Stochastic Processes and their Applications 82 (1999) 307–333

stochastic processes and their applications

www.elsevier.com/locate/spa

# Stability of nonlinear AR(1) time series with delay

Daren B.H. Cline *, Huay-min H. Pu

*Department of Statistics, Texas A&M University, College Station TX 77843, USA*

Received 22 April 1998; received in revised form 7 April 1999; accepted 9 April 1999

**Abstract**

The stability of generally defined nonlinear time series is of interest as nonparametric and other nonlinear methods are used more and more to fit time series. We provide sufficient conditions for stability or nonstability of general nonlinear AR(1) models having delay $d \geqslant 1$. Our results include conditions for each of the following modes of the associated Markov chain: geometric ergodicity, ergodicity, null recurrence, transience and geometric transience. The conditions are sharp for threshold-like models and they characterize parametric threshold AR(1) models with delay. 

*Keywords:* Nonlinear time series; Ergodicity; Transience

## 1. Introduction

In this paper we are interested in the stability of the first-order nonlinear time series model with delay lag $d \geqslant 1$. Specifically, the model is

$$\xi_t = \varphi(\xi_{t-1},\dots,\xi_{t-d})\xi_{t-1}+\vartheta(\xi_{t-1},\dots,\xi_{t-d})+c(e_t;\xi_{t-1},\dots,\xi_{t-d}), \quad t \geqslant 1, \tag{1.1}$$

where $\varphi$ and $\vartheta$ are bounded and measurable, $c$ is measurable and $\{e_t\}$ is an iid sequence of random variables, independent of the initial values $\xi_{1-d},\dots,\xi_0$. Examples studied in the literature include first-order threshold models with $\varphi$ depending only on the sign of $\xi_{t-d}$ (Chen and Tsay, 1991; Lim, 1992) and first-order amplitude-dependent exponential autoregressive (EXPAR) processes where $\varphi(\xi_{t-1},\dots,\xi_{t-d}) = \alpha + \beta e^{-\delta\xi_{t-d}^2}$ (cf. for example, Tong, 1990). However, $\varphi$ could be defined more generally than either of these.

As models such as (1.1) are being fit to nonlinear autoregressive time series, understanding their stability has become increasingly important. (cf. Chen and Tsay, 1993a,b; Tjøstheim and Auestad, 1994a,b; Härdle et al., 1997). In a series of work, Chan (1990,1993) and Chan and Tong (1985,1994) pioneered the study of the stability of general nonlinear time series, applying well known drift conditions for Markov chains. (See also Tong, 1990). In particular, they identified stability conditions for models

* Corresponding author. Tel.: +1-409-845-3181; fax: +1-409-845-3144.
*E-mail address:* dcline@stat.tamu.edu (D.B.H. Cline)

0304-4149/99/$ - see front matter 
PII: S0304-4149(99)00042-3

in which the autoregression function is Lipschitz continuous. Others (Chan and Tong, 1986; Chen and Tsay, 1993a; Guégan and Diebolt, 1994; An and Huang, 1996; Lu, 1996) have identified conditions without the continuity assumptions but the conditions can be fairly strong when either the autoregression order $p$ or the delay lag $d$ is greater than 1.

These efforts either do not include or do not characterize models with discontinuous regression functions, such as parametric threshold models for which either $p$ or $d$ is greater than 1. The only threshold models which have been characterized are the self-exciting threshold autoregression (SETAR) model of order 1 and no delay (Petrucelli and Woolford, 1984; Chan et al., 1985; Guo and Petrucelli, 1991) and the simplest threshold models of order 1 and delay $d > 1$ (Chen and Tsay, 1991; Lim, 1992). This paper will characterize the stability of more general threshold models with order 1 and delay $d > 1$.

The "coefficient function" $\varphi(x)$ in (1.1) and the "intercept function" $\vartheta(x)$ can be construed either nonparametrically or parametrically. For example, the otherwise nonparametric model (1.1) can be a partially parameterized "threshold-like" model as follows. Let $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ and $u = (u_1, \dots, u_d) \in \mathbb{U} = \{1, -1\}^d$ and suppose there exist

$$\phi_u = \lim_{\min_i u_i x_i \to \infty} \varphi(x), \quad \theta_u = \lim_{\min_i u_i x_i \to \infty} \vartheta(x) \quad \text{for all} \quad u \in \mathbb{U}. \tag{1.2}$$

The EXPAR models, for example, are threshold like. A fully parametric threshold model would have

$$\varphi(x) = \phi_u, \quad \vartheta(x) = \theta_u \quad \text{for} \quad u_i x_i > 0, \quad u \in \mathbb{U}. \tag{1.3}$$

The partially parametric model (1.2) is quite general because it makes no assumptions about $\varphi(x)$ near the thresholds (i.e., when one component of $x$ is "small"). In this paper we provide sharp conditions for stability of the partially parameterized model, the characterization being complete for the fully parameterized model. These results are simple consequences of our conditions for stability and nonstability of the nonparametric model.

The results are presented in Sections 2 and 3. In Section 2 we provide conditions for the Markov chain associated with $\{\xi_t\}$ to be geometrically ergodic and conditions for it to be geometrically transient. Such conditions do not rely heavily on the intercept term $\vartheta(x)$ or on the error term $c(e_1; x)$. In Section 3 we look at more refined conditions for ergodicity and transience, as well as for null recurrence. These conditions are much more sensitive to $\vartheta(x)$ and $c(e_1; x)$. Examples are provided and the proofs are given in Section 4.

## 2. Conditions for geometric ergodicity and geometric transience

The Markov chain associated with the autoregression process in (1.1) is

$$X_t = (\xi_t, \xi_{t-1}, \dots, \xi_{t-d+1}). \tag{2.1}$$

Stability of the time series (1.1) is determined in terms of Harris recurrence of (2.1). By *ergodicity* we mean positive Harris recurrence, as we will assume aperiodicity and irreducibility throughout. Conditional probability and expectation, given the initial state of the chain are denoted as $P_x(\cdot)=\boldsymbol{P}(\cdot|X_0=x)$ and $E_x(\cdot)=\boldsymbol{E}(\cdot|X_0=x)$, respectively.

In this section we provide conditions for $\{X_t\}$ to be geometrically ergodic or geometrically transient. By the latter we mean there is a positive probability the process will grow geometrically fast, for any initial state. The conditions are stronger than those for simply proving ergodicity or transience but much less can be assumed about the error term. In particular, the results in this section can be applied to the $\beta$-ARCH models of Guégan and Diebolt (1994). These models are characterized by $|c(v;x)|\leqslant K(1+||x||^{\beta}|v|)$ for some $K<\infty$ and $0<\beta<1$.

We first identify conditions applicable to general nonlinear models satisfying (1.1) and then adapt them to nonparametric models exhibiting cyclic behavior. We will apply the results to obtain sharp conditions for the partially parametric model given by (1.2) and compare our results to those for similar cyclical models in the literature. The last result of the section does not assume the cyclic behavior but it does assume more smoothness.

Define $a(x)=\varphi(x)x_1$ for $x=(x_1,\ldots,x_d)\in\mathbb{R}^d$. We will assume throughout that $\varphi(x)$ and $\vartheta(x)$ are bounded and that, for each $x\in\mathbb{R}^d$, $c(e_t;x_1,\ldots,x_d)$ has a lower semicontinuous density positive everywhere on $\mathbb{R}$. These conditions ensure that $\{X_t\}$ is aperiodic and $\psi$-irreducible with Lebesgue measure ($\mu_d$) as the irreducibility measure (Cline and Pu, 1998b; cf. Meyn and Tweedie, 1993 for definitions and related conditions). We also assume $\{X_t\}$ is a $T$-chain (again, cf. Meyn and Tweedie, 1993). This is so, for example, if $c(e_1;x)=b(x)e_1$ where both $b$ and the density of $e_1$ are locally bounded away from 0 (Cline and Pu, 1998b).

In addition, our results in this section will refer to the following assumptions about the error term $c(e_1;x)$. Note that the assumptions hold trivially for additive errors or even when $\{|c(e_1;x)|^r\}$ is uniformly integrable for some $r>0$. The assumptions also hold for the $\beta$-ARCH models of Guégan and Diebolt (1994) when $\beta<1$.

**Assumptions.** Let $r>0$ and let $||\cdot||$ be a norm defined on $\mathbb{R}^d$.

(A.1) For each $M<\infty$, $\sup_{||x||\leqslant M}\boldsymbol{E}(|c(e_1;x)|^r)<\infty$ and

$$\limsup_{||x||\to\infty}\frac{\boldsymbol{E}(|c(e_1;x)|^r)}{||x||^r}=0.$$

(A.2) The function $s(x)=\min(|a(x)|,|x_1|,\ldots,|x_d|)$ is unbounded on $\mathbb{R}^d$ and for each $\delta>0$,

$$\limsup_{s(x)\to\infty}s^r(x)\boldsymbol{P}(|c(e_1;x)|>\delta|a(x)|)=0.$$

(A.3) $|c(e_1;x)|/|a(x)|\to 0$ in probability, as $|a(x)|\to\infty,\min_i|x_i|\to\infty$.

Our results rely on well-known drift conditions for Markov chains, involving carefully crafted test functions. We start by establishing a general condition for geometric ergodicity of the Markov chain (2.1). Note that by definition $X_1=(\xi_1,x_1,\ldots,x_{d-1})$ when $X_0=x=(x_1,\ldots,x_d)$. Define $s(x)=\min(|a(x)|,|x_1|,\ldots,|x_d|)$.

**Theorem 2.1.** *Assume* (A.1). *If* $s(x)$ *is bounded or if there exists* $\lambda$: $\mathbb{R}^d \to (0,\infty)$, *bounded and bounded away from* 0, *and* $M < \infty$ *such that*

$$\limsup_{\substack{|a(x)| \to \infty \\ \min_i |x_i| \to \infty}} E_x\left(\frac{\lambda(X_1)}{\lambda(x)} |\varphi(X_1)|^r 1_{|\xi_1|>M, \operatorname{sgn}(\xi_1)=\operatorname{sgn}(a(x))}\right) < 1, \tag{2.2}$$

*then* $\{X_t\}$ *is geometrically ergodic.*

**Remark.** An obvious condition for geometric ergodicity is $\limsup_{||x|| \to \infty} |\varphi(x)| < 1$, which ensures the process shrinks anytime it becomes too large. This, in fact, is the condition one gets when applying known results for general autoregressive models of order $p$ to the order 1 model with delay $d$ (e.g., Chan and Tong, 1986; Chen and Tsay, 1993a; Guégan and Diebolt, 1994; An and Huang, 1996). Many stable nonlinear time series do not have this trait, however, and instead one need only have that $|\varphi(X_{t+m}) \cdots \varphi(X_{t+1})|$ is small, in some average sense and for some $m$, when $\xi_t$ is large. In fact, by Cline and Pu (1999, Lemma 4.1), (2.2) is equivalent to

$$\limsup_{\substack{|a(x)| \to \infty \\ \min_i |x_i| \to \infty}} E_x\left(\prod_{j=1}^{m} |\varphi(X_j)|^r 1_{|\xi_j|>M}\right) < 1 \quad \text{for some } m \geqslant 1.$$

The indicator $1_{|\xi_1|>M, \operatorname{sgn}(\xi_1)=\operatorname{sgn}(a(x))}$ within the expectation in (2.2) looks a bit cumbersome but it makes it possible to restrict consideration of the behavior of $\varphi$ to a suitable subset of $\mathbb{R}^d$.

Next, we have a general condition for geometric transience of the process (2.1).

**Theorem 2.2.** *Assume* (A.2). *If there exists* $\lambda$: $\mathbb{R}^d \to [0,\infty)$, *bounded, and* $R \subset \mathbb{R}^d$ *such that* $\mu_d(\{x \in R$: $\lambda(x)|a(x)|^r > M, \min_i |x_i| > M\}) > 0$ *for every* $M < \infty$ *and*

$$\limsup_{s(x) \to \infty, x \in R} s^r(x) P_x(X_1 \notin R) = 0 \tag{2.3}$$

*and if there exists* $M < \infty$ *and* $q < 1$ *such that*

$$\limsup_{\substack{\lambda(x)|a(x)|^r \to \infty \\ \min_i |x_i| \to \infty, x \in R}} E_x\left(\frac{\lambda(x)|a(x)|^r}{1 + \lambda(X_1)|\varphi(X_1)a(x)|^r} 1_{|\xi_1|>M, \operatorname{sgn}(\xi_1)=\operatorname{sgn}(a(x)), X_1 \in R}\right) < q^r, \tag{2.4}$$

*then* $\{X_t\}$ *is transient and* $P_x(\lim_{t\to\infty} q^t |\xi_t| \to \infty) > 0$ *for all* $x \in \mathbb{R}^d$.

