prod_{j=1}^{k} A_j^{(i)} X_{n^*} \|^{-(k-1)r}) \\ & + (K_2 V^{'}(X_{n^*}) + K_3) O(\| \prod_{j=1}^{n^*} A_{i_j} x \|^{-(n^*-1)r}) \\ & + K_4 V^{'}[h(X_{n^*})] + K_5. \end{aligned} \tag{5.17}$$

Since $h(x) = O(\| x \|^{\epsilon})$, Jensen's Inequality has $E_x[(V^{'}[h(X_{n^*})])^{\epsilon}] \leq K_6(E_x[V^{'}(X_{n^*})])^{\epsilon}$ for some $K_6 < \infty$. Using this and (5.17),

$$\limsup_{E_x[V^{'}(X_{n^*})] \to \infty} \frac{E_x[E(V^{'}(X_{n^*+k})|X_{n^*})]}{E_x[V^{'}(X_{n^*})]} < 1. \tag{5.18}$$

Let $V(x) := 1 + E_x[V^{'}(X_{n^*})]$. Then $V \geq 1$ satisfies (2.2), is locally bounded, unbounded and measurable. The level sets are petite by Lemma 5; by Lemma 2, the process $\{X_t\}$ is V-uniformly ergodic.

(2) Suppose w.l.o.g. $\mathbb{C}_1$ has length k_1 and has $\rho := \rho(\prod_{j=1}^{k_1} A_j^{(1)}) > 1$, and that $X_0 = x \in R_1^{(1)} \in \mathbb{C}_1$ with $\| x \| > M := \max(M_1, M_2, M_3)$. Define $a_1(x), I_{k_1}, I_{k_1}^C, I_g, I_g^C$ as in the proof of Theorem 1. Pick $q < 1/2 \min(r, [k_1 - 1]\epsilon)$.

By (3.4) we can write, when $X_j \in R_{j+1}^{(1)}, j = 1, \ldots, k_1$,

$$X_{k_1} = X_{k_1}^{'} + [\sum_{j=1}^{k_1} \prod_{s=j+1}^{k_1} A_s^{(1)} h(x) + \sum_{j=1}^{k_1} \prod_{s=j+2}^{k_1} A_s^{(1)} h(X_j)], \tag{5.19}$$

where $\{X_t^{'}\}$ is defined by (1.2). By the assumptions on h,

$$\phi(|a_1(\sum_{j=1}^{k_1} \prod_{s=j+1}^{k_1} A_s^{(1)} h(x))|) = O([\phi(|a_1(x)|)]^{(k_1-1)\epsilon}),$$

$$\phi(|a_1(\sum_{j=1}^{k_1} \prod_{s=j+1}^{k_1} A_s^{(1)} h(X_j))|) = O_p([\phi(|a_1(x)|)]^{(k_1-1)\epsilon}).$$

We have similar to (5.15) by (5.19) that when $|a_1(x)|$ is large,

$$E_x(I_{k_1} I_g) < P_x(I_g = 1 | I_{k_1} = 1) = O([\phi(|a_1(x)|)]^{-r}), \tag{5.20}$$

Define the test function $V(x) = (1+[\phi(|a_1(x)|)]^q)$. Then by (5.20) and arguments similar to those leading to (5.16),

$$\limsup_{V(x)\to\infty} E_x\Big[\frac{V(x)}{V(X_{k_1})}\Big] = \limsup_{|a_1(x)|\to\infty} E_x\Big[\frac{(1+[\phi(|a_1(x)|)]^q)}{(1+[\phi(|a_1(X_{k_1})|)]^q)}\Big] \quad < \quad 1. \tag{5.21}$$

By assumption $\{X_t\}$ is aperiodic and ψ-irreducible, with ψ being Lebesgue measure. Clearly $\psi(\{x : V(x) > M\}) > 0$ for all $M < \infty$. By Lemma 2, $\{X_t\}$ is transient. □

References

Chan, K. S., Petruccelli, J. D., Tong, H. and Woolford, S.W. (1985). A multiple threshold AR(1) model. *J. Appl. Prob.* **22**, 267-279.

Chen, R. and Tsay, S. R. (1991). On the ergodicity of TAR(1) processes. *Ann. Appl. Prob.* **1**, 613-634.

Ciarlet, P. G. (1982). *Introduction a la Analyse Numerique Matricielle et a l'Optimisation.* Masson, Paris.

Cline, D. B. H. and Pu, H. H. (1998). Verifying irreducibility and continuity of a nonlinear time series. *Statist. Probab. Lett.* **40**, 139-148.

Cline, D. B. H. and Pu, H. H. (1999). Stability of nonlinear AR(1) time series with delay. *Stochastic Process. Appl.* **82**, 307-333.

Cline, D. B. H. and Pu, H. H. (2001a). Geometric transience of nonlinear time series. *Statist. Sinica* **11**, 273-287.

Guo, M. and Petruccelli, J. D. (1991). On the null recurrence and transience of a first order SETAR model. *J. Appl. Prob.* **28**, 584-592.

Kunitomo, N. (2001). On ergodicity of some TAR(2) processes. Dept. of Economics, University of Tokyo, Japan.

Lewis, P. A. W. and Stevens, J. G. (1991). Nonlinear modeling of time series using multivariate adaptive regression splines (MARS). *J. of Amer. Stat. Assoc.* **86**, 864-877.

Liebscher, Eckhard (2005). Towards a unified approach for proving geometric ergodicity and mixing properties of nonlinear autoregressive processes. *J. Time Ser. Anal.* **26**, 669-689.

Lim, K. S. (1992). On the stability of a threshold AR(1) without intercepts. *J. Time Ser. Anal.* **13**, 119-132.

Martelli, M. (1992). *Discrete Dynamical Systems and Chaos.* J. Wiley and Sons, New York.

Meyn, S. P. and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability.* Springer-Verlag, London.

Nummelin, Esa (1984). *General Irreducible Markov Chains and Non-Negative Operators.* Cambridge University Press.

Petruccelli, J. D. and Woolford, S. W. (1984). A threshold AR(1) model. *J. Appl. Prob.* 21, 270-286.

Tjostheim, D. (1990). Non-linear time series and Markov chains. *Adv. Appl. Probab.* **22**, 587-611.

Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society B* **42**, 245-292.

Tong, H. (1990). *Non-linear Time Series Analysis: A Dynamical System Approach.* Oxford University Press, London.

Tweedie, R. L. (1976). Criteria for classifying general Markov chains. *Adv. Appl. Probab.* **8**, 737-771.

Department of Mathematics, Plymouth State University, Plymouth NH 03264, U.S.A.

E-mail: tboucher1@plymouth.edu

Department of Statistics, Texas A&M University, College Station TX 77843, U.S.A.

E-mail: dcline@stat.tamu.edu