Statistica Sinica **11**(2001), 273-287

GEOMETRIC TRANSIENCE OF NONLINEAR TIME SERIES

Daren B. H. Cline and Huay-min H. Pu

Texas A&M University

Abstract: In this paper we provide conditions for nonlinear time series to be geometrically explosive with positive probability. The paper complements earlier work by the authors on geometrically ergodic nonlinear time series, showing that the conditions for examples in the earlier paper are sharp. We also study the transience of polynomial autoregression models.

Key words and phrases: Markov chain, nonlinear time series, transience.

1. Introduction

Nonlinear time series modeling is attracting increasing interest, particularly with the development of nonparametric methods for estimating such models (eg., Chen and Tsay (1993a,b), Tjøstheim and Auestad (1994a,b)). The stability of general nonlinear models is only partially understood, however. Groundbreaking work appeared in Petruccelli and Woolford (1984), Chan and Tong (1985, 1986), Chan (1990) and Tjøstheim (1990). These papers, and the many that followed, dealt primarily with determining sufficient conditions for stability of nonlinear time series. Only a few also have results on nonstability, including Chan (1990), Tjøstheim (1990), Guo and Petruccelli (1991), Bhattacharya and Lee (1995) and Cline and Pu (1999b).

The conditions for stability/nonstability fall into two classes: conditions that are not necessarily sharp but are applicable to relatively general time series (including nonparametric time series), and conditions that are usually sharp and are applicable to specific parametric models. Except for ARCH models (Lu (1998a,b)) and bilinear models (Pham (1985)), the error distribution usually plays no role in the derived stability condition other than ensuring irreducibility of the process. In Cline and Pu (1999a,c) we showed that the error distribution nevertheless can have a significant effect on stability. In particular, we obtained drift conditions sufficient for general nonlinear models to be geometrically ergodic and applied them to models that failed the drift conditions of the earlier papers.

This article is complementary to Cline and Pu (1999a), hereafter referred to as CP. We now seek sufficient conditions for transience of nonlinear time series models, including some studied in CP. Noting that stable series can have characteristics that, at face value, appear to be explosive, a study of transience helps to identify what is in fact necessary for stability. We also note there is a clear preference for proving geometric ergodicity (when possible) over simply showing ergodicity, because the conclusion is stronger and the drift conditions are more readily checked. We therefore want to provide corresponding "geometric" conditions for transience. Our examples include multivariate threshold models, models with a periodic "coefficient function" and polynomial autoregression models.

As in CP, $\{X_t\}$ denotes a time homogeneous Markov chain defined on $\mathbb{X} \subset \mathbb{R}^p$ by

$$X_t = \alpha(X_{t-1}) + \gamma(e_t; X_{t-1}), \qquad t \geq 1, \tag{1.1}$$

where $\alpha(x)$ and $\gamma(e;x)$ are measurable functions and $\{e_t\}$ is an i.i.d. sequence of random variables, independent of the initial state X_0. The process may be a $\mathbb{R}^p$-valued nonlinear autoregression of order 1, or it may be the state space vector $X_t = (\xi_t, \ldots, \xi_{t-p+1})$ for a nonlinear autoregression $\{\xi_t\}$ of order p. We operate under the assumption that $\mathbb{X}$ is an unbounded subspace of $\mathbb{R}^p$ and that α is essentially unbounded on $\mathbb{X}$. Otherwise, with minimal assumptions on γ, the process would be recurrent.

In Section 2 we provide general drift conditions for transience of a Markov chain, complementing those for geometric ergodicity, and show they imply a type of geometric transience. Specific results and time series examples are in Section 3 and the proofs and lemma are in Section 4.

2. Geometric Transience

We are principally interested in forms of transience for which $||X_t||$ becomes unbounded (as this is most relevant in a time series setting), obtaining useful but readily applied drift conditions that mimic those for geometric ergodicity. As a bonus, we show that the conditions imply the time series explodes geometrically with positive probability. Specifically, we say the time series is *geometrically transient* if there exists $q < 1$ such that $\mathbf{P}(q^t||X_t|| \to \infty \mid X_0 = x) > 0$, as $t \to \infty$, for all $x \in \mathbb{X}$.

The notation and terminology used here are the same as in CP. In particular, $P_x(\cdot) = \mathbf{P}(\cdot|X_0 = x)$ and $E_x(\cdot) = \mathbf{E}(\cdot|X_0 = x)$. We assume throughout that $\{X_t\}$ is a ψ-irreducible Markov chain defined by (1.1) on $\mathbb{X} \subset \mathbb{R}^p$ with maximal irreducibility measure ψ (cf. Meyn and Tweedie (1993)). Conditions for ψ-irreducibility can also be found in Cline and Pu (1998), as well as examples. We also assume

$$\limsup_{||x|| \to \infty} \mathbf{E}(||\gamma(e_1; x)||^r) < \infty, \qquad \text{for some } r > 0. \tag{2.1}$$

Let V be a nonnegative measurable function on $\mathbb{X}$. A drift condition for transience is one which determines that $E_x(1/V(X_1))$ is sufficiently smaller than

$1/V(x)$, when $V(x)$ is large. While $V(x) = ||x||^r$ is an obvious choice for such a test function, model specific choices often work better. As a result, a "directional method" uses $V(x) = ||\alpha(x)||^r$ or $V(x) = |\nu' x|^r$, or even modifies such V with $\lambda(x)V(x)$, where λ and/or ν are chosen suitably for each model. Authors who have used handpicked versions of the directional approach for showing ergodicity include Petruccelli and Woolford (1984) and Chan, Petruccelli, Tong and Woolford (1985). The method performed excellently in determining conditions for both ergodicity and transience of threshold AR(1) models with delay (Cline and Pu (1999b)). Also, checking for transience differs from checking for ergodicity in the sense that, for transience to hold, $V(X_1)$ often need only be large when $V(x)$ is large in a particular section, say $D \subset \mathbb{X}$, while $V(X_1)$ must be more globally small for ergodicity to hold. This means conditions for transience may take a more restrictive look at $V(X_1)$ than do conditions for ergodicity. This directional approach is illustrated next.