**Remark.** Again, the indicator variable in (2.4) is designed to make application of the theorem easier, despite appearances to the contrary.

The assumption that $\mu_d(\{x \in R$: $\lambda(x)|a(x)|^r > M, \min_i |x_i| > M\}) > 0$ for every $M < \infty$ is valid, for example, if $\mu_d(R) > 0$ and $\inf_{x \in R, \min_i |x_i| > M} \lambda(x)|a(x)|^r \to \infty$ as $M \to \infty$. In particular, the latter will hold if $R$ and $\lambda$ are chosen so that $\lambda(x)|\varphi(x)|^r$ is bounded away from 0 on $R$.

We now turn our attention to models with cyclic behavior, of which the fully parametric model (1.3) is an example: If $\xi_t, \ldots, \xi_{t-d+1}$ are large then $\operatorname{sgn}(\xi_{t+1})$ is nearly certain to be the same as $\mathrm{sgn}(\phi_{\mathrm{sgn}(X_t)}\xi_t)$ (where $\mathrm{sgn}(x)$ is taken componentwise). Thus, the process $\mathrm{sgn}(\xi_t)$ will tend to follow cycles determined by the signs of the parameters $\phi_u$, at least as long as $\xi_t$ remains large. In fact the partially parameterized model (1.2) has this characteristic as well and it takes only a little imagination to see that many nonparametric models do also. We therefore identify this kind of cyclic behavior for a general process.

Recall we have defined $\mathbb{U} = \{-1, 1\}^d$. For each $u \in \mathbb{U}$ and $M < \infty$ we also define

$$Q_{u,M} = \{x \in \mathbb{R}^d\colon\ u_i x_i > M,\ i = 1, \ldots, d\}.$$

**Assumption**

(A.4) For some $M < \infty$ and each $u \in \mathbb{U}$, either $\varphi(x) \geqslant 0$ for all $x \in Q_{u,M}$ or $\varphi(x) \leqslant 0$ for all $x \in Q_{u,M}$.

Suppose Assumption (A.4) is valid and let

$$Q^*_{u,M} = \{x \in Q_{u,M}\colon\ |\varphi(x)x_1| > M\}.$$

If $Q^*_{u,M}$ is not empty then there exists $u^* \in \mathbb{U}$ such that $(\varphi(x)x_1, x_1, \ldots, x_{d-1}) \in Q_{u^*,M}$ for all $x \in Q^*_{u,M}$. We call $u^*$ the *successor* of $u$ if $Q^*_{u,M}$ is nonempty for all $M < \infty$. If $|\varphi(x)x_1|$ is bounded on $x \in Q_{u,M}$ for some $M < \infty$ then we say $u$ has no successor. Therefore, every $u \in \mathbb{U}$ satisfies exactly one of the following:

(i) $u$ has no successor,

(ii) $u$ is in a cycle $C = \{u^{(1)}, \ldots, u^{(k)}\}$ where $u^{(j)}$ succeeds $u^{(j-1)}$ for $j = 2, \ldots, k$ and $u^{(1)}$ succeeds $u^{(k)}$,

(iii) $u$ has a successor but $u$ is not in a cycle.

**Example 2.1.** Consider the case $d = 2$, $\mathbb{U} = \{-1, 1\}^2$ and

$$\varphi(x) = \left(\sum_{u \in \mathbb{U}} \phi_u 1_{\mathrm{sgn}(x)=u}\right) p(x),$$

where $p(x)$ is nonnegative and $\mathrm{sgn}(x) = (\mathrm{sgn}(x_1), \mathrm{sgn}(x_2))$. If $\phi_{1,-1} > 0$, $\phi_{1,1} < 0$, $\phi_{-1,-1} < 0$ and $\phi_{-1,1} > 0$ then there is only one cycle, $\{(1, 1), (-1, 1), (-1, -1), (1, -1)\}$. If $\phi_{1,-1} < 0$, $\phi_{1,1} > 0$, $\phi_{-1,-1} = 0$ and $\phi_{-1,1} < 0$ then there are two cycles, $\{(1, 1)\}$ and $\{(1, -1), (-1, 1)\}$, and $(-1, -1)$ has no successor. Several other cases are discussed in examples below and the rest are left to the reader.

We denote the class of cycles with $\mathscr{C}$. When large, the time series $\{\xi_t\}$ behaves as if $\mathrm{sgn}(X_t)$ follows the rules of succession outlined above. The time series will be unstable (i.e., grow in magnitude) only if there is a cycle for which the effect on $\xi_t$ in a complete circuit of that cycle is to make $\xi_t$ grow. With this in mind we have the following definitions and theorem.

For $u \in \mathbb{U}$ with successor $u^*$ and fixed $r > 0$ define

$$\Phi_u = \limsup_{\substack{|a(x)| \to \infty \\ \min_i u_i x_i \to \infty}} E_x(|\varphi(X_1)|^r 1_{X_1 \in Q_{u^*,M}})$$

and

$$\Phi'_u = \liminf_{\substack{|a(x)| \to \infty \\ \min_i u_i x_i \to \infty}} (E_x(|\varphi(X_1)|^{-r} 1_{X_1 \in Q_{u^*,M}}))^{-1}.$$

**Remark.** If $u$ has successor $u^* = (u_1^*, \dots, u_d^*)$ then

$$\Phi_u \leqslant \limsup_{\min_i u_i^* x_i \to \infty} |\varphi(x)|^r \quad \text{and} \quad \Phi'_u \geqslant \liminf_{\min_i u_i^* x_i \to \infty} |\varphi(x)|^r.$$

Bhattacharya and Lee (1995) use the limits on the right to prove results for general first-order models with no delay. In practice, the bounds we use are generally better but harder to compute. The bounds coincide in the partially parametrized model (1.2) considered in Corollary 2.4 below.

We now provide conditions for a general first-order, cyclical model with delay.

**Theorem 2.3.** *Let* $r > 0$. *Assume* (A.4) *and define the cycles* $C \in \mathscr{C}$ *and the limits* $\Phi_u$, $\Phi'_u$ *according to the discussion following* (A.4).

(i) *Assume* (A.1). *If* $\mathscr{C}$ *is empty or*

$$\max_{C \in \mathscr{C}} \prod_{u \in C} \Phi_u < 1,$$

*then* $\{X_t\}$ *is geometrically ergodic.*

(ii) *Assume Assumption* (A.2). *If there is a cycle* $C$ *and* $q < 1$ *such that for some* $u \in C$,

$$\mu_d \left( \left\{ x\colon \min_i u_i x_i > M, |a(x)| > M \right\} \right) > 0 \quad \textit{for all } M < \infty$$

*and*

$$\prod_{u \in C} \Phi'_u > q^{-r}, \tag{2.5}$$

*then* $\{X_t\}$ *is transient and* $P_x(\lim_{t \to \infty} q^t |\xi_t| = \infty) > 0$ *for all* $x \in \mathbb{R}^d$.

**Remark.** For the proof of the geometric ergodicity result Theorem 2.3(i), $\vartheta(x)$ need not be bounded, but $\lim_{||x|| \to \infty} \vartheta(x)/||x|| = 0$ is required. Also, we may weaken (A.4) by assuming there exists $\varphi^* : \mathbb{R}^d \to \mathbb{R}$ satisfying the condition given for $\varphi$ in (A.4) and

$$\limsup_{\min_i |x_i| \to \infty} |\varphi^*(x) - \varphi(x)| = 0.$$

This is possible since $(\varphi(x) - \varphi^*(x))x_1$ may be absorbed into $\vartheta(x)$, without changing the assumptions.

The proof of the geometric transience result, part (ii) of Theorem 2.3, requires $\varphi(x)x_1$ to be locally bounded but it does not actually require $\varphi(x)$ to be bounded.

**Corollary 2.4.** *Assume partially parametric model* (1.2) (*and hence* (A.4)).

(i) *Assume Assumption* (A.1). *If* $\mathscr{C}$ *is empty or*

$$\max_{C\in\mathscr{C}} \prod_{u\in C} \phi_u < 1,$$

*then* $\{X_t\}$ *is geometrically ergodic.*

(ii) *Assume Assumption* (A.2). *If there exists* $q < 1$ *such that*

$$\max_{C\in\mathscr{C}} \prod_{u\in C} \phi_u > \frac{1}{q},$$

*then* $\{X_t\}$ *is transient and* $P_x(\lim_{t\to\infty} q^t|\xi_t| = \infty) > 0$ *for all* $x \in \mathbb{R}^d$.

**Example 2.1** [*cont.*] Suppose $p(x) \to 1$ as $\min(|x_1|, |x_2|) \to \infty$ in Example 2.1 above. It is not difficult to enumerate all the cases and to determine sharp conditions. Specifically, let

$$\rho = \max(\phi_{-1,-1}, \phi_{1,1}, \min(\phi_{-1,1}, \phi_{-1,1}\phi_{-1,-1}, 0)\min(\phi_{1,-1}, \phi_{1,-1}\phi_{1,1}, 0)).$$

Then geometric ergodicity occurs if $\rho < 1$ and geometric transience occurs if $\rho > 1$.

**Example 2.2** (cf. Chen and Tsay, 1991; Lim, 1992). Consider the simple two-parameter TAR(1) model with delay $d > 1$ defined by

$$\xi_t = \begin{cases} \phi_1\xi_{t-1} + e_t & \text{if } \xi_{t-d} \leqslant 0, \\ \phi_2\xi_{t-1} + e_t & \text{if } \xi_{t-d} > 0. \end{cases}$$

Using different algebraic methods, the above-mentioned authors have shown that the precise stability conditions are as we have described in Corollary 2.4(i). Specifically, if $\phi_1\phi_2 \geqslant 0$ the condition is the same as the well-known condition for a TAR(1) model with no delay ($d = 1$): $\max(\phi_1, \phi_2, \phi_1\phi_2) < 1$. But if $\phi_1\phi_2 < 0$ the condition is $\max(\phi_1^{s_d}\phi_2^{t_d}, \phi_1^{t_d}\phi_2^{s_d}) < 1$ where $s_d$ and $t_d$ are integers (respectively, odd and even) that the authors have computed and tabulated for $d = 1, \ldots, 27$ by finding all the possible cycles.

By our results, the same conditions apply for partially parametrized models as well. This would include order 1 models similar to EXPAR models where, for example, $\varphi(x) = \phi_1 + (\phi_2 - \phi_1)G(x_d)$ and $G$ is a univariate distribution function. Also, our results show that the model may include an intercept term and nonadditive errors.

**Remark.** Tjøstheim (1990, Theorem 4.5)) considers parametric $d$-dimensional threshold processes with coefficients constrained so that the process, when large, follows a single cycle from one region to another. Our methods could be used to generalize his results to cases with multiple cycles.

Not all models have the cyclical behavior of Assumption (A.4). The final result for this section provides an alternative condition for geometric ergodicity. For $x \in \mathbb{R}^d$, let $\alpha_1(x) = (a(x), x_1, \ldots, x_{d-1})$ and $\alpha_j(x) = \alpha_1(\alpha_{j-1}(x))$ for $j \geqslant 2$. Also, let $\varphi_1(x) = \varphi(x)$ and $\varphi_j(x) = \varphi(\alpha_{j-1}(x)) = \varphi_{j-1}(\alpha_1(x))$ for $j \geqslant 2$. A number of results (e.g., Chan and Tong (1985,1994)) state that geometric ergodicity holds when the dynamical system

$\{\alpha_n(x)\}$ is exponentially stable. The results usually also require smoothness of $\alpha_1$ such as Lipschitz continuity. As our result shows, it suffices to consider the behavior of $\varphi$ on a more restricted set. We also define, for $x, y \in \mathbb{R}^d$,

$$s(x) = \min(|a(x)|, |x_1|, \ldots, |x_d|) \quad \text{and} \quad \eta(x, y) = \max_i \left( \frac{|x_i - y_i|}{|x_i|} \right).$$

**Theorem 2.5.** *Assume* (A.1) *and* (A.3). *If* $\varphi$ *is such that*

$$\lim_{\substack{s(x) \to \infty \\ \eta(x,y) \to 0}} |\varphi(x) - \varphi(y)| = 0 \tag{2.6}$$

*and*

$$\limsup_{\min_i |x_i| \to \infty} \prod_{j=1}^{n} |\varphi_j(x)| 1_{|\varphi_j(x) x_1| > M} < 1 \quad \textit{for some } n \geq 1,\ M < \infty, \tag{2.7}$$

*then* $\{X_t\}$ *is geometrically ergodic.*

**Remark.** Essentially, (2.7) is exponential stability of the dynamical system defined by $x_t = a(x_{t-1}) 1_{s(x_{t-1}) > M}$, while (2.6) weakens the continuity assumption. Both conditions depend on the values of $\varphi(x)$ only for $x$ such that $s(x)$ is arbitrarily large.

