GEOMETRIC ERGODICITY OF A DOUBLY STOCHASTIC TIME SERIES MODEL

SEAN P. MEYN AND LEI GUO

University of Illinois and Institute of Systems Science, Beijing

First version received June 1990

Abstract. We demonstrate that a large class of doubly stochastic time series models are geometrically ergodic, and hence admit second-order stationary solutions.

We also establish a version of the strong law of large numbers, the law of the interated logorithm and the central limit theorem for the stochastic processes under consideration.

Keywords. Doubly stochastic models; bilinear models; stationarity; Harris recurrence; geometric ergodicity.

1. INTRODUCTION AND DEFINITIONS

In this paper we consider the time series model

$$y_{k+1} = \theta_k^{\mathrm{T}}\varphi_k + v_{k+1} \tag{1}$$

where $\boldsymbol{v}$ is an independent and identically distributed (i.i.d.) process on $\mathbb{R}$, $\boldsymbol{\theta}$ is a stochastic process taking values in $\mathbb{R}^m$, both defined on a probability space $(\Omega, \mathcal{F}, P)$, and $\boldsymbol{\varphi}$ is the regression sequence $\varphi_k^{\mathrm{T}} = (y_k, \ldots, y_{m-1})$, with initial condition $\varphi_0 \in \mathbb{R}^m$ given.

The process y defined by (1) is called a doubly stochastic autoregressive process of order m (DSAR (m)) (Tjøstheim, 1986). Tjøstheim shows that this class of models contains a broad range of popular nonlinear time series models. A survey of recent results may be found in the introduction of Pourahmadi (1986).

Pourahmadi also presents sufficient conditions under which a stationary second-order process satisfying (1) exists in the scalar case. The approach taken is to impose conditions on the stochastic process $\{Y_k \triangleq \log \theta_k^2, k \geqslant 1\}$. Under either of the hypotheses (i) $\boldsymbol{Y}$ is the output of a finite dimensional linear system driven by white noise and (ii) $\boldsymbol{Y}$ is a stationary Gaussian process, it is shown that a sufficiently strong moment condition on $\boldsymbol{Y}$ implies the existence of a second-order stationary solution to the time series model under consideration.

Recently, Karlsen (1990) used a decoupling inequality due to Klein *et al.* (1981) to obtain an elegant stability proof (and hence also a second-order stationary solution for the model) for a class of stationary Gaussian parameter processes.

0143-9782/93/01 93–108 JOURNAL OF TIME SERIES ANALYSIS Vol. 14, No. 1.

The goal of this paper is to find explicit moment conditions on the process $\boldsymbol{\theta}$ which ensure the existence of an L_p-bounded solution to (1) for a given $p \geqslant 2$. We are principally interested in the existence of stationary solutions, and conditions under which a stationary solution is unique. Our approach is to construct a Markov chain $\boldsymbol{\Phi}$ which can serve as a nonlinear state process for (1) and apply recent results from the ergodic theory of Markov chains on topological state spaces.

A barrier in any analysis of this kind is stability. In the stability proof below we construct a test function on the process $\boldsymbol{\Phi}$ which can be seen as a stochastic generalization of the Lyapunov functions used in the stability theory of, for instance, La Salle and Lefschetz (1961). The existence of the appropriate test function will be seen to imply stability of the model and the existence of a unique invariant probability for $\boldsymbol{\Phi}$.

This approach takes us far beyond our goal. We find that the underlying distributions of the process $\boldsymbol{y}$ will converge to the unique stationary solution at a geometric rate for arbitrary initial conditions of the model. Furthermore, the central limit theorem, the law of large numbers and the law of the iterated logorithm are obtained as a by-product. These sample path results are significant since it is unlikely that tight bounds on the moments of $\boldsymbol{y}$ can be obtained analytically, but these results show that they can be estimated by simulations.