Theorem 2.1. *Let Λ be the class of nonnegative functions on $\mathbb{X}$, bounded and bounded away from 0. Assume $V : \mathbb{X} \to \mathbb{R}$ is nonnegative and $D \subset \mathbb{X}$. If*

$$\psi(\{x \in D \colon V(x) > M\}) > 0, \qquad \text{for all } M < \infty, \tag{2.2}$$

$$\limsup_{\substack{V(x)\to\infty \\ x\in D}} (V(x))^r P_x(X_1 \notin D) = 0, \qquad \text{for some } r > 0, \tag{2.3}$$

and there exists $q \in (0,1)$ such that

$$\limsup_{\substack{V(x)\to\infty \\ x\in D}} E_x\Big(\Big(\frac{\lambda(x)}{\lambda(X_1)}\frac{V(x)}{V(X_1)}\Big)^r 1_{\{X_1\in D\}}\Big) < q^r, \quad \text{for some } r > 0 \text{ and some } \lambda \in \Lambda, \tag{2.4}$$

then $\{X_t\}$ is transient and $P_x(q^t V(X_t)1_{\{X_t\in D\}} \to \infty) > 0$ for all $x \in \mathbb{X}$.

When $V(x)/||x||$ remains bounded as $||x|| \to \infty$, say, then one may conclude from Theorem 2.1 that the process has a positive probability of increasing geometrically fast, hence our use of the term "geometric transience". Also, condition (2.4) is completely analogous to the condition for geometric ergodicity (for which the ratio in the expectation is inverted), except it allows for a look at the process restricted to a set D. It does not, however, imply a geometric rate of convergence of return probabilities (to 0), a notion of geometric transience studied by Kendall (1959), Vere-Jones (1962) and Kingman (1963) for countable Markov chains.

In fact, the relationship between these two notions is open to study. We briefly mention two examples to illustrate the issues involved, although a full study is beyond the scope of this paper. First, consider a random walk with drift, $X_t = X_{t-1} + b + e_t$, with $b \neq 0$ and having Gaussian errors. It is easy to see that $P_x(|X_t| \le y)$ decreases geometrically fast, as $t \to \infty$, even though the time series does not increase geometrically fast. Condition (2.4), however, is met with $V(x) = e^x$, $\lambda(x) = 1$, $D = \mathbb{R}$ and r sufficiently small. Second, consider the threshold model, $X_t = \max(X_{t-1}, aX_{t-1}) - b + e_t$, with $a > 1$, $b > 0$ and Cauchy errors. Here, condition (2.4) is met with $V(x) = |x|$, $\lambda(x) = 1$, $D = (0, \infty)$ and $r < 1$, so there is a positive chance that $q^t X_t \to \infty$ for $q > 1/a$. Despite this, the Cauchy errors ensure that $P_x(|X_t| \leq y)$ decreases at a polynomial rate for any x, as $t \to \infty$, since the chance that the process eventually drifts off to $-\infty$ decreases slowly as $x \to \infty$.

Models which are transient but which do not satisfy a geometric drift condition such as (2.4) usually require a more subtle and careful analysis. For example, in Cline and Pu (1999b) we determined sharp conditions for nongeometric ergodicity and nongeometric transience of a parametric threshold model with delay.

The directional approach used in Theorem 2.1 often requires a selective choice of the test function V, or equivalently, of the modifying function λ. An alternative approach was suggested by Tjøstheim (1990), who noted that it suffices to apply drift conditions to the m-step process $\{X_{mt}\}$. A generalization of this "m-step" approach, we show below, is actually equivalent to the directional method.

Theorem 2.2. *Assume $V : \mathbb{X} \to \mathbb{R}$ is nonnegative and $D \subset \mathbb{X}$. If*

$$\sup_x E_x\Big(\Big(\frac{V(x)}{V(X_1)}\Big)^r 1_{\{X_1 \in D\}}\Big) < \infty, \qquad \textit{for some } r > 0, \tag{2.5}$$

and, for some $q \in (0,1)$, either

$$\limsup_{\substack{V(x)\to\infty \\ x \in D}} E_x\Big(\sum_{j=m}^{n} \log\Big(\frac{V(X_{j-1})}{qV(X_j)}\Big) 1_{\{X_j \in D\}}\Big) < 0, \qquad \textit{for some } n \geq m \geq 1 \tag{2.6}$$

or

$$\limsup_{\substack{V(x)\to\infty \\ x \in D}} E_x\Big(\prod_{j=m}^{n} \Big(\frac{V(X_{j-1})}{qV(X_j)}\Big)^{r1_{\{X_j \in D\}}}\Big) < 1, \qquad \textit{for some } r > 0 \textit{ and } n \geq m \geq 1, \tag{2.7}$$

then (2.4) *holds (with a possibly different value for r).*

3. Results and Examples

We start with a surprisingly sharp condition for models with periodic coefficient functions. Suppose X_t is real valued and satisfies

$$X_t = \beta_0(X_{t-1}) + \beta_1(X_{t-1})X_{t-1} + e_t, \tag{3.1}$$

where β_0 and β_1 are bounded and β_1 is periodic with period τ. While such a model may be unusual (but see the estimated coefficient functions for the sunspot example in Chen and Tsay (1993a)), it is a useful illustration of the improvement gained by carefully applying the approaches discussed above. It is well-known, and easily shown, that the time series in (3.1) is geometrically ergodic if $\sup_x |\beta_1(x)| < 1$. In fact, by CP (Theorem 3.4), a sufficient condition for geometric ergodicity is

$$\frac{1}{\tau}\int_0^\tau \log(|\beta_1(u)|)du < 0. \tag{3.2}$$

We now proceed to show that the condition in (3.2) is sharp.

Assumption 3.1. Let $\beta\colon \mathbb{R} \to \mathbb{R}$ be periodic with period τ. For each $\epsilon > 0$, there is an open set $A_\epsilon \subset [0,\tau]$ with $\mu(A_\epsilon) \geq \tau - \epsilon$, where μ is Lebesgue measure, such that β is continuously differentiable on A_ϵ and its derivative is bounded away from 0 on A_ϵ.