## 3. Conditions for ergodicity, transience and null recurrence

Weaker conditions are possible for both ergodicity and transience but they do not necessarily ensure geometric behavior of the process. Bhattacharya and Lee (1995) have provided such conditions for fairly general first order models with no delay. In this section we investigate conditions for ergodicity and transience of the model (1.1), as well as conditions for null recurrence. The behavior of $\vartheta(x)$ is crucial and stronger assumptions are required for $c(e_1; x)$. We continue to assume that the Markov chain $\{X_t\}$ is an aperiodic, $\mu_d$-irreducible $T$-chain. The results also assume some form of the following:

**Assumption**

(A.5) $\boldsymbol{E}(c(e_1; x)) = 0$ and for some $r \geq 1$, $\{|c(e_1; x)|^r\}_{x \in \mathbb{R}^d}$ is uniformly integrable.

Recall the cyclic behavior described in Assumption (A.4) and the definitions which follow it. Given a cycle, say $C = \{u^{(1)}, u^{(2)}, \ldots, u^{(k)}\}$, and associated constants, $\phi_{u^{(1)}}, \ldots, \phi_{u^{(k)}}$, we define

$$\lambda_{u^{(j)}} = \prod_{i=j}^{k} |\phi_{u^{(i)}}| \left( \prod_{i=1}^{k} |\phi_{u^{(i)}}| \right)^{j/k}. \tag{3.1}$$

Note that for all $u \in C$ with successor $u^*$,

$$\frac{\lambda_{u^*} |\phi_u|}{\lambda_u} = \left( \prod_{v \in C} |\phi_v| \right)^{1/k}. \tag{3.2}$$

The constants $\lambda_u$ play important roles both in the proofs of our results and in determining the part of the drift of the time series which is influenced by $\vartheta(x)$.

We first give conditions for stability and then, in Theorem 3.3, conditions for non-stability.

**Theorem 3.1.** *Assume Assumption* (A.4) *and make use of the definitions which follow it. Let $\phi_u$, $\theta_u$ be constants satisfying $\phi_u\varphi(x)\geqslant 0$ for $x\in Q_{u,M}$, $u\in C$, $C\in\mathscr{C}$, and*

$$0<\prod_{u\in C}\phi_u\leqslant 1 \quad \text{and} \quad \sum_{u\in C}\lambda_{u^*}\sigma_u\theta_u\leqslant 0 \quad \textit{for all } C\in\mathscr{C}, \tag{3.3}$$

*where the $\lambda_u$'s are given by* (3.1), *$\sigma_u=\operatorname{sgn}(\phi_u u_1)$ and $u^*$ is the successor to $u$.*

(i) *Assume Assumption* (A.5) *holds for some $r\geqslant 1$. If*

$$\limsup_{\substack{|a(x)|\to\infty\\ \min_i u_i x_i\to\infty}} (|\varphi(x)x_1+\vartheta(x)|-|\phi_u x_1+\theta_u|)|x_1|^s<0 \quad \textit{for all } u\in C,\ C\in\mathscr{C} \tag{3.4}$$

*for $s=0$ or for some $s\in(0,\min(1,r-1))$, then $\{X_t\}$ is ergodic.*

(ii) *Assume Assumption* (A.5) *holds with $r=2$ and*

$$\liminf_{\substack{|a(x)|\to\infty\\ \min_i u_i x_i\to\infty}} \boldsymbol{E}(c^2(e_1;x))>0 \quad \textit{for all } u\in C,\ C\in\mathscr{C}. \tag{3.5}$$

*If*

$$\limsup_{\substack{|a(x)|\to\infty\\ \min_i u_i x_i\to\infty}} (|\varphi(x)x_1+\vartheta(x)|-|\phi_u x_1+\theta_u|)|x_1|\leqslant 0 \quad \textit{for all } u\in C,\ C\in\mathscr{C}, \tag{3.6}$$

*then $\{X_t\}$ is Harris recurrent.*

**Example 3.1.** Suppose $d=2$, $\phi_{1,1}=1$, $\phi_{-1,-1}<0<\phi_{-1,1}$ and $\phi_{-1,1}\phi_{-1,-1}\phi_{1,-1}=1$. Then there are two cycles, $\{(1,1)\}$ and $\{(-1,1),(-1,-1),(1,-1)\}$, and the condition in Corollary 2.4 for geometric ergodicity is not met. However, if there exist $M<\infty$, $\varepsilon>0$ and either $s=0$ or $s\in(0,\min(1,r-1))$ such that

$$\varphi(x)x_1+\vartheta(x)\begin{cases}\leqslant x_1-\varepsilon|x_1|^{-s} & \text{if } x_1>M \text{ and } x_2>M,\\ \geqslant\phi_{-1,1}x_1+\theta_{-1,1}+\varepsilon|x_1|^{-s} & \text{if } x_1<-M \text{ and } x_2>M,\\ \leqslant\phi_{-1,-1}x_1+\theta_{-1,-1}-\varepsilon|x_1|^{-s} & \text{if } x_1<-M \text{ and } x_2<-M,\\ \geqslant\phi_{1,-1}x_1+\theta_{1,-1}+\varepsilon|x_1|^{-s} & \text{if } x_1>M \text{ and } x_2<-M,\end{cases}$$

where $-\phi_{-1,-1}\phi_{1,-1}\theta_{-1,1}-\phi_{1,-1}\theta_{-1,-1}-\theta_{1,-1}\leqslant 0$ then (3.3) and (3.4) are valid and the process is ergodic.

**Example 3.2.** (cf. Bhattacharya and Lee, 1995; Pu, 1995, Chapter V.) Suppose $d=1$ so the model has no delay. Suppose also $\vartheta(x)=0$, $c(e_1;x)=b(x)e_1$ where $b(x)$ is bounded and bounded away from 0, and $\boldsymbol{E}(e_1)=0$. Under these assumptions, it is possible to get slightly weaker conditions for ergodicity and sometimes to dispense with the cyclical assumption. For example, Pu shows that if $b(x)=1$ and $|x|^2(1-|\varphi(x)|)>\boldsymbol{E}(e_1^2)/2$ for all large $|x|$ then the process is ergodic. Without assuming a second moment, Bhattacharya and Lee show that if $|x|(1-|\varphi(x)|)>\varepsilon>0$ for all large $|x|$ then the process is ergodic. Bhattacharya and Lee also show if $\varepsilon\leqslant\varphi(\varphi(x)x)\varphi(x)|x|\leqslant|x|-c$ for some $\varepsilon > 0$ and a certain constant $c$ (defined by them but depending on $\varphi(x)$, $b(x)$ and the error distribution) and for all large $|x|$ then the process is ergodic. In this case, if $\varphi(x) < 0$, the time series cycles from positive to negative and back again, at least while it is large.

Such conditions are also possible for models with $d > 1$ if the error is additive, but they would involve the coefficients $\phi_u$ and $\theta_u$ and the error distribution in a complicated way.

**Remark.** Theorem 3.1(ii) shows, as did Lamperti (1960) for random walks, that a Markov chain may have a small drift away from the origin and still be recurrent.

Also, in certain cases the condition for ergodicity may in fact imply geometric ergodicity, even if $\prod_{u \in C} \phi_u = 1$ for some cycle $C$. We state this next. (See also Spieksma and Tweedie, 1994.)

**Corollary 3.2.** *Assume there exists* $\eta > 0$ *such that* $\sup_x \boldsymbol{E}(\mathrm{e}^{\eta|c(e_1;x)|}) < \infty$. *If the conditions of Theorem* 3.1(i) *hold with* $s = 0$ *then* $\{X_t\}$ *is geometrically ergodic.*

Now we turn to nonstability.

**Theorem 3.3.** *Assume Assumption* (A.4) *and make use of the definitions which follow it. For some cycle* $C$, *let* $\phi_u$, $\theta_u$ *be constants satisfying* $\phi_u \varphi(x) \geqslant 0$ *for* $x \in Q_{u,M}$, $u \in C$, *and*

$$\prod_{u \in C} \phi_u \geqslant 1 \quad \text{and} \quad \sum_{u \in C} \lambda_{u^*} \sigma_u \theta_u \geqslant 0, \tag{3.7}$$

*where the* $\lambda_u$*'s are given by* (3.1), $\sigma_u = \mathrm{sgn}(\phi_u u_1)$ *and* $u^*$ *is the successor to* $u$.

(i) *Assume Assumption* (A.5) *holds for some* $r > 1$. *If*

$$\liminf_{\min_i u_i x_i \to \infty} (|\varphi(x)x_1 + \vartheta(x)| - |\phi_u x_1 + \theta_u|)|x_1|^s > 0 \quad \text{for all } u \in C \tag{3.8}$$

*for some* $s \in (0, \min(1, r-1))$, *then* $\{X_t\}$ *is transient.*

(ii) *Assume Assumption* (A.5) *holds with* $r = 1$. *If there exists* $L_1, L_2, M < \infty$, *such that*

$$|\varphi(x)x_1 + \vartheta(x)| - |\phi_u x_1 + \theta_u| \geqslant -\Delta(x) \qquad \text{for all } x \in Q_{u,M},\ u \in C, \tag{3.9}$$

*where*

$$\Delta(x) = \int_{|\varphi(x)x_1 + \vartheta(x)|}^{\infty} (1 - F_x(y))\,\mathrm{d}y + L_1(1 - F_x(|\varphi(x)x_1 + \vartheta(x)| - L_2))$$

*and* $F_x$ *is the distribution of* $-\sigma_u c(e_1; x)$, *then* $\{X_t\}$ *is not positive recurrent.*

**Remark.** The condition in (3.9) allows nonpositive processes to have a slight negative drift but even for additive errors, and depending on the error distribution, $\Delta(x)$ may converge rapidly to 0 as $|x_1| \to \infty$. Either term in the definition of $\Delta(x)$ can dominate. Pu (1995, Chapter V) obtained similar results for TAR(1) models with no delay and additive errors.

The proof of Theorem 3.3 requires that we improve the drift condition for a Markov chain to be nonpositive (cf. Meyn and Tweedie, 1993, Theorem 11.5.1). The new condition is stated here.

**Lemma 3.4.** *Suppose $\{X_t\}$ is a homogeneous $\psi$-irreducible Markov chain on $\mathbb{R}^d$ and suppose $V:\mathbb{R}^d \to [0,\infty)$ vanishes on $R^c$. Assume there exist $K_1<\infty$, $K_2<\infty$ and function $w(x,y)$ such that*

(i) $V(X_1)-V(x) \geqslant w(x,X_1)-K_1 1_{X_1 \notin R}$, *if* $X_0=x\in R$,
(ii) $\boldsymbol{E}(|w(x,X_1)|\,|X_0=x)<K_2$ *and* $\boldsymbol{E}(w(x,X_1)|X_0=x)\geqslant 0$ *for all* $x\in R$, *and*
(iii) $\psi(\{x\in R\colon V(x)>K_1\})>0$ *and* $\psi(R^c)>0$.
*Then $\{X_t\}$ is not positive recurrent.*

The conditions in Theorems 3.1 and 3.3 are sharp for the partially parametric model of (1.2). With stronger assumptions they fully classify the model, as well as the fully parametric model (1.3).