There is at present a great deal of interest in the application of Markov chain techniques to the analysis of time series models and, in particular, conditions which ensure geometric ergodicity of the model subject to analysis (Chan, 1986; Mokkadem, 1987; Diebolt and Guégan, 1990; Tjøstheim, 1990). It is believed that the techniques introduced in this paper are sufficiently general that they may be extended to other stochastic models such as are commonly found throughout the applied sciences. Some extensions of the results presented here may be found in Meyn (1991). Further extensions of the theory presented herein, as well as a number of examples from time series, queueing theory and estimation and control theory will be forthcoming (Meyn and Tweedie, 1991).

We assume in this paper that $\boldsymbol{\theta}$ itself is the output of a stable linear state space model. We begin with the scalar case so that the joint process $(\boldsymbol{y}, \boldsymbol{\theta})$ can be expressed as

$$\theta_{k+1} = \alpha\theta_k + e_{k+1} \qquad |\alpha| < 1, \tag{2}$$

$$y_{k+1} = \theta_k y_k + v_{k+1}. \tag{3}$$

The original formulation (1) will be treated in Section 6. We remark that more general parameter models, including autoregressive moving-average (ARMA) models, can be treated using the same methods. See Meyn and Guo (1990) for a set of related results for adaptive systems.

We henceforth assume that the model (2), (3) satisfies the following conditions.

(A1) $w \triangleq (e, v)^{\mathrm{T}}$ is i.i.d. and independent of (θ_0, y_0), and e and v are mutually independent.

(A2) The distribution μ_w of w_k, with $k \geq 1$, is nonsingular with respect to Lebesgue measure.

(A3) For some $p \geq 2$,

$$E(|v_1|^p) < \infty \qquad \gamma^p \triangleq E\left\{\exp\left(\frac{p}{1-|\alpha|}|e_1| - p\right)\right\} < 1.$$

Under these conditions, the joint process $\Phi_k = \binom{\theta_k}{y_k}$, $k \geq 0$, is a Markov chain with stationary transition probabilities P^k, $k \geq 1$, defined so that for any bounded measurable function f on the state space $X \triangleq \mathbb{R}^2$

$$E[f(\Phi_{k+n})|\sigma\{\Phi_0, \ldots, \Phi_n\}] = \int_X P^k(\Phi_n, dy)f(y) \text{ almost surely (a.s.).}$$

It is well known that a stationary solution to (3) will exist if the existence of an invariant probability can be established for Φ, i.e. a probability π on $\mathcal{B}(X)$ satisfying the defining property

$$\pi(B) = \int \pi(dx)P(x, B) \qquad B \in \mathcal{B}(X).$$

We say that a Markov chain is geometrically ergodic if an invariant probability exists and is unique and there exists $\rho < 1$ and a function $R: X \to \mathbb{R}_+$ such that

$$|P^k(x, A) - \pi(A)| \leq R(x)\rho^k \qquad x \in X,\ A \in \mathcal{B}(X),\ k \in \mathbb{Z}_+.$$

We state here the main results of the paper. Our first result establishes geometric ergodicity of the model.

THEOREM 1.1. *Suppose that conditions (A1)–(A3) hold. Then Φ is geometrically ergodic, and hence possesses a unique invariant probability π. For every initial condition $x \in X$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} y_k^2 = \int y^2 d\pi \qquad \text{a.s. } [P_x]$$

$$|E_x(y_k^2) - \int y^2 d\pi| \leq M(x)\rho^k \qquad k \geq 0$$

where M is a continuous function on X and $0 < \rho < 1$.

In addition, we obtain a version of the central limit theorem and the law of the iterated logorithm.

THEOREM 1.2. *Suppose that conditions (A1)–(A3) hold. Then the limit*

$$\omega^2 = \lim_{n\to\infty} \frac{1}{n} E_\pi\left(\left[\sum_{k=1}^{n} \{y_k - E_\pi(y_k)\}\right]^2\right) \tag{4}$$

exists and is finite. The quantity ω^2 may also be expressed as

$$\omega^2 = \sum_{k=-\infty}^{\infty} E_\pi[\{y_0 - E_\pi(y_0)\}(y_k - E_\pi(y_k))\}] \tag{5}$$

where the sum converges absolutely. The following limits hold for each initial condition.