Theorem 3.2. *Assume the model is given by* (3.1)*, where β_0 and β_1 are bounded and $\beta = \beta_1$ satisfies Assumption* 3.1. *Assume also e_t has density f on $\mathbb{R}$ which is locally bounded away from* 0 *and locally Riemann integrable,* $\sup_x E_x(|\beta_1(X_1)|^{-r}) < \infty$ *and* $\mathbf{E}(|e_t|^r) < \infty$ *for some $r > 0$. If*

$$\log\rho \stackrel{\text{def}}{=} \frac{1}{\tau}\int_0^\tau \log(|\beta_1(u)|)du > 0 \tag{3.3}$$

then $\{X_t\}$ is transient and $P_x(q^t|X_t| \to \infty) > 0$ for all $q > 1/\rho$ and all $x \in \mathbb{R}$.

Example 3.1. (Cf. Example 3.1 in CP.) Suppose X_t is as stated in Theorem 3.2 with $\beta_1(x) = a + b\cos(x)$ and f is the density function for e_t. If $\sup_x f(x) < \infty$ and $f(x) \leq f_1(|x|)$ where f_1 is monotone and integrable on $(0,\infty)$, then $\sup_x E_x(|a + b\cos(X_1)|^{-r}) < \infty$ for any $r < 1/2$. By Gradshteyn and Ryzhik (1980, eq. 4.226.1),

$$\frac{1}{2\pi}\int_0^{2\pi} \log(|a + b\cos(u)|)du = \begin{cases} \log(\frac{|b|}{2}), & \text{if } |a| \leq |b|, \\ \log(\frac{|a|+\sqrt{a^2-b^2}}{2}), & \text{if } |a| > |b|. \end{cases}$$

Thus (3.3) holds if $|b| > \max(2, |a|)$ or $|a| + \sqrt{a^2 - b^2} > 2$.

Now consider a ψ-irreducible chain on $\mathbb{R}^p$ defined by (1.1) with $\alpha(x) = A(x)x$, where $A(\cdot)$ is a $(p \times p)$-matrix valued function on $\mathbb{R}^p$. If (2.1) holds and there exists a constant matrix A_0 with an eigenvalue of modulus greater than 1 such that $\sup_x ||(A(x) - A_0)x|| < \infty$, then it is easy to show that $\{X_t\}$ is geometrically transient. (Cf. Tjøstheim (1990) for a related result.)

It is thus tempting to think that $\{X_t\}$ is transient if the maximum modulus of eigenvalues of $A(x)$ is larger than, and bounded away from, 1. This is not true, however, as is shown by the following simple threshold model on $\mathbb{R}^2$. Suppose

$$X_t = \begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix} = A(X_{t-1}) \begin{pmatrix} X_{t-1,1} \\ X_{t-1,2} \end{pmatrix} + \begin{pmatrix} e_{t,1} \\ e_{t,2} \end{pmatrix}, \tag{3.4}$$

where

$$A(X_{t-1}) = \begin{pmatrix} \phi_1 & 0 \\ 0 & \Phi_1 \end{pmatrix} 1_{\{X_{t-1,1} \le 0\}} + \begin{pmatrix} \phi_2 & 0 \\ 0 & \Phi_2 \end{pmatrix} 1_{\{X_{t-1,1} > 0\}}. \tag{3.5}$$

Note that the sub-process $\{X_{t,1}\}$ is itself an ordinary threshold model of order 1 (cf. Petruccelli and Woolford (1984)) and it drives the nonlinearity of $\{X_{t,2}\}$. If, for example, $\phi_1 < -1$, $0 < \phi_2 < 1$ and $\Phi_2 > 1$, then there is Φ_1 small enough such that $\{X_t\}$ is ergodic, not transient. (See CP Example 3.2.) We revisit this example after Theorem 3.4 below.

To generalize, suppose X_t may be represented as

$$X_t = \begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix} = \begin{pmatrix} A_1(X_{t-1,1})X_{t-1,1} + \gamma_1(e_t; X_{t-1,1}) \\ A_2(X_{t-1,1})X_{t-1,2} + \gamma_2(e_t; X_{t-1}) \end{pmatrix}, \tag{3.6}$$

where $X_{t,i} \in \mathbb{R}^{p_i}$, $A_i(\cdot)$ is a $p_i \times p_i$-matrix valued function on $\mathbb{R}^{p_1}$ and $p_1 + p_2 = p$. Note that $\{X_{t,1}\}$ is also a Markov chain while the nonlinearity of $\{X_{t,2}\}$ is driven entirely by $\{X_{t,1}\}$. If $\{X_{t,1}\}$ is known to be ergodic then stability/nonstability of $\{X_{t,2}\}$ will depend on the invariant distribution of $\{X_{t,1}\}$.

For the next theorem and example, $x' = (x_1', x_2')$ where $x_i \in \mathbb{R}^{p_i}$, $i = 1, 2$. Also, $E_{x_1}(\cdot)$ will refer to expectations for the process $\{X_{t,1}\}$ conditioned on $X_{0,1} = x_1$. If $\{X_t\}$ is ψ-irreducible then $\{X_{t,1}\}$ is ψ_1-irreducible for some ψ_1, as is easily checked. We need the following assumption.

Assumption 3.3. The complex vector $\nu \in \mathbb{C}^{p_2}$ satisfies $\psi(\{x\colon |\nu' x_2| > M, x_1 \in C\}) > 0$ for all $M < \infty$ and all $C \subset \mathbb{R}^{p_1}$ such that $\psi_1(C) > 0$.

Theorem 3.4. *Suppose $\{X_t\}$ is an aperiodic, ψ-irreducible Markov chain on $\mathbb{X} \subset \mathbb{R}^p$ satisfying* (3.6)*. Assume $\{X_{t,1}\}$ is aperiodic and geometrically ergodic, and has invariant distribution G_1. Fix $\nu \in \mathbb{C}^{p_2}$ and let $\varphi(x_1) = \min_{x_2} \frac{|\nu' A_2(x_1) x_2|}{|\nu' x_2|}$. Suppose* (2.1) *and Assumption* 3.3 *hold and $\sup_{x_1} E_{x_1}((\varphi(X_{1,1}))^{-r}) < \infty$ for some $r > 0$. If*

$$\log \rho \stackrel{\text{def}}{=} \int \log(\varphi(x_1)) G_1(dx_1) > 0 \tag{3.7}$$

then $\{X_t\}$ is transient and $P_x(q^t |\nu' X_{t,2}| \to \infty) > 0$ for all $q > 1/\rho$ and all $x \in \mathbb{X}$.