**Corollary 3.5.** *Assume Assumption* (A.5) *holds with $r=2$. Assume there exists constants $\phi_u$, $\theta_u$ and $s>0$ such that*

$$\lim_{\min_i u_i x_i \to \infty} (\varphi(x)-\phi_u)|x_1|^{1+s}=0 \quad \textit{for all } u \textit{ in a cycle.} \tag{3.10}$$

*and*

$$\lim_{\min_i u_i x_i \to \infty} (\vartheta(x)-\theta_u)|x_1|^{s}=0 \quad \textit{for all } u \textit{ in a cycle.} \tag{3.11}$$

*Define $\lambda_u$ by* (3.1) *and $\sigma_u=\operatorname{sgn}(\phi_u u_1)$ for each $u$ in a cycle.*

(i) *If for every $C\in\mathscr{C}$ either $\prod_{u\in C}\phi_u<1$ or both $\prod_{u\in C}\phi_u=1$ and $\sum_{u\in C}\lambda_{u^*}\sigma_u\theta_u<0$ then $\{X_t\}$ is ergodic.*

(ii) *Assume* (3.10) and (3.11) *hold with $s=1$ and assume* (3.5) *holds. If for every $C\in\mathscr{C}$ both $\prod_{u\in C}\phi_u\leqslant 1$ and $\sum_{u\in C}\lambda_{u^*}\sigma_u\theta_u\leqslant 0$, and for some $C\in\mathscr{C}$* (3.9) *holds and both $\prod_{u\in C}\phi_u=1$ and $\sum_{u\in C}\lambda_{u^*}\sigma_u\theta_u=0$, then $\{X_t\}$ is null Harris recurrent.*

(iii) *If for some $C\in\mathscr{C}$ either $\prod_{u\in C}\phi_u>1$* or both *$\prod_{u\in C}\phi_u=1$ and $\sum_{u\in C}\lambda_{u^*}\sigma_u\theta_u>0$ then $\{X_t\}$ is transient.*

**Example 3.3.** Consider the model with $d=2$ and $\varphi$ as defined in Example 2.1 but suppose we also have

$$\vartheta(x)=\left(\sum_{u\in\mathbb{U}}\theta_u 1_{\operatorname{sgn}(x)=u}\right)q(x).$$

Assume $p(x)\to 1$ and $q(x)\to 1$ as $\min(x_1,x_2)\to\infty$ sufficiently fast to satisfy Corollary 3.5. Enumerating all the possibilities would be lengthy but, for example, suppose $\phi_{1,1}<0$, $\phi_{-1,1}<0$, $\phi_{1,-1}>0$ and $\phi_{-1,-1}<0$ so that there is one cycle $\{(1,1),(-1,1),(1,-1)\}$. If $\rho=\phi_{1,1}\phi_{-1,1}\phi_{1,-1}<1$ the process $\{(\xi_t,\xi_{t-1})\}$ is geometrically ergodic and it is geometrically transient if $\rho>1$. If, instead, $\rho=1$ then the process is ergodic, null recurrent or transient as $\phi_{-1,1}\phi_{1,-1}\theta_{1,1}+\phi_{1,-1}\theta_{-1,1}+\theta_{1,-1}$ is negative, zero or positive.

In practice, constructing the parameter space for ergodic models can be complicated. But checking the conditions is fairly straightforward once one has estimated values for the parameters. Even finding all the cycles is easily automated.

## 4. Proofs

**Proof of Theorem 2.1.** The proof consists of demonstrating the existence of a test function $V_1 : \mathbb{R}^d \to [0,\infty)$ with which we may apply the drift condition of Meyn and Tweedie (1993, Theorem 15.0.1). (See (4.12) and (4.13) below.) There is no loss here in assuming that $\vartheta$ is identically 0 as the mean of the error term $c(e_t;x)$ is inconsequential for this argument. Assume first that $s(x)$ is unbounded. Without any loss we assume $r \leqslant 1$ and is small enough that $\{|c(e_1;x)|^r/1+||x||^r\}$ is uniformly integrable. Let $V(x) = \lambda(x)|a(x)|^r + \sum_{i=1}^d \varepsilon^i |x_i|^r$, $\varepsilon \in (0,1)$ to be fixed later. Thus, given $X_0 = x$,

$$V(X_1) = (\lambda(X_1)|\varphi(X_1)|^r + \varepsilon)|a(x) + c(e_1;x)|^r + \sum_{i=1}^{d-1} \varepsilon^{i+1}|x_i|^r.$$

Choose $L < \infty$ so that $1/L < \lambda(x) < L$ and $|\varphi(x)| < L$ for all $x$. Then, by substituting in the definitions for $V(x)$ and $V(X_1)$ and using (A.1),

$$\begin{aligned}
&\limsup_{\substack{|a(x)|\to\infty \\ \min_i |x_i| \to \infty}} E_x\left(\frac{V(X_1)}{V(x)} 1_{|\xi_1|>M,\,\mathrm{sgn}(\xi_1)=\mathrm{sgn}(a(x))}\right) \\
&\quad \leqslant \limsup_{\substack{|a(x)|\to\infty \\ \min_i |x_i| \to \infty}} E_x\left(\frac{\lambda(X_1)}{\lambda(x)}|\varphi(X_1)|^r 1_{|\xi_1|>M,\,\mathrm{sgn}(\xi_1)=\mathrm{sgn}(a(x))}\right) + \varepsilon L
\end{aligned} \tag{4.1}$$

and

$$\limsup_{||x||\to\infty} E_x\left(\frac{V(X_1)}{V(x)}\right) \leqslant L(L^{1+r}+\varepsilon) < 2L^{2+r}. \tag{4.2}$$

We now fix $\delta$, $M$, $K$ and $\varepsilon$, in that order, so that according to (2.2), (4.1) and (4.2),

$$\sup_{\substack{|a(x)|>M \\ \min_i |x_i|>M}} E_x\left(\frac{V(X_1)}{V(x)} 1_{|\xi_1|>M,\,\mathrm{sgn}(\xi_1)=\mathrm{sgn}(a(x))}\right) < 1-\delta, \tag{4.3}$$

$$\sup_{|a(x)|>M} E_x\left(\frac{V(X_1)}{V(x)}\right) \leqslant 2(1-\delta)L^{2+r} \tag{4.4}$$

and

$$K > 2L^{2+r}, \qquad (L^{1+r}+1)\varepsilon K^d < \delta. \tag{4.5}$$

Define

$$v(x) = \begin{cases} 1 & \text{if } |x_1| \leqslant M, \\ K^{-i} & \text{if } |x_j| > M \text{ for } j \leqslant i \text{ and } |x_{i+1}| \leqslant M,\ i = 1,\dots,d-1, \\ K^{-d} & \text{if } |x_j| > M \text{ for } j \leqslant d, \end{cases} \tag{4.6}$$

so that

$$\frac{v(X_1)}{v(x)} \leqslant \begin{cases} 1 & \text{if } |\xi_1| > M, \quad \min_i |x_i| > M, \\ \frac{1}{K} & \text{if } |\xi_1| > M, \quad \min_i |x_i| \leqslant M, \\ K^d & \text{if } |\xi_1| \leqslant M. \end{cases}$$

Now define $V_1(x) = v(x)V(x)$. This is our test function. Note that by our choice of $r$, $\{V_1(X_1)/(1+V_1(x))\}$ is uniformly integrable and by Assumption (A.1),

$$\limsup_{\substack{||x|| \to \infty \\ |a(x)| \leqslant M}} P_x(|\xi_1|^r > \varepsilon V(x)) \leqslant \limsup_{||x|| \to \infty} \boldsymbol{P}(|c(e_1;x)|^r > \varepsilon V(x) - M^r) = 0.$$

Thus $(V_1(X_1)/V_1(x))1_{|\xi_1|^r > \varepsilon V(x)} \to 0$ in probability, as $||x|| \to \infty$, $|a(x)| \leqslant M$, and

$$\limsup_{\substack{||x|| \to \infty \\ |a(x)| \leqslant M}} E_x\left(\frac{V_1(X_1)}{V_1(x)} 1_{|\xi_1|^r > \max(M^r, \varepsilon V(x))}\right) = 0. \tag{4.7}$$

Likewise, $(V_1(X_1)/V_1(x))1_{|c(e_1;x)|^r > \varepsilon V(x)} \to 0$ in probability, as $||x|| \to \infty$, and

$$\limsup_{||x|| \to \infty} E_x\left(\frac{V_1(X_1)}{V_1(x)} 1_{|c(e_1;x)|^r > \varepsilon V(x)}\right) = 0. \tag{4.8}$$

Next, we note $|\xi_1|^r = |a(x) + c(e_1;x)|^r > \varepsilon V(x)$ implies that either $\mathrm{sgn}(\xi_1) = \mathrm{sgn}(a(x))$ or $|c(e_1;x)|^r > \varepsilon V(x)$. Thus, by (4.3) and (4.8),

$$\begin{aligned} &\limsup_{\substack{||x|| \to \infty \\ |a(x)| > M, \min_i |x_i| > M}} E_x\left(\frac{V_1(X_1)}{V_1(x)} 1_{|\xi_1|^r > \max(M^r, \varepsilon V(x))}\right) \\ &\quad \leqslant \sup_{\substack{|a(x)| > M \\ \min_i |x_i| > M}} E_x\left(\frac{V(X_1)}{V(x)} 1_{|\xi_1| > M, \mathrm{sgn}(\xi_1) = \mathrm{sgn}(a(x))}\right) \\ &\qquad + \limsup_{||x|| \to \infty} E_x\left(\frac{V_1(X_1)}{V_1(x)} 1_{|c(e_1;x)|^r > \varepsilon V(x)}\right) \\ &\quad < 1 - \delta. \end{aligned} \tag{4.9}$$

Also, by (4.4) and (4.5),

$$\sup_{\substack{|a(x)| > M \\ \min_i |x_i| \leqslant M}} E_x\left(\frac{V_1(X_1)}{V_1(x)} 1_{|\xi_1| > M}\right) \leqslant \frac{2(1-\delta)L^{2+r}}{K} < 1 - \delta. \tag{4.10}$$

Additionally, we may compute

$$\limsup_{||x|| \to \infty} E_x\left(\frac{V_1(X_1)}{V_1(x)} 1_{|\xi_1|^r \leqslant \max(M^r, \varepsilon V(x))}\right) \leqslant (L^{1+r} + 1)\varepsilon K^d < \delta. \tag{4.11}$$

Therefore, combining (4.7) and (4.9)–(4.11),

$$\limsup_{||x|| \to \infty} E_x\left(\frac{V_1(X_1)}{V_1(x)}\right) < 1. \tag{4.12}$$

Clearly, by Assumption (A.1) and the fact $K > 1$,

$$\sup_{||x|| \leqslant M_1} E_x(V_1(X_1)) \leqslant \sup_{||x|| \leqslant M_1} E_x(V(X_1)) < \infty \quad \text{for all } M_1 < \infty. \tag{4.13}$$

Since compact sets are petite (Meyn and Tweedie, 1993, Theorem 6.2.5) and (4.12) and (4.13) hold, geometric ergodicity follows from the drift condition of Meyn and Tweedie (1993, Theorem 15.0.1).

If $s(x)$ is bounded, the above argument is valid provided $M$ is chosen larger than $\sup s(x)$ and we use the convention that any supremum over the empty set has value 0. $\square$

**Proof of Theorem 2.2.** We demonstrate that there exists a test function $V: \mathbb{R}^d \to [0,\infty)$ with which we may apply the drift condition for geometric transience. That is, we will verify (4.18) below. As in the previous proof, we may assume $\vartheta$ is identically 0. Choose $L < \infty$ such that $\lambda(x) \leqslant L$ for all $x$ and choose $M < \infty$ and $\delta > 0$ such that, according to (2.4),

$$\sup_{\substack{\lambda(x)|a(x)|^r > M^* \\ \min_i |x_i| > M, x \in R}} E_x\left(\frac{\lambda(x)|a(x)|^r}{1+\lambda(X_1)|\varphi(X_1)a(x)|^r} 1_{|\xi_1|>M,\, \mathrm{sgn}(\xi_1)=\mathrm{sgn}(a(x)),\, X_1\in R}\right)$$
$$< (1-\delta)^{1+r} q^r, \tag{4.14}$$

where $M^* = (1-\delta)^{-r} L M^r$. Now choose $K > (1+(1-\delta)^{-r}L)/\delta q^r$ and define $Q_M = \{x \in \mathbb{R}^d\colon \min_{1\leqslant i\leqslant d}|x_i| > M\}$. The test function we use is

$$V(x) = \min(\lambda(x)|a(x)|^r, K|x_1|^r, \ldots, K^d|x_d|^r) 1_{Q_M}(x) 1_R(x).$$

By the definition of $V$ and Assumption (A.2),

$$\limsup_{V(x)\to\infty} E_x\left(\frac{V(x)}{1+V(X_1)} 1_{|c(e_1;x)|>\delta|a(x)|}\right)$$
$$\leqslant \limsup_{s(x)\to\infty} K^d s^r(x) P_x(|c(e_1;x)| > \delta|a(x)|) = 0. \tag{4.15}$$

Likewise, by (2.3),

$$\limsup_{V(x)\to\infty} E_x\left(\frac{V(x)}{1+V(X_1)} 1_{X_1\notin R}\right) \leqslant \limsup_{s(x)\to\infty} K^d s^r(x) P_x(X_1 \notin R) = 0. \tag{4.16}$$