(i)

$$\lim_{n\to\infty} P_x\left[\frac{1}{(n\omega^2)^{1/2}} \sum_{k=1}^{n} \{y_k - E_\pi(y_k)\} \leq t\right] = \int_{-\infty}^{t} \frac{1}{2\pi} \exp\left(-\frac{x^2}{2}\right) dx$$

(ii) The limit inifimum and limit supremum of the normalized sum

$$\frac{1}{\{2\omega^2 n \log\log(n)\}^{1/2}} \sum_{k=1}^{n} \{y_k - E_\pi(y_k)\}$$

are respectively −1 and +1 with probability one.

In addition to these results, we can establish geometric mixing for the process y. For details the reader is referred to Meyn and Tweedie (1992).

In the next section we establish a number of results for the Markov chain $\mathbf{\Phi}$ which will lead to a proof of these results.

2. ANALYSIS OF A MARKOVIAN STATE PROCESS

Let μ_w be decomposed as

$$\mu_w = \lambda\mu_{ac} + (1 - \lambda)\mu_\perp \tag{6}$$

where the probability μ_{ac} is absolutely continuous with respect to Lebesgue measure, the probability $\mu_\perp$ is singular with respect to Lebesgue measure and $0 < \lambda \leq 1$. Define the Markov transition functions P and T on $(X, \mathcal{B}(X))$ for $(\theta, y) \in X$ and a measurable function $f: X \to \mathbb{R}$ by

$$Pf(\theta, y) = \int f(\alpha\theta + e_1, \theta y + v_1)\, d\mu_w(e_1, v_1)$$

$$Tf(\theta, y) = \int f(\alpha\theta + e_1, \theta y + v_1)\, d\mu_{ac}(e_1, v_1).$$

It is easy to check that $P = P^1$ is the Markov transition function for the Markov chain $\mathbf{\Phi}$, and that for all $x \in X$, $A \in \mathcal{B}(X)$ we have

$$P(x, A) \geq \lambda T(x, A) \tag{7}$$

The following result shows that in fact λT is a continuous component (Tuomineu and Tweedie, 1979; Meyn and Tweedie, 1992, 1991) of the Markov transition function P.

LEMMA 2.1. *The Markov transition function T has the strong Feller property. That is, for any bounded measurable function $f: X \to \mathbb{R}$, the function Tf is continuous.*

PROOF. If $f: X \to \mathbb{R}$ is bounded and measurable, then it follows from a result of Rudin (1974) that, if $(\theta_k, y_k) \to (\theta_\infty, y_\infty)$,

$$f(\alpha\theta_k + e_1, \theta_k y_k + v_1) \to f(\alpha\theta_\infty + e_1, \theta_\infty y_\infty + v_1) \text{ as } k \to \infty$$

for almost every $(e_1, v_1) \in \mathbb{R}^2[\mu^{leb}]$. Hence by the dominated convergence theorem and the definition of T,

$$\lim_{k\to\infty} Tf(\theta_k, y_k) = \lim_{k\to\infty} \int f(\alpha\theta_k + e_1, \theta_k y_k + v_1)\, d\mu_{ac}(e_1, v_1)$$
$$= Tf(\theta_\infty, y_\infty)$$

which is the desired result. ∎

Lemma 2.1, together with the results of Tuominen and Tweedie (1979) and Meyn and Tweedie (1992), allows an application of the rich theory of irreducible Markov chains, as described in Nummelin (1984) for instance, which is based upon the notion of status sets (Tweedie, 1976). A particular instance of a status set is a small set (Nummelin, 1984), where we call a set $S \in \mathcal{B}(X)$ *small* if there exists a non-trivial measure ν on $\mathcal{B}(X)$ and $l \in \mathbb{Z}_+$ such that

$$P^l(x, B) \geq \nu(B) \qquad B \in \mathcal{B}(X), x \in S.$$

We say that $\mathbf{\Phi}$ is ν-irreducible (or just irreducible if the specific irreducibility measure ν is unimportant) if

$$G(x, B) \triangleq \sum_{k=1}^{\infty} P^k(x, B) > 0$$

for every $x \in X$ and every set $B \in \mathcal{B}(X)$ of positive ν-measure. An irreducible Markov chain can be further classified as either periodic or aperiodic, just as in the discrete state space case (Nummelin, 1984).