Example 3.2. (Cf. Example 3.2 in CP.) Let $p_1 = p_2 = 1$ and suppose $\{X_t\}$ is an aperiodic, μ_2-irreducible chain given by (3.4) and $\mathbf{E}(|e_{1,i}|^r) < \infty$ for $i = 1, 2$ and some $r > 0$. This defines a bivariate threshold model with coefficient matrix given in (3.5). Thus $\{X_{t,1}\}$ is the univariate TAR(1) model known to be geometrically ergodic if

$$\max(\phi_1, \phi_2, \phi_1\phi_2) < 1. \tag{3.8}$$

Assume (3.8) and let G_1 be the invariant distribution for $\{X_{t,1}\}$. If $\Phi_1\Phi_2 \neq 0$ then $\sup_{x_1} E_{x_1}(\varphi^{-r}(x_1)) < \infty$, and the condition

$$|\Phi_1|^{G_1(0)}|\Phi_2|^{1-G_1(0)} > 1 \tag{3.9}$$

implies (3.7) and therefore transience. In case $\Phi_1 = 0$, (3.9) implies $G_1(0) = 0$ and $|\Phi_2| > 1$. Hence $\psi_1(\{x_1 : x_1 \leq 0\}) = 0$ and $X_{t,2} = \Phi_2 X_{t-1,2} + e_{t,2}$ for all t large enough. In this case, $\{X_t\}$ is clearly transient. Similarly, (3.9) is also sufficient for transience if $\Phi_2 = 0$. In CP we showed that if (3.8) holds and $|\Phi_1|^{G_1(0)}|\Phi_2|^{1-G_1(0)} < 1$ the time series is geometrically ergodic.

Now we turn to nonlinear autoregressive time series of order $p > 1$, namely

$$\xi_t = a(\xi_{t-1}, \ldots, \xi_{t-p}) + c(e_t; \xi_{t-1}, \ldots, \xi_{t-p}) \tag{3.10}$$

for some $a\colon \mathbb{R}^p \to \mathbb{R}$. We assume also that

$$\limsup_{\|x\|\to\infty} \mathbf{E}(|c(e_1; x)|^r) < \infty, \qquad \text{for some } r > 0, \tag{3.11}$$

and that $X_t = (\xi_t, \ldots, \xi_{t-p+1})'$ is μ_p-irreducible (that is, Lebesgue irreducible). (See Cline and Pu (1998).) For $x = (x_1, \ldots, x_p)' \in \mathbb{R}^p$, $y \in \mathbb{R}$, let $\alpha(x) = (a(x), x_1, \ldots, x_{p-1})'$ and $\gamma(y; x) = (c(y; x), 0, \ldots, 0)'$, so that $\{X_t\}$ satisfies (1.1) and (2.1).

Example 3.3. When $a(x) = \beta_1(x)x_1 + \cdots + \beta_p(x)x_p$ with bounded "coefficient functions" $\beta_1, \ldots, \beta_p$, we have the FAR model of order p defined by Chen and Tsay (1993a):

$$\xi_t = \beta_1(\xi_{t-1}, \ldots, \xi_{t-p})\xi_{t-1} + \cdots + \beta_p(\xi_{t-1}, \ldots, \xi_{t-p})\xi_{t-p} + c(e_t; \xi_{t-1}, \ldots, \xi_{t-p}). \tag{3.12}$$

If $\sup_x |\sum_{i=1}^p (\beta_i(x) - b_i)x_i| < \infty$, where b_i are the coefficients of a nonstable linear AR(p) time series, then (3.12) is nonstable. (This is easily shown.)

Chen and Tsay (1993a) showed that (3.12) is geometrically ergodic if $|\beta_i(x)| \leq b_i$ for $i = 1, \ldots, p$ and $b_1 + \cdots + b_p < 1$. In fact, $\sup_x \sum_{i=1}^p |\beta_i(x)| < 1$ suffices (Chan and Tong (1985, 1986), Chan (1990) and An and Huang (1996)). We cannot, however, expect that $|\beta_i(x)| \geq b_i > 0$ for $i = 1, \ldots, p$ and $b_1 + \cdots + b_p > 1$ imply transience since there are stable linear time series satisfying these. Nevertheless, we have the following.

Theorem 3.5. *Consider the model* (3.12) *and let* $X_t = (\xi_t, \ldots, \xi_{t-p+1})'$*. Assume* $\{X_t\}$ *is* μ_p*-irreducible and* (3.11) *holds. Suppose* $b_1, \ldots, b_p$ *are all positive constants such that* $b_1 + \cdots + b_p > 1$ *and let* ρ *be the maximal root of* $z^p - b_1 z^{p-1} - \cdots - b_p = 0$*. Then any one of the following implies* $\{X_t\}$ *is transient and* $P_x(q^t|\xi_t| \to \infty) > 0$ *for all* $q > 1/\rho$ *and all* $x \in \mathbb{R}^p$*.*

(i) *For some* $M < \infty$, $\min_{1 \le i \le p} x_i > M$ *implies* $\beta_i(x) \ge b_i$, $i = 1, \ldots, p$.
(ii) *For some* $M < \infty$, $\max_{1 \le i \le p} x_i < -M$ *implies* $\beta_i(x) \ge b_i$, $i = 1, \ldots, p$.
(iii) *For some* $M < \infty$, $\max(\min_{1 \le i \le p}((-1)^i x_i), \min_{1 \le i \le p}((-1)^{i+1} x_i)) > M$ *implies* $(-1)^i \beta_i(x) \ge b_i$, $i = 1, \ldots, p$.

Our final result captures explosive behavior of nonlinear time series in a way useful for studying polynomial autoregressions.