If $x \in R$, $\lambda(x)|a(x)|^r > M^*$ and $\min_i |x_i| > M$ then, given $X_0 = x$, $|c(e_1;x)| \leqslant \delta|a(x)|$ implies $|\xi_1| > M$, $X_1 \in Q_M$ and

$$\frac{V(x)}{1+V(X_1)} 1_{|c(e_1;x)|\leqslant\delta|a(x)|,\, X_1\in R}$$
$$= \frac{V(x)}{1+V(X_1)} 1_{|c(e_1;x)|\leqslant\delta|a(x)|,\, X_1\in Q_M,\, X_1\in R}$$
$$\leqslant \frac{1}{K} + \frac{\lambda(x)|a(x)|^r}{1+\min(\lambda(X_1)|\varphi(X_1)|^r, K)|\xi_1|^r} 1_{|c(e_1;x)|\leqslant\delta|a(x)|,\, X_1\in Q_M,\, X_1\in R}$$
$$\leqslant \frac{1}{K} + (1-\delta)^{-r}\left(\frac{L}{K} + \frac{\lambda(x)|a(x)|^r}{1+\lambda(X_1)|\varphi(X_1)a(x)|^r} 1_{|\xi_1|>M,\, \mathrm{sgn}(\xi_1)=\mathrm{sgn}(a(x)),\, X_1\in R}\right).$$

Therefore, by (4.14) and the choice of $K$,

$$\limsup_{V(x)\to\infty} E_x\left(\frac{V(x)}{1+V(X_1)}1_{|c(e_1;x)|\leqslant\delta|a(x)|,X_1\in R}\right)$$
$$\leqslant\frac{1}{K}+(1-\delta)^{-r}\frac{L}{K}+(1-\delta)q^r<q^r. \tag{4.17}$$

Combining (4.15)–(4.17),

$$\limsup_{V(x)\to\infty} E_x\left(\frac{1+V(x)}{1+V(X_1)}\right)=\limsup_{V(x)\to\infty} E_x\left(\frac{V(x)}{1+V(X_1)}\right)<q^r<1. \tag{4.18}$$

Also, by assumption we have $\mu_d(\{x\colon V(x)>M\})>0$ for every $M<\infty$. Transience follows from Meyn and Tweedie (1993, Theorem 8.4.2). The conclusion of geometric transience follows from Cline and Pu (1998a, Lemma 4.1). □

**Proof of Theorem 2.3.** (i) The objective is to identify a function $\lambda(x)$ in order to apply Theorem 2.1. If $s(x)=\min(|a(x)|,|x_1|,\dots,|x_d|)$ is bounded then the conclusion follows immediately from Theorem 2.1. So we assume $s$ is not bounded. Let $\alpha(x)=(a(x),x_1,\dots,x_{d-1})$.

For some $\varepsilon>0$,

$$\max_{C\in\mathscr{C}}\prod_{u\in C}(\Phi_u+\varepsilon)<1. \tag{4.19}$$

(This is vaccuous if there are no cycles.) It suffices to choose $M$ large enough so that both Assumption (A.4) is valid and

$$\Phi_{u,M}=\sup_{x\in Q^*_{u,M}} E_x(|\varphi(X_1)|^r 1_{X_1\in Q_{u^*,M}})\leqslant\Phi_u+\varepsilon,$$

where $u^*$ is the successor of $u$ and $Q^*_{u,M}$ is defined following Assumption (A.4). Note that $x\in Q^*_{u,M}$ implies $\alpha(x)\in Q_{u^*,M}$. For any $u$ which has no successor, define

$$\Phi_{u,M}=\sup_{x\in Q_{u,M}} E_x(|\varphi(X_1)|^r),$$

which is bounded by assumption. Now for $x\in\mathbb{R}^d$, let

$$\lambda(x)=\begin{cases}\dfrac{\prod_{i=j}^k\Phi_{u^{(i)},M}}{\left(\prod_{i=1}^k\Phi_{u^{(i)},M}\right)^{-j/k}} & \text{if } x\in Q_{u^{(j)},M} \text{ and } u^{(j)} \text{ is in the cycle } \{u^{(1)},\dots,u^{(k)}\},\\ 1 & \text{if } \min_i |x_i|\leqslant M \text{ or } x\in Q_{u,M} \text{ and } u \text{ has no successor},\\ \lambda(\alpha(x))(\Phi_{u,M}+\varepsilon) & \text{if } x\in Q_{u,M} \text{ and } u \text{ has a successor but } u \text{ is not in a cycle.}\end{cases}$$

Note that the third part of the definition is recursive in that $\lambda(x)$ must first be defined for the cycle cases and for the cases with no successor and then in reverse order of succession for the cases where a successor exists but $u$ is not in a cycle. For $u$ in a cycle $C$, $\lambda$ is defined so that it is constant on $Q_{u,M}$ and if $k_C$ is the length of $C$ then

$$\frac{\lambda(\alpha(x))\Phi_{u,M}}{\lambda(x)} = \left(\prod_{v\in C}\Phi_{v,M}\right)^{1/k_C} < 1 \qquad \text{for all } x \in Q^*_{u,M}.$$

Indeed, from the definitions of $\lambda(x)$ and $\Phi_{u,M}$ and (4.19), it is now a simple matter to determine that for each $u \in \mathbb{U}$,

$$\sup_{x\in Q^*_{u,M}} \frac{\lambda(\alpha(x))}{\lambda(x)} E_x(|\varphi(X_1)|^r 1_{X_1\in Q_{u^*,M}}) < 1.$$

We have, therefore,

$$\begin{aligned}
&\limsup_{\substack{|a(x)|\to\infty\\ \min_i|x_i|\to\infty}} E_x\left(\frac{\lambda(X_1)}{\lambda(x)}|\varphi(X_1)|^r 1_{|\xi_1|>M,\,\mathrm{sgn}(\xi_1)=\mathrm{sgn}(a(x))}\right)\\
&\quad = \sup_{u\in\mathbb{U}} \limsup_{\substack{|a(x)|\to\infty\\ \min_i u_i x_i\to\infty}} E_x\left(\frac{\lambda(X_1)}{\lambda(x)}|\varphi(X_1)|^r 1_{X_1\in Q_{u^*,M}}\right)\\
&\quad \leqslant \sup_{u\in\mathbb{U}} \sup_{x\in Q^*_{u,M}} \frac{\lambda(\alpha(x))}{\lambda(x)} E_x(|\varphi(X_1)|^r 1_{X_1\in Q_{u^*,M}}) < 1.
\end{aligned} \tag{4.20}$$

So geometric ergodicity holds by Theorem 2.1.

(ii) Here we will find $\lambda(x)$ and $R$ to satisfy (2.4). Identify the cycle in (2.5) as $C = \{u^{(1)},\dots,u^{(k)}\}$. Define

$$\Phi'_{u,M} = \inf_{x\in Q^*_{u,M}} (E_x(|\varphi(X_1)|^{-r} 1_{X_1\in Q_{u^*,M}}))^{-1},$$

where it suffices to choose $M$ large enough so that $\prod_{j=1}^k \Phi'_{u^{(j)},M} > q^{-r}$. Let $R = \bigcup_{j=1}^k Q_{u^{(j)},M}$. From Assumption (A.2) it is easy to see that if $u$ has successor $u^*$ then

$$\limsup_{s(x)\to\infty,\, x\in Q_{u,M}} s^r(x) P_x(X_1 \notin Q_{u^*,M}) = 0$$

and therefore (2.3) holds.

Now define

$$\lambda(x) = \begin{cases} \prod_{i=j}^k \Phi'_{u^{(i)},M} \left(\prod_{i=1}^k \Phi'_{u^{(i)},M}\right)^{j/k} & \text{if } x \in Q_{u^{(j)},M},\ j = 1,\dots,k,\\ 0 & \text{if } x \notin R. \end{cases}$$

Thus, for each $u \in C$, $\lambda$ is constant on $Q_{u,M}$ and

$$\frac{\lambda(\alpha(x))\Phi'_{u,M}}{\lambda(x)} = \left(\prod_{v\in C}\Phi'_{v,M}\right)^{1/k} \qquad \text{for all } x \in Q^*_{u,M}.$$

Since $x \in Q_{u,M}$, $|\xi_1| > M$ and $\operatorname{sgn}(\xi_1) = \operatorname{sgn}(a(x))$ imply $X_1 \in Q_{u^*,M}$, it is simple to verify that (2.4) holds, analogous to verifying (4.20) above.

By (2.3) and the assumption $\mu_d(\{x: \min_i u_i x_i > M, |a(x)| > M\}) > 0$ for all $M < \infty$ and some $u \in C$ we also have

$$\mu_d(\{x \in R: \lambda(x)|a(x)|^r > M, \min_i |x_i| > M\}) > 0$$

for every $M < \infty$. The conclusion thus follows from Theorem 2.2. □

**Proof of Corollary 2.4.** (i) Assumption (A.4) follows from (1.2). We note that $\Phi_u \leqslant |\phi_{u^*}|^r$. Since the number of sign changes in a cycle must be even, the condition in Theorem 2.3(i) is satisfied.

(ii) This follows from Theorem 2.3(ii) in the same way that part (i) follows from Theorem 2.3(i). Note that

$$\mu_d\left(\left\{x:\ \min_i u_i x_i > M, |a(x)| > M\right\}\right) > 0$$

for each $u \in C$ and each $M < \infty$. □

**Proof of Theorem 2.5.** This is also a corollary to Theorem 2.1 and again the objective is to find an appropriate $\lambda(x)$. Let $\alpha_0(x) = x$ and $h_M(x) = |\varphi(x)|^r 1_{s(x)>M}$ where $r > 0$ satisfies Assumption (A.1). Condition (2.7) implies there exists $K < \infty$, $M < \infty$ and $\rho < 1$ such that

$$\prod_{j=0}^{n} h_M(\alpha_j(x)) \leqslant K\rho^n \quad \text{for all } n \text{ and all } x. \tag{4.21}$$

Let $\rho_1 \in (\rho, 1)$ and $\varepsilon \in (0, \rho_1^{-1} - 1)$ and define

$$\lambda(x) = \sup_{n \geqslant 1} \rho_1^{-n} \prod_{j=1}^{n} h_M(\alpha_j(x)) + \varepsilon. \tag{4.22}$$

Note that (4.21) implies there exists $n_0 \geqslant 1$ such that

$$\lambda(x) = \max_{1 \leqslant n \leqslant n_0} \rho_1^{-n} \prod_{j=1}^{n} h_M(\alpha_j(x)) + \varepsilon \qquad \text{for all } x. \tag{4.23}$$

Condition (2.6) implies, by way of induction,

$$\lim_{\substack{\min_i |x_i| \to \infty \\ \eta(x,y) \to 0}} \min\left(\min_{1 \leqslant j \leqslant n} |\varphi_j(x)|, \max_{1 \leqslant j \leqslant n} |\varphi_j(x) - \varphi_j(y)|\right) = 0 \qquad \text{for all } n \geqslant 1$$

and thus,

$$\lim_{\substack{\min_i |x_i| \to \infty \\ \eta(x,y) \to 0}} \min\left(\prod_{j=0}^{n} h_M(\alpha_j(x)), \left|\prod_{j=0}^{n} h_M(\alpha_j(x)) - \prod_{j=0}^{n} h_M(\alpha_j(y))\right|\right) = 0$$

for all $n \geqslant 0$. By Assumption (A.3), it follows that

$$\min\left(\prod_{j=0}^{n} h_M(\alpha_j(X_1)), \left|\prod_{j=0}^{n} h_M(\alpha_j(X_1)) - \prod_{j=0}^{n} h_M(\alpha_{j+1}(x))\right|\right) \to 0$$

in probability, as $s(x) \to \infty$, for all $n \geqslant 0$.