Here we establish the existence of a family of small sets for the Markov chain $\mathbf{\Phi}$ and prove that the chain is irreducible.

LEMMA 2.2. *Suppose that conditions (A1) and (A2) hold for the model (2), (3). Then*

(i) there exists $x^* \in X$, $(e^*, v^*) \in \operatorname{supp} \mu_w$, such that, when $(e_k, v_k) \equiv (e^*, v^*)$ for all k, then for all initial conditions $\Phi_k \to x^*$ as $k \to \infty$;

(ii) the Markov chain $\mathbf{\Phi}$ is φ-irreducible with $\varphi(\cdot) \triangleq T(x^*, \cdot)$;

(iii) every compact subset of X is small, and hence $\mathbf{\Phi}$ is aperiodic.

PROOF. From condition (A3) and Jensen's inequality it follows that $\exp[pE\{|e_1|/(1 - |\alpha|)\}] < \exp(p)$ and hence $E(|e_1|) < 1 - |\alpha|$. This implies the existence of $(e^*, v^*) \in \operatorname{supp} \mu_w$ with $|e^*| < 1 - |\alpha|$. If, with this choice of (e^*, v^*), we set $(e_k, v_k) = (e^*, v^*)$ for all $k \geq 1$ then, by (2),

$$\theta^* \triangleq \lim_{k\to\infty} \theta_k = \frac{e^*}{1-\alpha} \tag{8}$$

for every initial condition $\Phi_0 \in X$. Hence, by (1),

$$y^* \triangleq \lim_{k\to\infty} y_k = \frac{(1-\alpha)v^*}{1-\alpha-e^*}, \tag{9}$$

and this proves (i). To prove (ii), let $A \in \mathcal{B}(X)$ have positive φ-measure. By Lemma 2.1 there exists an open set N containing x^* such that $\inf_{x\in N} P(x, A) > 0$. Also, by (i), $G(y, N) > 0$ for all $y \in X$ which shows that

$$G(y, A) \geq \int_N G(y, dz)P(z, A) > 0,$$

proving (ii).

Since $\mathbf{\Phi}$ is φ-irreducible, there exists a small set $A \in \mathcal{B}(X)$ with $\varphi(A) > 0$, and hence an open set N containing x^* with $\inf_{x\in N} P(x, A) > 0$. It follows from Proposition 2.11 of Nummelin (1984) that N is also small.

Let $F \subseteq X$ be compact. A slight strengthening of the results (8) and (9) may be used to show that, for some $k \in \mathbb{Z}_+$, $P^k(x, N) > 0$ for every $x \in F$. Since P has the Feller property, the function $P^k(\cdot, N)$ is lower semi-continuous (Cogburn, 1975) and hence $\inf_{x\in F} P^k(x, N) > 0$. Since N is small, Proposition 2.11 of Nummelin (1984) may be applied once more to show that F is small. ∎

3. L^p STABILITY

Here we establish stability of the nonliner difference equations (2), (3).

PROPOSITION 3.1. *Suppose that conditions (A1)–(A3) hold. Then for every (deterministic) initial condition*

$$\limsup_{k\to\infty} \|y_k\|_p \leq \frac{\|v_1\|_p}{1-\gamma} \exp\left(\frac{1}{1-|\alpha|}\right)$$

where p and γ are defined in condition (A3) and $\|x\|_p \triangleq E(|x|^p)^{1/p}$.

To prove this result we first expand the representation (3) and then make a number of estimates based upon this expansion.