Theorem 3.6. *Let* $\{\xi_t\}$ *be defined by* (3.10) *and let* $X_t = (\xi_t, \ldots, \xi_{t-p+1})'$*. Assume* $\{X_t\}$ *is* μ_p*-irreducible and* (3.11) *holds, and there exists* $M < \infty$, $b > 1$ *and* $d > b$ *so that*

$$a(x) \ge d\, a(z) \text{ whenever } x_j \ge b z_j \text{ for } j = 1, \ldots, p \text{ and } z_p > M. \tag{3.13}$$

Let $\rho \in (b, d)$ *and*

$$D_m = \{x\colon a(x) \ge \rho x_1, x_j \ge b x_{j+1}, \text{ for } j = 1, \ldots, p-1, \text{ and } x_p > m\}. \tag{3.14}$$

Assume also $\mu_p(D_m) > 0$ *for all* $m \ge M$*. Then* $\{X_t\}$ *is transient and* $P_x(q^t|\xi_t| \to \infty) > 0$ *for all* $q > 1/\rho$ *and all* $x \in \mathbb{R}^p$*.*

Example 3.4. For a particular example, suppose we have the model of Theorem 3.6 such that, for some i and j,

$$a(x) = \sum_{k=i}^{p} \sum_{l=j}^{p} a_{kl} x_k x_l$$

with $a_{ij} > 0$. (If $a_{ij} < 0$ we can consider the time series $\{-\xi_t\}$ instead.) Choose K so that $|a_{kl}|/|a_{ij}| \le K$ for $k \ge i, l \ge j$ and choose $b > \max(3,\, 2p^2 K)$. Thus, if $x_p > 0$ and $x_k \ge b x_{k+1}$ for $k = 1, \ldots, p-1$, then $a_{ij} x_i x_j / 2 \le a(x) \le 3 a_{ij} x_i x_j / 2$.

Let $d = b^2/3$. Then $d > b$ and it can be checked easily that (3.13) is satisfied for any M. Let $\rho \in (b, d)$. If $M \ge 2b^{i+j+2-p}/a_{ij}$, $x_k \ge b x_{k+1}$ for $k = 1, \ldots, p-1$ and $x_p > M$, then

$$a(x) > \frac{1}{2} a_{ij} x_i x_j > b^{p+2} x_p > \rho b^p x_p.$$

Clearly, D_m as defined in (3.14) has nonempty interior and hence $\mu_p(D_m) > 0$. Thus, the condition for geometric transience given in Theorem 3.6 holds and $P_x(q^t|\xi_t| \to \infty) > 0$ for all $q > 0$ and all x.

Tjøstheim (1990) considered a special case of this model with $a(x) = x_1x_2$. He speculated that a generalization would be possible by proving directly the transience of $\{X_{pt}\}$. We can now see that such a "multi-step" method is not necessarily required.

The method described above clearly can be modified to include linear and constant terms in $a(x)$. Also, it may be generalized for higher order polynomials of the form

$$a(x) = \sum_{i_j \le k_j \le p} a_{k_1\ldots k_d} x_{k_1} \cdots x_{k_d}, \qquad \text{where } a_{i_1\ldots i_d} \neq 0.$$

Chan and Tong (1994) use a related approach to show geometric transience in models for which $a(x)$ is a d-degree polynomial having a nontrivial x_1^d term, $d \ge 2$. Unfortunately, their generalization to all polynomial autoregression models is incomplete and thus the general result appears to be unresolved.

4. Proofs

We first reformulate a drift condition for transience given in Meyn and Tweedie (1993, Theorem 8.4.2) to identify conditions for geometric transience. Recall that $\{X_t\}$ is assumed to be ψ-irreducible on $\mathbb{X}$.

Lemma 4.1. *Assume $V : \mathbb{X} \to \mathbb{R}$ is nonnegative. Suppose $\psi(\{x\colon V(x) > M\}) > 0$ for all $M < \infty$. If*

$$\limsup_{V(x)\to\infty} E_x\Big(\Big(\frac{V(x)}{V(X_1)}\Big)^r\Big) < q^r, \qquad \text{for some } r > 0,\ q < 1, \tag{4.1}$$

then $\{X_t\}$ is transient and $P_x(q^tV(X_t) \to \infty) > 0$ for all $x \in \mathbb{X}$.

Proof. First note that, for all M large enough, both $C_M = \{x\colon V(x) \le M\}$ and C_M^c have positive ψ-measure. Also, $V^*(x) = 1 - (1 + V^r(x))^{-1}$ is nonnegative and bounded on $\mathbb{X}$. Clearly (4.1) implies the existence of $M < \infty$ such that $E_x(V^*(X_1)) - V^*(x) > 0$ for $x \in C_M^c$. By Theorem 8.4.2 of Meyn and Tweedie (1993), we conclude $\{X_t\}$ is transient.

Now we show the condition implies $q^tV(X_t) \to \infty$ with positive probability. Note that, if $b_1, b_2, \ldots$ are positive constants and $\mathbf{E}(Y) < a$, Jensen's inequality gives

$$\mathbf{E}\Big(\prod_{i=1}^m (1 - b_iY)_+\Big) \ge \prod_{i=1}^m (1 - b_ia)_+. \tag{4.2}$$

Choose $M < \infty$ and $\delta > 0$ such that

$$\sup_{V(x)>M} E_x\Big(\Big(\frac{V(x)}{V(X_1)}\Big)^r\Big) \le ((1-\delta)q)^r < 1.$$

Assume $X_0 = x$ almost surely with $V(x) > M$. For $n \geq 1$ let

$$Y_n = \Big(\frac{V(X_{n-1})}{(1-\delta)(V(X_n))}\Big)^r \qquad \text{and} \qquad W_n = \Big(\frac{V(x)}{(1-\delta)^n(V(X_n))}\Big)^r = Y_1 \cdots Y_n.$$

Thus, $W_{n-1} < 1$ implies $V(X_{n-1}) > (1-\delta)^{1-n}M$ and

$$\mathbf{E}\Big(Y_n 1_{\{W_{n-1}<1\}} \;\Big|\; X_{n-1}\Big) \leq q^r \qquad \text{a.s. } (P_x). \tag{4.3}$$

Using the Markov property and recursively applying the inequality (4.2) by way of (4.3),

$$\begin{aligned} P_x\left(W_1 < 1, \ldots, W_n < 1\right) &\geq E_x\Big(\prod_{i=1}^{n}(1-W_i)_+\Big) \\ &\geq E_x\Big(\prod_{i=1}^{n-1}(1-W_i)_+(1-q^r W_{n-1})_+\Big) \\ &\geq E_x\Big(\prod_{i=1}^{n-2}(1-W_i)_+(1-q^r W_{n-2})_+(1-q^{2r}W_{n-2})_+\Big) \\ &\geq \cdots \geq \prod_{i=1}^{n}(1-q^{ir}). \end{aligned}$$

Hence, for $V(x) > M$,

$$P_x(q^n V(X_n) \to \infty) \geq \lim_{n\to\infty} P_x\left(W_1 < 1, \ldots, W_n < 1\right) \geq \prod_{n=1}^{\infty}(1-q^{nr}) > 0.$$

Since $\psi(\{x\colon V(x) > M\}) > 0$ and $\{X_t\}$ is ψ-irreducible, we conclude $P_x(q^t V(X_t) \to \infty) > 0$ for all $x \in \mathbb{X}$.