Next, we note that by (4.23)

$$\lambda(x)h_M(x) \leqslant (1+\varepsilon) \max_{0\leqslant n\leqslant n_0} \rho_1^{-n} \prod_{j=0}^{n} h_M(\alpha_j(x))$$

and by (4.22)

$$\max_{0\leqslant n\leqslant n_0} \rho_1^{-n} \prod_{j=0}^{n} h_M(\alpha_{j+1}(x)) \leqslant \rho_1 \lambda(x).$$

Therefore,

$$\begin{aligned}
&\limsup_{\substack{|a(x)|\to\infty \\ \min_i |x_i|\to\infty}} E_x\left(\frac{\lambda(X_1)}{\lambda(x)}|\varphi(X_1)|^r 1_{|\xi_1|>M}\right) \\
&\quad\leqslant \limsup_{\substack{|a(x)|\to\infty \\ \min_i |x_i|\to\infty}} E_x\left(\frac{(1+\varepsilon)}{\lambda(x)} \max_{0\leqslant n\leqslant n_0} \rho_1^{-n} \prod_{j=0}^{n} h_M(\alpha_j(X_1)) + \frac{\lambda(X_1)|M/\xi_1|^r}{\lambda(x)} 1_{|\xi_1|>M}\right) \\
&\quad\leqslant \limsup_{\substack{|a(x)|\to\infty \\ \min_i |x_i|\to\infty}} \frac{(1+\varepsilon)}{\lambda(x)} \max_{0\leqslant n\leqslant n_0} \rho_1^{-n} \prod_{j=0}^{n} h_M(\alpha_{j+1}(x)) \\
&\quad\leqslant (1+\varepsilon)\rho_1 < 1.
\end{aligned}$$

The conclusion holds by Theorem 2.1. □

Before we prove Theorem 3.1 we need the next two lemmas. A nonnegative function $v$ on $\mathbb{R}^d$ is said to be *unbounded off petite sets* if $\{x\colon v(x)\leqslant K\}$ is petite for all $K<\infty$ (cf. Meyn and Tweedie, 1993).

**Lemma 4.1.** *Suppose $\{X_t\}$ is a Markov chain with representation*

$$X_t = \alpha(X_{t-1}) + \gamma(e_t; X_{t-1}). \tag{4.24}$$

*Suppose also $s_0$, $s_1$ and $w$ are nonnegative functions on $\mathbb{R}^d$ such that for all finite $M_1, M_2$,*

$$s_1(x)\leqslant M_1 \quad \text{and} \quad w(\gamma(e_1;x))\leqslant M_2 \Rightarrow s_0(\alpha(x)+\gamma(e_1;x))\leqslant K \quad \text{for some } K<\infty \tag{4.25}$$

*and*

$$\sup_{s_1(x)\leqslant M_1} \boldsymbol{E}(w^r(\gamma(e_1;x))) < \infty \quad \text{for some } r>0.$$

*If $s_0$ is unbounded off petite sets then $s_1$ is unbounded off petite sets.*

**Proof.** Given $M_1<\infty$, choose $M_2$ large enough so that

$$\sup_{s_1(x)\leqslant M_1} \boldsymbol{E}(w^r(\gamma(e_1;x))) < M_2^r.$$

Then choose $K$ large enough so that the implication in (4.25) holds. Using Markov's inequality,

$$\inf_{s_1(x)\leqslant M_1} P_x(s_0(X_1)\leqslant K)\geqslant \inf_{s_1(x)\leqslant M_1} \boldsymbol{P}(w(\gamma(e_1;x))\leqslant M_2) > 0.$$

Since $\{x\colon s_0(x)\leqslant K\}$ is petite, it follows that $\{x:\ s_1(x)\leqslant M_1\}$ is also petite (Meyn and Tweedie, 1993, Proposition 5.5.4(i)). $\square$

For $x\in\mathbb{R}^d$, let $a_0(x)=x_1$, $\alpha_0(x)=x$ and $\alpha(x)=(a(x),x_1,\dots,x_{d-1})$. Also, let $\alpha_j(x)=\alpha(\alpha_{j-1}(x))$ and $a_j(x)=a(\alpha_{j-1}(x))$ for $j\geqslant 1$. Note that

$$\alpha_j(x)=\begin{cases}(a_j(x),a_{j-1}(x),\dots,a(x),x_1,\dots,x_{d-j}) & \text{if } j<d,\\ (a_j(x),a_{j-1}(x),\dots,a_{j-d+1}(x)) & \text{if } j\geqslant d.\end{cases}$$

Now let $||\cdot||$ be any norm on $\mathbb{R}^d$. If $\{X_t\}$ is a $\mu_d$-irreducible $T$-chain then compact sets are petite (Meyn and Tweedie, 1993, Theorem 6.2.5). In particular, $\{x\colon ||x||\leqslant K\}$ is petite for each finite $K$. Using the previous lemma, we now bootstrap from this to show that $a(x)=\varphi(x)x_1$ is unbounded off petite sets.

**Lemma 4.2.** *Assume $\{X_t\}$ is a $\mu_d$-irreducible $T$-chain defined by* (1.1) and (2.1) *and let $r>0$. If*

$$\sup_{|a(x)|\leqslant M} \boldsymbol{E}(|c(e_1;x)|^r)<\infty \quad \textit{for each } M<\infty, \tag{4.26}$$

*then $s_j(x)=||\alpha_j(x)||$ is unbounded off petite sets for $j=1,\dots,d$ and $|a(x)|$ is unbounded off petite sets.*

**Proof.** Let $\alpha(x)=(a(x),x_1,\dots,x_{d-1})$ and $\gamma(e_1;x)=(c(e_1;x)+\vartheta(x),0,\dots,0)$ so that $\{X_t\}$ satisfies (4.24). Without loss of generality, we assume the norm is such that $||(x_1,0,\dots,0)||=|x_1|$. Choose $L$ so that $|\varphi(x)|\leqslant L$ for all $x$. Clearly $|a_j(x)|\leqslant L^j|x_1|$ for all $x$. Hence it is easily shown that, for some $L_j<\infty$ and for each $j=1,\dots,d$,

$$\begin{aligned} s_{j-1}(X_1)&=||\alpha_{j-1}(X_1)||\leqslant L_j(||\alpha_j(x)||+||\gamma(e_1;x)||)\\ &=L_j(s_j(x)+|c(e_1;x)|+L). \end{aligned}$$

Also $|a(x)|\leqslant K_j s_j(x)$ for some finite $K_j$, each $j=1,\dots,d$. Now let $s_0(x)=w(x)=||x||$. Since compact sets are petite, $s_0$ is unbounded off petite sets. By (4.26), Lemma 4.1, and induction, $s_j(x)$ is unbounded off petite sets for $j=1,\dots,d$. Since $||\alpha_d(x)||\to\infty$ if and only if $|a(x)|\to\infty$ it also follows that $|a(x)|$ is unbounded off petite sets. $\square$

**Proof of Theorem 3.1.** As in earlier results, the proof of each part consists of defining an appropriate test function and checking a drift condition. Because the intercept function $\vartheta(x)$ plays a critical role, however, the computations are more intricate. Throughout we rely the simple observation that if $|a|>|b|$ then $|a+b|=|a|+\operatorname{sgn}(a)b$.

(i) We may assume without loss that $r\leqslant 2$. Choose $M<\infty$ and $\varepsilon>0$ to satisfy Assumption (A.4) and, by (3.4), to satisfy $\varepsilon\lambda_{u^*}/\lambda_u<1$,

$$\begin{aligned} &|\varphi(x)x_1+\vartheta(x)|<|\phi_u x_1+\theta_u|-\varepsilon|x_1|^{-s},\\ &|\varphi(x)x_1|>4|\vartheta(x)| \quad\text{and}\quad |\phi_u x_1|>|\theta_u| \end{aligned} \tag{4.27}$$

if $|a(x)| > M$ and $x \in Q_{u,M}$, for each $u \in C$, $C \in \mathscr{C}$. Let $\phi_u = \sup_{x\in Q_{u,M}} |\varphi(x)|$ for all $u$ not in a cycle.

Let $k_C$ be the length of cycle $C$. It is possible to determine constants $\delta_u \geqslant 0$ satisfying

$$\delta_{u^*} + \lambda_{u^*}\sigma_u\theta_u - \delta_u = \frac{1}{k_C}\sum_{v\in C}\lambda_{v^*}\sigma_v\theta_v \quad \text{for } u \in C,\ C \in \mathscr{C}, \tag{4.28}$$

where $\sigma_u = \mathrm{sgn}(\phi_u u_1)$ and $u^*$ is the successor to $u$. Note that the number of negative $\phi_u$ in each cycle must be even. By (3.2), (3.3) and (4.28),

$$\lambda_{u^*}|\phi_u| \leqslant \lambda_u \quad \text{and} \quad \delta_{u^*} + \lambda_{u^*}\sigma_u\theta_u \leqslant \delta_u \quad \text{for } u \in C,\ C \in \mathscr{C}. \tag{4.29}$$

As before, let $\alpha(x) = (a(x), x_1, \ldots, x_{d-1})$. Next, define

$$\lambda(x) = \begin{cases} \lambda_u & \text{if } x \in Q_{u,M} \text{ and } u \text{ is in a cycle,} \\ 1 & \text{if } \min_i |x_i| \leqslant M \text{ or } x \in Q_{u,M} \text{ and } u \text{ has no successor,} \\ \lambda(\alpha(x))(|\phi_u| + \varepsilon) & \text{if } x \in Q_{u,M} \text{ and } u \text{ has a successor but } u \text{ is not in} \\ & \text{a cycle,} \end{cases}$$

the third part of the definition being recursive (similar to that in the proof of Theorem 2.3). Also define

$$\delta(x) = \begin{cases} \delta_u & \text{if } x \in Q_{u,M} \text{ and } u \text{ is in a cycle,} \\ 0 & \text{otherwise.} \end{cases}$$

Let $L$ be such that $|\varphi(x)| \leqslant L$ and $1/L \leqslant \lambda(x) \leqslant L$ for all $x$. Define $v(x)$ as in (4.6) (with $K > L^2$) and

$$V(x) = K^d v(x)(\lambda(x)|x_1| + \delta(x)).$$

Now suppose $X_0 = x \in Q_{u,M}$ where $u \in C$, $C \in \mathscr{C}$. If $|a(x)| > 2M$ and $|c(e_1;x)| \leqslant |a(x)|/3$ then $X_1 \in Q_{u^*,M}$, $|c(e_1;x)| < |\varphi(x)x_1 + \vartheta(x)|$ and $\mathrm{sgn}(\xi_1) = \mathrm{sgn}(\varphi(x)x_1 + \vartheta(x)) = \sigma_u$. Also, by (4.27),

$$\begin{aligned} |\xi_1| &= |\varphi(x)x_1 + \vartheta(x)| + \sigma_u c(e_1;x) \leqslant |\phi_u x_1 + \theta_u| + \sigma_u c(e_1;x) - \varepsilon|x_1|^{-s} \\ &= |\phi_u x_1| + \sigma_u\theta_u + \sigma_u c(e_1;x) - \varepsilon|x_1|^{-s}. \end{aligned}$$

Hence, by (4.29),

$$\begin{aligned} &\lambda(X_1)|\xi_1| + \delta(X_1) - \lambda(x)|x_1| - \delta(x) \\ &\quad \leqslant \lambda_{u^*}(|\phi_u x_1| + \sigma_u\theta_u + \sigma_u c(e_1;x) - \varepsilon|x_1|^{-s}) + \delta_{u^*} - \lambda_u|x_1| - \delta_u \\ &\quad \leqslant \lambda_{u^*}(\sigma_u c(e_1;x) - \varepsilon|x_1|^{-s}). \end{aligned} \tag{4.30}$$

Furthermore, if $|a(x)|$ is large enough and $|c(e_1;x)| \leqslant |a(x)|/3$, then (4.27) and (4.29) imply

$$\begin{aligned} \lambda_{u^*}|\sigma_u c(e_1;x) - \varepsilon|x_1|^{-s}| &\leqslant \lambda_{u^*}(|a(x)|/3 + \varepsilon|x_1|^{-s}) \\ &\leqslant \lambda_{u^*}(4|\varphi(x)x_1 + \vartheta(x)|/9 + \varepsilon|x_1|^{-s}) \\ &\leqslant \lambda_{u^*}(4|\phi_u x_1 + \theta_u|/9 + \varepsilon|x_1|^{-s}) \leqslant V(x)/2. \end{aligned} \tag{4.31}$$