Proof of Theorem 2.1. Let $r > 0$ and $\lambda \in \Lambda$ be such that (2.3) and (2.4) hold, and define $V_\lambda(x) = 1 + \lambda(x)V(x)1_{\{x\in D\}}$. Also $\psi(\{x\colon V_\lambda(x) > M\}) > 0$ for all $M < \infty$. By (2.3) and (2.4),

$$\begin{aligned} &\limsup_{V_\lambda(x)\to\infty} E_x\Big(\Big(\frac{V_\lambda(x)}{V_\lambda(X_1)}\Big)^r\Big) \\ &\leq \limsup_{\substack{V(x)\to\infty \\ x\in D}} \Big(E_x\Big(\Big(\frac{\lambda(x)V(x)}{\lambda(X_1)V(X_1)}\Big)^r 1_{\{X_1\in D\}}\Big) + E_x\Big((1+\lambda(x)V(x))^r 1_{\{X_1\notin D\}}\Big)\Big) \\ &< q^r < 1. \end{aligned}$$

So, by Lemma 4.1, $\{X_t\}$ is transient and $P_x(q^t V_\lambda(X_t) \to \infty) > 0$ for all $x \in \mathbb{X}$. Since $\lambda(x)$ is bounded, $P_x(q^t V(X_t)1_{\{X_t\in D\}} \to \infty) > 0$ as well.

Proof of Theorem 2.2. By Jensen's inequality, (2.7) implies (2.6). Now we show that (2.6) implies (2.4). Fix m, n and r such that the inequalities in (2.5) and (2.6) hold. By (2.5), Jensen's inequality and conditioning,

$$\sup_x E_x\Big(\log\Big(\frac{V(X_{j-1})}{V(X_j)}\Big)1_{\{X_j\in D\}}\Big)<\infty,\qquad\text{for every } j.$$

Thus, (2.6) implies that there exists $\delta>0$ such that

$$\limsup_{\substack{V(x))\to\infty\\ x\in D}} E_x\Big(\sum_{j=m}^{n}\log\Big(\delta+\frac{V(X_{j-1})}{qV(X_j)}\Big)1_{\{X_j\in D\}}\Big)<0. \tag{4.4}$$

By the proof of Lemma 4.1(i) of CP, (4.4) implies that for some $r_1>0$ and $\lambda\in\Lambda$,

$$\limsup_{\substack{V(x))\to\infty\\ x\in D}} E_x\Big(\Big(\frac{\lambda(x)}{\lambda(X_1)}\frac{V(x)}{qV(X_1)}\Big)^{r_1 1_{\{X_1\in D\}}}\Big)<1.$$

Then (2.4) follows.

Proof of Theorem 3.2. Let $\alpha(x)=\beta_1(x)x$ and $\gamma(e_1;x)=\beta_0(x)+e_1$. Clearly $\{X_t\}$ is μ-irreducible. Choose $q\in(1/\rho,1)$. In the proof of Theorem 3.4 of CP, we showed that $\mathrm{frac}(X_2/\tau)$ is asymptotically uniform in distribution as $|\alpha(x)|\to\infty$. As a result of that fact and the boundedness of $E_x(|\beta_1(X_1)|^{-r})$,

$$\lim_{|\alpha(x)|\to\infty} E_x(\log(|q\beta_1(X_2)|^{-1}))=-\frac{1}{\tau}\int_0^\tau\log(q|\beta_1(u)|)du<0. \tag{4.5}$$

By (4.5) and a result contained in the proof of Lemma 4.1(i) of CP, there exists $r_1\in(0,r)$ and a nonnegative function $\lambda(x)$, bounded and bounded away from 0, such that

$$\limsup_{|\alpha(x)|\to\infty} E_x\Big(\frac{\lambda(X_1)}{\lambda(x)}|q\beta_1(X_1)|^{-r_1}\Big)<1. \tag{4.6}$$

Now let $D=\mathbb{R}$ and $V(x)=(1+|\alpha(x)|^{r_1})/\lambda(x)$. Then (2.2) and (2.3) hold trivially. Choose $\epsilon\in(0,1-q)$ and L so that $1/L<\lambda(x)<L$. By (4.6) and the fact $\mathbf{E}(|e_1|^r)<\infty$,

$$\begin{aligned}\limsup_{V(x)\to\infty} E_x\Big(\frac{V(x)}{V(X_1)}\Big) &\le \limsup_{|\alpha(x)|\to\infty} L^2|\alpha(x)|^{r_1}P_x(|\beta_0(x)+e_1|>\epsilon|\alpha(x)|)\\ &\quad+\limsup_{|\alpha(x)|\to\infty}(1-\epsilon)^{-r_1}E_x\Big(\frac{\lambda(X_1)}{\lambda(x)}|\beta_1(X_1)|^{-r_1}1_{\{|\beta_0(x)+e_1|\le\epsilon|\alpha(x)|\}}\Big)\\ &<\Big(\frac{q}{1-\epsilon}\Big)^{r_1}<1.\end{aligned}$$

By Theorem 2.1, $\{X_t\}$ is transient and, since q and ϵ are arbitrary, $P_x(q^t|\alpha(X_t)| \to \infty)$ for all $q > 1/\rho$ and all x. Furthermore, since β_1 is bounded, we conclude $P_x(q^t|X_t| \to \infty)$ for all $q > 1/\rho$ and all x.

Proof of Theorem 3.4. This argument parallels that for CP (Theorem 3.5) and is a sophisticated use of the generalized m-step approach. In the interest of space the proof is sketched and the reader is referred to the authors' complete technical version.