Therefore, by (4.30), if $x \in Q_{u,M}$ then

$$
\begin{aligned}
&V(X_1) - V(x)\\
&\quad \leqslant (\lambda(X_1)|\xi_1| + \delta(X_1) - \lambda(x)|x_1| - \delta(x))1_{|c(e_1;x)|\leqslant|a(x)|/3} + V(X_1)1_{|c(e_1;x)|>|a(x)|/3}\\
&\quad \leqslant \lambda_{u^*}(\sigma_u c(e_1;x) - \varepsilon|x_1|^{-s})1_{|c(e_1;x)|\leqslant|a(x)|/3} + 5K^d L|c(e_1;x)|1_{|c(e_1;x)|>|a(x)|/3},
\end{aligned}
$$

when $|a(x)|$ is large. Let $H(y) = y^{1+s}$. Then, for $|z| \leqslant y/2$ and $r \leqslant 2$,

$$
H(y+z) - H(y) \leqslant (1+s)y^s z + s|z|^r (y/2)^{1+s-r}.
$$

Applying this inequality with $y = V(x)$ and $z = \lambda_{u^*}(\sigma_u c(e_1;x) - \varepsilon|x_1|^{-s})$ and noting that by (4.31), $|z|$ is no more than $y/2$ if $|c(e_1;x)| \leqslant |a(x)|/3$ and $|a(x)|$ is large enough,

$$
\begin{aligned}
&H(V(X_1)) - H(V(x))\\
&\quad \leqslant (H(y+z) - H(y))1_{|c(e_1;x)|\leqslant|a(x)|/3} + H(V(X_1))1_{|c(e_1;x)|>|a(x)|/3}\\
&\quad \leqslant (1+s)V^s(x)\lambda_{u^*}(\sigma_u c(e_1;x) - \varepsilon|x_1|^{-s})1_{|c(e_1;x)|\leqslant|a(x)|/3}\\
&\qquad + s\lambda_{u^*}^r|\sigma_u c(e_1;x) - \varepsilon|x_1|^{-s}|^r (V(x)/2)^{1+s-r}\\
&\qquad + (5K^d L|c(e_1;x)|)^{1+s}1_{|c(e_1;x)|>|a(x)|/3}.
\end{aligned}
\tag{4.32}
$$

Since either $s = 0$ and $r \geqslant 1$ or $r > 1+s$, Assumption (A.5) and (4.32) imply

$$
\begin{aligned}
&\limsup_{\substack{|a(x)|\to\infty\\ x\in Q_{u,M},|\varphi(x)|>\varepsilon}} E_x(H(V(X_1)) - H(V(x)))\\
&\quad \leqslant -(1+s)\varepsilon\lambda_{u^*}\lambda_u^s < 0 \quad \text{for } u \in C,\ C \in \mathscr{C}.
\end{aligned}
\tag{4.33}
$$

Requiring $|\varphi(x)| > \varepsilon$ in (4.33) ensures that $V^s(x)\boldsymbol{E}(|c(e_1;x)|1_{|c(e_1;x)|>|a(x)|/3})$ vanishes. However,

$$
\limsup_{\substack{|a(x)|\to\infty\\ x\in Q_{u,M},|\varphi(x)|\leqslant\varepsilon}} E_x\left(\frac{H(V(X_1)) - H(V(x))}{H(V(x))}\right) \leqslant (\varepsilon\lambda_{u^*}/\lambda_u)^{1+s} - 1 < 0. \tag{4.34}
$$

Using the definitions of $\lambda, \delta, v$ and $V$ and the choice for $K$, it is easily shown that

$$
\limsup_{\substack{|a(x)|\to\infty\\ \min_i|x_i|\leqslant M}} E_x\left(\frac{H(V(X_1)) - H(V(x))}{H(V(x))}\right) \leqslant (L^2/K)^{1+s} - 1 < 0 \tag{4.35}
$$

and, for $u$ not in any cycle,

$$
\limsup_{\substack{|a(x)|\to\infty\\ x\in Q_{u,M}}} E_x\left(\frac{H(V(X_1)) - H(V(x))}{H(V(x))}\right) \leqslant (|\phi_u|/(|\phi_u| + \varepsilon))^{1+s} - 1 < 0. \tag{4.36}
$$

From (4.33)–(4.36), we conclude there exists $M^* < \infty$ such that

$$\sup_{|a(x)|>M^*} E_x(H(V(X_1)) - H(V(x))) < 0.$$

Clearly, also,

$$\sup_{|a(x)|\leqslant M^*} E_x(H(V(X_1))) < \infty. \tag{4.37}$$

By Lemma 4.2, $\{x: |a(x)|\leqslant M^*\}$ is petite. From Meyn and Tweedie (1993, Theorem 13.0.1), therefore, $\{X_t\}$ is ergodic.

(ii) Using (3.5) and (3.6), choose $M$ large enough and $\varepsilon > 0$ small enough so that

$$\inf_{\substack{|a(x)|>M \\ x\in Q_{u,M}}} \boldsymbol{E}(c^2(e_1;x)) > 6\varepsilon\lambda_u/\lambda_{u^*} \quad \text{and} \quad \varepsilon\lambda_{u^*}/\lambda_u < 1 \tag{4.38}$$

for each $u \in C$, $C \in \mathscr{C}$, and if $|a(x)| > M$ and $x \in Q_{u,M}$ then

$$|\varphi(x)x_1+\vartheta(x)|<|\phi_u x_1+\theta_u|+\varepsilon|x_1|^{-1}, \quad |\varphi(x)x_1| > 4|\vartheta(x)| \quad \text{and} \quad |\phi_u x_1|>|\theta_u|.$$

The proof is similar to part (i), defining $\lambda$, $\delta$, $v$ and $V$ as before.

Note that $\log(1+z)\leqslant z - z^2/3$ if $|z|\leqslant\frac{1}{2}$. Using inequalities analogous to (4.30) and (4.31) and using $z = (\lambda_{u^*}(\sigma_u c(e_1;x) + \varepsilon|x_1|^{-1}))/(1 + V(x))$, an argument similar to (4.32) above gives, for $X_0 = x \in Q_{u,M}$, $u \in C$, $C \in \mathscr{C}$ and $|a(x)|$ large enough,

$$\begin{aligned}
&(1 + V(x))^2\log\left(1 + \frac{V(X_1) - V(x)}{1 + V(x)}\right) \\
&\quad\leqslant \left(\lambda_{u^*}\sigma_u\left(1 + V(x) - \frac{2\varepsilon}{3|x_1|}\right)c(e_1;x)\right. \\
&\qquad \left.-\frac{\lambda_{u^*}^2 c^2(e_1;x)}{3} + 2\lambda_{u^*}\lambda_u\varepsilon\right)1_{|c(e_1;x)|\leqslant|a(x)|/3} \\
&\qquad + 5K^d L(1 + V(x))|c(e_1;x)|1_{|c(e_1;x)|>|a(x)|/3}.
\end{aligned} \tag{4.39}$$

Since Assumption (A.5) holds with $r = 2$, therefore, (4.38) and (4.39) imply

$$\begin{aligned}
&\limsup_{\substack{|a(x)|\to\infty \\ x\in Q_{u,M},|\varphi(x)|>\varepsilon}} V^2(x)E_x(\log(1 + V(X_1)) - \log(1 + V(x))) \\
&\quad\leqslant 2\lambda_{u^*}\lambda_u\varepsilon - \frac{\lambda_{u^*}^2}{3}\inf_{\substack{|a(x)|>M \\ x\in Q_{u,M}}} \boldsymbol{E}(c^2(e_1;x)) < 0 \quad \text{for } u \in C,\ C \in \mathscr{C}.
\end{aligned} \tag{4.40}$$

It is easy to show that (4.34)–(4.36) hold with $H(y) = y$ and $s = 0$ as well. By these three results and (4.40), it follows that

$$\limsup_{|a(x)|\to\infty} V^2(x)E_x(\log(1 + V(X_1)) - \log(1 + V(x))) < 0.$$

Along with (4.37) and Lemma 4.2, as in (i) above, this suffices to prove Harris recurrence (cf. Meyn and Tweedie, (1993, Theorem 9.1.8). $\square$

**Proof of Corollary 3.2.** Fix the constants $M$, $\varepsilon$, $K$ and $L$ as in the proof of Theorem 3.1. Define $V(x)$ as in that proof. Suppose $x \in Q_{u,M}$ where $u$ is in some cycle $C$. From the argument for Theorem 3.1 we have, for large enough $|a(x)|$,

$$V(X_1) - V(x) \leqslant W(x) \stackrel{\text{def}}{=} \lambda_{u^*}(\sigma_u c(e_1; x) - \varepsilon) + 6K^d L|c(e_1; x)| 1_{|c(e_1;x)| > |a(x)|/3}.$$

Note that $\limsup_{|a(x)| \to \infty} \boldsymbol{E}(W(x)) \leqslant -\varepsilon/L < 0$, $\{W(x)\}$ is uniformly integrable and there exists $\eta_1 > 0$ such that $\{e^{\eta_1 W(x)}\}$ is uniformly integrable. It follows that, for some $\eta_2 > 0$ and all $u$ in a cycle,

$$\limsup_{\substack{|a(x)| \to \infty \\ x \in Q_{u,M}}} E_x(\mathrm{e}^{\eta_2(V(X_1) - V(x))}) \leqslant \limsup_{\substack{|a(x)| \to \infty \\ x \in Q_{u,M}}} \boldsymbol{E}(\mathrm{e}^{\eta_2 W(x)}) < 1. \tag{4.41}$$

(See Cline and Pu, 1999, Lemma 4.2.)

We can also show that there exists $K_1 < \infty$, $K_2 < \infty$ and $\eta_3 > 0$ such that, for $u$ not in a cycle,

$$\limsup_{\substack{|a(x)| \to \infty \\ x \in Q_{u,M}}} E_x(e^{\eta_3(V(X_1) - V(x))}) \leqslant \limsup_{\substack{|a(x)| \to \infty \\ x \in Q_{u,M}}} \boldsymbol{E}(\mathrm{e}^{-\eta_3[\varepsilon\lambda_{u^*}|x_1| + K_1 + K_2|c(e_1;x)|]}) = 0 \tag{4.42}$$

and such that

$$\limsup_{\substack{|a(x)| \to \infty \\ \min_i |x_i| \leqslant M}} E_x(\mathrm{e}^{\eta_3(V(X_1) - V(x))}) \leqslant \limsup_{\substack{|a(x)| \to \infty \\ \min_i |x_i| \leqslant M}} \boldsymbol{E}(\mathrm{e}^{-\eta_3[(K - L^2)|x_1| + K_1 + K_2|c(e_1;x)|]}) = 0. \tag{4.43}$$

Furthermore, for some $\eta_4 > 0$ and all $M^* < \infty$,

$$\sup_{|a(x)| \leqslant M^*} E_x(\mathrm{e}^{\eta_4(V(X_1) - V(x))}) < \infty. \tag{4.44}$$

Let $\eta^* = \min(\eta_2, \eta_3, \eta_4)$. Geometric ergodicity follows from (4.41)–(4.44), Lemma 4.2 and the drift condition in Meyn and Tweedie (1993, Theorem 15.0.1) applied with the test function $V_1(x) = \mathrm{e}^{\eta^* V(x)}$.

**Proof of Theorem 3.3.** Again, each proof consists of verifying a drift condition. Also, we again observe that if $|a| > |b|$ then $|a + b| = |a| + \operatorname{sgn}(a)b$.

(i) We can assume without loss that $0 < s < r - 1 \leqslant 1$. According to (3.8), choose $M < \infty$ large enough and $\varepsilon \in (0, r - s - 1)$ small enough that

$$|\varphi(x)x_1 + \vartheta(x)| > |\phi_u x_1 + \theta_u| + \varepsilon|x_1|^{-s}, \qquad |\varphi(x)x_1| > |\vartheta(x)| \quad \text{and} \quad |\phi_u x_1| > 4|\theta_u|$$

for $x \in Q_{u,M}$, $u \in C$. Define constants $\delta_u \geqslant 0$ to satisfy (4.28). Then by (3.2) and (3.7), for each $u \in C$,

$$\lambda_{u^*}|\phi_u| \geqslant \lambda_u \quad \text{and} \quad \delta_{u^*} + \lambda_{u^*}\sigma_u\theta_u \geqslant \delta_u. \tag{4.45}$$

Define

$$V(x) = \begin{cases} \lambda_u|x_1| + \delta_u & \text{if } x \in Q_{u,M},\ u \in C, \\ 0 & \text{otherwise.} \end{cases}$$

Also let $R = \bigcup_{u \in C} Q_{u,M}$. Similar to that which gave (4.30), if $X_0 = x \in Q_{u,M}$, $u \in C$, $|c(e_1; x)| \leqslant |\phi_u x_1|/3$ and $|x_1|$ is large enough, then

$$V(X_1) - V(x) \geqslant \lambda_{u^*}(\sigma_u c(e_1; x) + \varepsilon |x_1|^{-s}). \tag{4.46}$$

Let $H(y) = 1 - (1+y)^{-\varepsilon}$ and note that for $|z| \leqslant y/2$,

$$H(y+z) - H(y) \geqslant \frac{\varepsilon z}{(1+y)^{1+\varepsilon}} - \frac{\varepsilon(1+\varepsilon)|z|^r}{2(1+y/2)^{r+\varepsilon}}. \tag{4.47}$$

Analogous to (4.31), $z = \lambda_{u^*}(\sigma_u c(e_1; x) + \varepsilon |x_1|^{-s})$ is no more than $V(x)/2$ in absolute value if $|c(e_1; x)| \leqslant |\phi_u x_1|/3$ and $|x_1|$ is large enough. Thus, we apply (4.46) and (4.47) with this $z$ and with $y = V(x)$ to obtain

$$\begin{aligned} &H(V(X_1)) - H(V(x)) \\ &\quad \geqslant (H(y+z) - H(y)) 1_{|c(e_1;x)| \leqslant |\phi_u x_1|/3} - 1_{|c(e_1;x)| > |\phi_u x_1|/3} \\ &\quad \geqslant \left( \frac{\varepsilon(\sigma_u c(e_1; x) + \varepsilon |x_1|^{-s})}{(1+V(x))^{1+\varepsilon}} - \frac{\varepsilon(1+\varepsilon)|\lambda_{u^*}(\sigma_u c(e_1; x) + \varepsilon |x_1|^{-s})|^r}{2(1+V(x)/2)^{r+\varepsilon}} \right) 1_{|c(e_1;x)| \leqslant |\phi_u x_1|/3} \\ &\qquad - 1_{|c(e_1;x)| > |\phi_u x_1|/3}. \end{aligned}$$

Therefore, since Assumption (A.5) holds with $r > 1 + s + \varepsilon$,

$$\limsup_{\substack{|x_1| \to \infty \\ x \in Q_{u,M}}} |x_1|^{1+s+\varepsilon} E_x(H(V(X_1)) - H(V(x))) > \varepsilon^2 \lambda_u^{-1-\varepsilon} > 0 \quad \text{for } u \in C.$$

Also, for any $M^*$, $\mu_d(\{x \in R\colon V(x) > M^*\}) > 0$. By (Meyn and Tweedie, 1993, Theorem 8.0.2) this shows that $\{X_t\}$ is transient.