Choose $q \in (1/\rho, 1)$. By Meyn and Tweedie (1993, Theorem 15.0.1), there exists $V_1 \colon \mathbb{R}^{p_1} \to [1, \infty)$ satisfying $\int V_1(x_1)G_1(dx_1) < \infty$ and the drift condition for geometric ergodicity. We can then find $\delta > 0$, $\epsilon \in (0, 1)$ and $s = (1-\epsilon)r$ and define $h(x_1) = E_{x_1}\left((\delta + 1/\varphi(X_{1,1}))^s\right)(V_1(x_1))^\epsilon$ so that $\sup_{x_1} E_{x_1}(h(X_{n,1}))/(V_1(x_1))^\epsilon < q^s$, for some $n \geq 1$, again by Meyn and Tweedie (1993, Theorem 15.0.1). Next, define

$$V(x) = \Big(\frac{(\delta+|\nu' A(x_1)x_2|)^s}{\prod_{j=0}^{n-1} E_{x_1}(h(X_{j,1}))}\Big)^{\frac{1}{n+2}},$$

$$R_1(x, X_1) = \Big(\frac{\delta+|\nu' A(x_1)x_2|}{(\delta+|\nu' A(X_{1,1})X_{1,2}|)(\delta+1/\varphi(X_{1,1}))}\Big)^s \text{ and } R_2(x, X_1) = \frac{(\delta+1/\varphi(X_{1,1}))^s}{E_{x_1}((\delta+1/\varphi(X_{1,1}))^s)}.$$

Using (2.1) we can show $\limsup_{V(x)\to\infty} E_x\left(R_1(x, X_1)\right) \leq 1$. By Hölder's inequality and the fact that $\{X_{t,1}\}$ is Markov, we conclude

$$\begin{aligned}
&\limsup_{V(x)\to\infty} E_x\Big(\frac{V(x)}{V(X_1)}\Big) \\
&= \limsup_{V(x)\to\infty} E_x\Big(\Big(R_1(x, X_1)R_2(x, X_1)\frac{E_{X_{1,1}}(h(X_{n,1}))}{(V_1(x_1))^\epsilon}\prod_{j=1}^{n-1}\frac{E_{X_{1,1}}(h(X_{j,1}))}{E_{x_1}(h(X_{j,1}))}\Big)^{\frac{1}{n+2}}\Big) \\
&\leq \limsup_{V(x)\to\infty} \Big(E_x(R_1(x, X_1))\Big)^{\frac{1}{n+2}} \quad \sup_{x_1}\Big(\frac{E_{x_1}(h(X_{n,1}))}{(V_1(x_1))^\epsilon}\Big)^{\frac{1}{n+2}} < q^{\frac{s}{n+2}}.
\end{aligned}$$

This verifies (4.1).

Assumption 3.3 ensures that $\psi(\{x \colon V(x) > M\}) > 0$ for any $M < \infty$. By Lemma 4.1, $\{X_t\}$ is transient and $P_x(q^{st/(n+2)}V(X_t) \to \infty) > 0$ for all x. From this, it follows that $P_x(q^t|\nu' X_t| \to \infty) > 0$ for all x and all $q > 1/\rho$.

Proof of Theorem 3.5. Here we have an expedient use of the directional method. Let

$$B = \begin{pmatrix} b_1 & b_2 & \cdots & b_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}$$

and assume (i). Since $b_p \neq 0$, B is irreducible. Also, the elements of B are nonnegative. Thus B has a real eigenvalue equal to ρ, the maximal root of

$z^p - b_1 z^{p-1} - \cdots - b_p = 0$, and there corresponds a left eigenvector v' with positive components (Rao (1973, p.46)). It is easy to show that in this case $\rho > 1$ since $\sigma = \sum_{i=1}^p b_i > 1$. Let $m = \min(b_1, \ldots, b_p) > 0$. There exists $\epsilon > 0$ such that $(1 - \frac{\epsilon}{m})\sigma > 1$ and $\rho - \epsilon > 1$. Let $a(x) = \beta_1(x)x_1 + \cdots + \beta_p(x)x_p$. Let $D = \{x\colon \min_{1\le i\le p} x_i > M\}$. Then for $x \in D$, we have

$$a(x) \ge b_1 x_1 + \cdots + b_p x_p \ge \max(\sigma M, m\|x\|). \tag{4.7}$$

We may assume that $||v|| = 1$. Now define $V(x) = |v'x|$. Then (2.2) is satisfied. Recall that $\gamma(e_1; x) = (c(e_1; x), 0, \ldots, 0)'$. If $X_0 = x \in D$ and $|c(e_1; x)| \le \epsilon V(x)$ then $|c(e_1; x)| \le \epsilon\|x\|$ and, by (4.7), $a(x) + c(e_1; x) \ge \big(1 - \frac{\epsilon}{m}\big)\, a(x) \ge \big(1 - \frac{\epsilon}{m}\big)\, \sigma M > M$. Hence $X_1 \in D$ when $X_0 \in D$ and $|c(e_1; x)| \le \epsilon V(x)$. Also,

$$V(X_1) \ge |v'A(x)x| - ||\gamma(e_1; x)|| \ge v'Bx - |c(e_1; x)| \ge (\rho - \epsilon)V(x), \tag{4.8}$$

where $A(x)$ is the matrix obtained from B by substituting $\beta_i(x)$ for b_i. Take $r_1 \in (0, r)$. We see from (3.11) and (4.8) that (2.3) holds with r replaced by r_1, and that

$$\begin{aligned}
&\limsup_{\substack{V(x)\to\infty \\ x\in D}} E_x\Big(\Big(\frac{1+V(x)}{1+V(X_1)}\Big)^{r_1} 1_{\{X_1\in D\}}\Big) \\
&\le \limsup_{\substack{V(x)\to\infty \\ x\in D}} E_x\Big(\Big(\frac{V(x)}{V(X_1)}\Big)^{r_1} 1_{\{|c(e_1;x)|\le\epsilon V(x)\}}\Big) + \limsup_{|v'x|\to\infty} |v'x|^{r_1}\mathbf{P}(|c(e_1;x)| > \epsilon|v'x|) \\
&= (\rho-\epsilon)^{-r_1} < 1.
\end{aligned} \tag{4.9}$$

Hence (2.4) is also satisfied (with $\lambda(x) = 1$). By Theorem 2.1 and the arbitrary choice of ϵ, the process is transient and $P_x(q^t|v'X_t| \to \infty) > 0$ for all x and all $q > 1/\rho$. Since the components of v are positive and (4.7) holds, we also have $P_x(q^t|a(X_t)| \to \infty) > 0$ for all x and all $q > 1/\rho$. By (3.11), $P_x(|\xi_1| > |a(x)|/2) > 1/2$ for $|a(x)|$ large enough and thus $P_x(q^t|\xi_t| \to \infty) > 0$ for all $q > 1/\rho$ and all x, proving the sufficiency of (i). The sufficiency of (ii) can be proved analogously.

To prove that (iii) is sufficient, we first define $D_1 = \{x\colon \min_{1\le i\le p}((-1)^i x_i) > M\}$ and $D_2 = \{x\colon \min_{1\le i\le p}((-1)^{i+1} x_i) > M\}$. Note that for $x \in D_1$, we have $|x_i| = (-1)^i x_i$ and $\beta_i(x)x_i > b_i|x_i|$ for each $i = 1, \ldots, p$, and hence $a(x) > \sigma M > M$ and $\alpha(x) \in D_2$. Similarly, if $x \in D_2$ then $\beta_i(x)x_i < -b_i|x_i|$ $a(x) < -\sigma M < -M$ and $\alpha(x) \in D_1$. Consequently, if $x \in D = D_1 \cup D_2$, then $|a(x)| > \sum_{i=1}^p b_i|x_i|$. Let $|x| = (|x_1|, \ldots, |x_p|)'$ for $x = (x_1, \ldots, x_p)'$. Define $V(x) = v'|x|$. Therefore, $x \in D$ implies

$$V(\alpha(x)) = v'|\alpha(x)| > v_1\Big(\sum_{i=1}^p b_i|x_i|\Big) + \sum_{i=2}^p v_i|x_{i-1}| = v'B|x| = \rho v'|x| = \rho V(x).$$

The rest of the proof is the same as the proof for (i).

Proof of Theorem 3.6. Fix M and ρ to satisfy the assumptions. Choose δ so that $d > \rho + \delta > \rho - \delta > b$ and let $D = D_M$ and $V(x) = \delta|x_1|$. Then (2.2) holds. If $X_0 = x \in D$ and $|c(e_1; x)| \le V(x)$, then from (3.13) and (3.14) we have $a(x) + c(e_1; x) > b|x_1|$,

$$a(\alpha(x) + \gamma(e_1; x)) \ge d\, a(x) > (\rho + \delta)a(x) > \rho a(x) + \delta\rho x_1 \ge \rho(a(x) + c(e_1; x)), \tag{4.10}$$

and hence $X_1 = \alpha(x) + \gamma(e_1; x) \in D$. Therefore, if $r_1 \in (0, r)$, (3.11) and (4.10) imply (2.3) and, as in the proof of (4.9),

$$\limsup_{\substack{V(x)\to\infty \\ x\in D}} E_x\Big(\Big(\frac{1+V(x)}{1+V(X_1)}\Big)^{r_1} 1_{\{X_1\in D\}}\Big) \le \frac{1}{(\rho-\delta)^{r_1}} < 1.$$

Hence, (2.4) is satisfied and the conclusion holds by Theorem 2.1.

References

An, H. Z. and Huang, F. C. (1996). The geometrical ergodicity of nonlinear autoregressive models. *Statist. Sinica* **6**, 943-956.

Bhattacharya, R. and Lee, C. (1995). Ergodicity of nonlinear first order autoregressive models. *J. Theory Probab.* **8**, 207-219.

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

Chan, K.-S. and 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. and Tong, H. (1986). On estimating thresholds in autoregressive models. *J. Time Ser. Anal.* **7**, 179-190.

Chan, K.-S. and Tong, H. (1994). A note on noisy chaos. *J. Roy. Statist. Soc. Ser. B* **56**, 301-311.

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

Chen, R. and Tsay, R. S. (1993a). Functional-coefficient autoregressive models. *J. Amer. Statist. Assoc.* **88**, 298-308.

Chen, R. and Tsay, R. S. (1993b). Nonlinear Additive ARX models. *J. Amer. Statist. Assoc.* **88**, 955-967.

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. (1999a). Geometric ergodicity of nonlinear time series. *Statist. Sinica* **9**, 1103-1118.

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

Cline, D. B. H. and Pu, H. H. (1999c). Stability of threshold-like ARMA time series. Statistics Department., Texas A&M University.

Gradshteyn, I. S. and Ryzhik, I. M. (1980). *Table of Integrals, Series and Products.* Academic Press, Orlando, FL.

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

Kendall, D. G. (1959). Unitary dilations of Markov transition operators and the corresponding integral representations for transition-probability matrices. In *Probab. Statist.* (Edited by U. Granander), 139-161. Almqvist and Wiksell, New York.

Kingman, J. F. C. (1963). The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* **13**, 337-358.

Lu, Z. (1998a). Geometric ergodicity of a general ARCH type model with applications to some typical models. In *Adv. Operations Research Systems Engineering* (Edited by J. Gu, G. Fan and S. Wang), 76-86. Global-Link Informatics Ltd.

Lu, Z. (1998b). On the geometric ergodicity of a non-linear autoregressive model with an autoregressive conditional heteroscedastic term. *Statist. Sinica* **8**, 1205-1217.

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

Petruccelli, J. D. and Woolford, S. W. (1984). A threshold AR(1) model. *J. Appl. Probab.* **21**, 587-611.

Pham D. T. (1985). Bilinear Markovian representation and bilinear models. *Stoch. Proc. Appl.* **20**, 295-306.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* 2nd edition. Wiley, New York.

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

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

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

Vere-Jones, D. (1962). Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford* **13**, 7-28.

Department of Statistics, Texas A&M University, College Station TX 77843-3143, U.S.A.

E-mail: dcline@stat.tamu.edu

Department of Statistics, Texas A&M University, College Station TX 77843-3143, U.S.A.

E-mail: pu@stat.tamu.edu

(Received September 1999; accepted June 2000)