(ii) In choosing $\delta_u$'s to satisfy (4.28), the choice is unique up to an additive constant. We choose them here to so that $\delta_u/\lambda_u \leqslant -L_1$ for all $u \in C$. By (3.9) we may choose $M > L_2$ large enough that $\lambda_u M + \delta_u \geqslant 0$ for all $u \in C$ and

$$|\varphi(x)x_1 + \vartheta(x)| \geqslant |\phi_u x_1 + \theta_u| - \Delta(x), \quad |\varphi(x)x_1| > \vartheta(x)| \quad \text{and} \quad |\phi_u x_1| > |\theta_u| \tag{4.48}$$

for $x \in Q_{u,M}$, $u \in C$. Let $K_1 = \max_{u \in C}(\lambda_u M + \delta_u)$. Now define

$$V(x) = \begin{cases} \lambda_u |x_1| + \delta_u & \text{if } x \in Q_{u,M},\ u \in C, \\ 0 & \text{otherwise} \end{cases}$$

and $R = \bigcup_{u \in C} Q_{u,M}$. By (4.45) and (4.48), for $x \in Q_{u,M}$,

$$V(x) \leqslant \lambda_{u^*} |\phi_u x_1 + \theta_u| + \delta_{u^*} \leqslant \lambda_{u^*}(|\varphi(x)x_1 + \vartheta(x)| + \Delta(x)) + \delta_{u^*}.$$

Recall that if $X_0 = x \in Q_{u,M}$, then $X_1 \in Q_{u^*,M}$ if and only if $|\xi_1| > M$ and $\mathrm{sgn}(\xi_1) = \sigma_u$. Also, if $x \in Q_{u,M}$,

$$|\xi_1| = \begin{cases} \sigma_u c(e_1; x) + |\varphi(x)x_1 + \vartheta(x)| & \text{if } \mathrm{sgn}(\xi_1) = \sigma_u, \\ -\sigma_u c(e_1; x) - |\varphi(x)x_1 + \vartheta(x)| & \text{if } \mathrm{sgn}(\xi_1) \neq \sigma_u. \end{cases}$$

Therefore we obtain, for $x \in Q_{u,M}$, $u \in C$,

$$\begin{aligned} V(X_1) - V(x) &\geqslant (\lambda_{u^*}|\xi_1| + \delta_{u^*}) 1_{X_1 \in Q_{u^*,M}} - V(x) 1_{X_1 \notin Q_{u^*,M}} \\ &\geqslant \lambda_{u^*}(\sigma_u c(e_1; x) - \Delta(x)) - (\lambda_{u^*}|\xi_1| + \delta_{u^*}) 1_{|\xi_1| \leqslant M, \mathrm{sgn}(\xi_1) = \sigma_u} \\ &\quad + (\lambda_{u^*}|\xi_1| - \delta_{u^*}) 1_{\mathrm{sgn}(\xi_1) \neq \sigma_u} \\ &\geqslant \lambda_{u^*}(\sigma_u c(e_1; x) - \Delta(x) + |\xi_1| 1_{\mathrm{sgn}(\xi_1) \neq \sigma_u}) + \delta_{u^*} 1_{|\xi_1| \leqslant M, \mathrm{sgn}(\xi_1) = \sigma_u} \\ &\quad - K_1 1_{X_1 \notin R}. \end{aligned} \tag{4.49}$$

Note that $\lambda_{u^*}$ and $\sigma_u$ are fixed, given $X_0 = x \in Q_{u,M}$. By Assumption (A.5), the definition of $\Delta(x)$ and the choice of $\delta_{u^*}$ and $M$, it is easily seen that

$$w(x, X_1) \stackrel{\text{def}}{=} \lambda_{u^*}(\sigma_u c(e_1; x) - \Delta(x) + |\xi_1| 1_{\mathrm{sgn}(\xi_1) \neq \sigma_u}) + \delta_{u^*} 1_{|\xi_1| \leqslant M, \mathrm{sgn}(\xi_1) = \sigma_u}$$

is uniformly bounded in absolute mean and has nonnegative mean. Also,

$$\mu_d(\{x \in R\colon V(x) > K_1\}) > 0. \tag{4.50}$$

By (4.49), (4.50) and Lemma 3.4, $\{X_t\}$ is not positive recurrent. □

**Proof of Lemma 3.4.** Let $\mathscr{F}_t$ be the $\sigma$-field generated by $(X_0, X_1, \ldots, X_t)$. Define the random variables $Y_t = w(X_{t-1}, X_t)$. Define $\tau = \inf\{t \geqslant 1\colon X_t \notin R\}$ and suppose $x \in R$ and $E_x(\tau) < \infty$. Then, by (i),

$$\begin{aligned} -V(x) = V(X_\tau) - V(x) &= \sum_{t=1}^{\infty} (V(X_t) - V(X_{t-1})) 1_{\tau \geqslant t} \\ &\geqslant \sum_{t=1}^{\infty} (Y_t - K_1 1_{X_t \notin R}) 1_{\tau \geqslant t}. \end{aligned}$$

Applying Fubini's theorem and (ii), we obtain

$$\begin{aligned} -V(x) &\geqslant \sum_{t=1}^{\infty} E_x(\boldsymbol{E}(Y_t 1_{\tau \geqslant t} - K_1 1_{\tau = t} | \mathscr{F}_{t-1})) \\ &\geqslant \sum_{t=1}^{\infty} E_x(-K_1 1_{\tau = t}) = -K_1. \end{aligned}$$

We have shown $x \in R$ and $E_x(\tau) < \infty$ implies $V(x) \leqslant K_1$. But if $\{X_t\}$ is positive recurrent and $\psi(R^c) > 0$ then $E_x(\tau) < \infty$ almost everywhere ($\psi$) and

$$\psi(\{x \in R\colon V(x) > K_1\}) \leqslant \psi(\{x \in R\colon E_x(\tau) = \infty\}) = 0,$$

contradicting (iii). So $\{X_t\}$ must not be positive recurrent. □

**Proof of Corollary 3.5.** Note that (3.10) and (3.11) imply (1.2). The result follows from Theorems 3.1 and 3.3 except that the constants $\phi_u$ and $\theta_u$, while being the limits in (1.2), do not necessarily satisfy the conditions in Theorem 3.1 or 3.3. So, for example, assume case (i) and let $C$ be any cycle. If $\prod_{u \in C} \phi_u < 1$, there is $\varepsilon$ small enough that (3.3) and (3.4) both hold with $\phi_u$ replaced by $\phi'_u = \phi_u + \mathrm{sgn}(\phi_u)\varepsilon$, $\theta_u$ replaced by $\theta'_u = 0$ and $\lambda_u$ recalculated accordingly. If $\prod_{u \in C} \phi_u = 1$ and $\sum_{u \in C} \lambda_{u^*} \sigma_u \theta_u < 0$ then a small adjustment to the $\theta_u$'s is all that is needed. Cases (ii) and (iii) are dealt with likewise. □

## References


An, H.Z., Huang, F.C., 1996. The geometrical ergodicity of nonlinear autoregressive models. Statist. Sinica 6, 943–956.

Bhattacharya, R., Lee, C., 1995. Ergodicity of first order autoregressive models. J. Theoret. Probab. 8, 207–219.

Chan, K.S., 1990. Deterministic stability, stochastic stability, and ergodicity. In: H. Tong, Appendix 1 in Non-linear Time Series Analysis: A Dynamical System Approach. Oxford University Press, London.

Chan, K.S., 1993. A review of some limit theorems of Markov chains and their applications. In: H. Tong (Ed.), Dimensions, Estimation and Models. World Scientific, Singapore, pp. 108–1135.

Chan, K.S., Petrucelli, J.D., Tong, H., Woolford, S.W., 1985. A multiple threshold AR(1) model. J. Appl. Probab. 22, 267–279.

Chan, K.S., Tong, H., 1985. On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. Adv. Appl. Probab. 17, 666–678.

Chan, K.S., Tong, H., 1986. On estimating thresholds in autoregressive models. J. Time Series Anal. 7, 179–190.

Chan, K.S., Tong, H., 1994. A note on noisy chaos. J. Royal Statist. Soc. 56, 301–311.

Chen, R., Tsay, R.S., 1991. On the ergodicity of TAR(1) process. Ann. Appl. Probab. 1, 613–634.

Chen, R., Tsay, R.S., 1993a. Functional-coefficient autogressive models. J. Amer. Statist. Assoc. 88, 298–308.

Chen, R., Tsay, R.S., 1993b. Nonlinear additive ARX models. J. Amer. Statist. Assoc. 88, 955–967.

Cline, D.B.H., Pu, H.H., 1998a. Geometric transience of nonlinear time series. Technical report, Department of Statistics, Texas A&M University, College Station TX.

Cline, D.B.H., Pu, H.H., 1998b. Verifying irreducibility and continuity of a nonlinear time series. Statist. Probab. Lett. 40, 139–148.

Cline, D.B.H., Pu, H.H., 1999. Geometric ergodicity of nonlinear time series. Statistica Sinica, to appear.

Guégan, D., Diebolt, J., 1994. Probabilistic properties of the $\beta$-ARCH model. Statist. Sinica 4, 71–87.

Guo, M., Petrucelli, J.D., 1991. On the null-recurrence and transience of a first order SETAR model. J. Appl. Probab. 28, 584–592.

Härdle, W., Lütkepohl, H., Chen, R., 1997. A review of nonparametric time series analysis. Internat. Statist. Rev. 65, 49–72.

Lamperti, J., 1960. Criteria for the recurrence or transience of stochastic processes I. J. Math. Anal. Appl. 1, 314–330.

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

Lu, Z., 1996. A note on the geometric ergodicity of autoregressive conditional heteroscedasticity (ARCH) model. Statist. Probab. Lett. 30, 305–311.

Meyn, S.P., Tweedie, R.L., 1993. Markov Chains and Stochastic Stability. Springer, London.

Petrucelli, J.D., Woolford, S.W., 1984. A threshold AR(1) model. J. Appl. Probab. 21, 270–286.

Pu, H.H., 1995. Recurrence and transience for autoregressive nonlinear time series. Dissertation, Texas A&M University, College Station, Texas.

Spieksma, F.M., Tweedie, R.L., 1994. Strengthening ergodicity to geometric ergodicity for Markov chains. Stochast. Models 10, 45–74.

Tjøstheim, D., 1990. Non-linear time series and Markov chains. Adv. Appl. Probab. 22, 587–611.

Tjøstheim, D., Auestad, B.H., 1994a. Nonparametric identification of nonlinear time series: projections. J. Amer. Statist. Assoc. 89, 1398–1409.

Tjøstheim, D., Auestad, B.H., 1994b. Nonparametric identification of nonlinear time series: selecting significant lags. J. Amer. Statist. Assoc. 89, 1410–1419.

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