The following equality is obtained by iterating equation (3):

$$y_{k+1} = \sum_{j=1}^{k}\left(\prod_{i=j}^{k} \theta_i\right)v_j + \left(\prod_{i=0}^{k} \theta_i\right)y_0 + v_{k+1}. \tag{10}$$

First let us consider the term $\prod_{i=j}^{k} \theta_i$.

LEMMA 3.1. *We have the bound*

$$\left|\prod_{i=j}^{k} \theta_i\right| \leq \exp\{-(k-j+1)\} \exp\left(\frac{|\alpha|}{1-|\alpha|}\,|\theta_{j-1}|\right) \exp\left(\frac{1}{1-|\alpha|} \sum_{i=j}^{k} |e_i|\right) \quad (j \geq 1).$$

PROOF. We have, by the estimate $x \leq \exp(-1)\exp(x)$,

$$\left|\prod_{i=j}^{k} \theta_i\right| \leq \exp\{-(k-j+1)\}\left(\prod_{i=j}^{k} \exp|\theta_i|\right)$$
$$= \exp\{-(k-j+1)\} \exp\left(\sum_{i=j}^{k} |\theta_i|\right). \tag{11}$$

But since, by (2),

$$\sum_{i=j}^{k} |\theta_i| \leq |\alpha| \sum_{i=j}^{k} |\theta_i| + |\alpha||\theta_{j-1}| + \sum_{i=j}^{k} |e_i|,$$

we have

$$\sum_{i=j}^{k} |\theta_i| \leq \frac{|\alpha|}{1-|\alpha|}\,|\theta_{j-1}| + \frac{1}{1-|\alpha|} \sum_{i=j}^{k} |e_i|. \tag{12}$$

Equations (11) and (12) imply the result. ■

In order to apply (10) and Lemma 3.1 to infer the L^p boundedness of y, we shall require the following result.

LEMMA 3.2. *Under conditions (A1)–(A3) we have for all* $k \geq 0$

$$E\left\{\exp\left(\frac{p}{1-|\alpha|}\,|\theta_k|\right)\right\} \leq \exp\left(\frac{p|\alpha|^k}{1-|\alpha|}\,|\theta_0|\right) \exp\left(\frac{p}{1-|\alpha|}\right)$$

PROOF. By (2) we have

$$|\theta_k| \leq |\alpha||\theta_{k-1}| + |e_k|$$
$$\leq |\alpha|^k|\theta_0| + \sum_{i=1}^{k} |\alpha|^{k-i}|e_i|$$

Hence by condition (A1)

$$E\left\{\exp\left(\frac{p}{1-|\alpha|}\,|\theta_k|\right)\right\} \leq \exp\left(\frac{p|\alpha|^k}{1-|\alpha|}\,|\theta_0|\right) \prod_{i=1}^{k} E\left\{\exp\left(\frac{p|\alpha|^{k-i}}{1-|\alpha|}\,|e_i|\right)\right\}. \tag{13}$$

By Jensen's inequality and condition (A3) we have

$$\sum_{i=1}^{k} \log E\left\{\exp\left(\frac{p|\alpha|^{k-i}}{1-|\alpha|}\,|e_i|\right)\right\} \leq \sum_{i=1}^{k} |\alpha|^{k-i} \log E\left\{\exp\left(\frac{p}{1-|\alpha|}\,|e_i|\right)\right\} \leq p \sum_{i=0}^{\infty} |\alpha|^i$$
$$= \frac{p}{1-|\alpha|}$$

and this combined with (13) implies that, for all $k \geqslant 1$,

$$E\left\{\exp\left(\frac{p}{1-|\alpha|}|\theta_k|\right)\right\} \leqslant \exp\left(\frac{p|\alpha|^k}{1-|\alpha|}|\theta_0|\right)\exp\left(\frac{p}{1-|\alpha|}\right) \qquad \blacksquare$$

PROOF OF PROPOSITION 3.1 From (10) and condition (A1) we have

$$\|y_k\|_p \leqslant \left(1+\sum_{i=1}^{k}\left\|\prod_{i=j}^{k}\theta_i\right\|_p\right)\|v_1\|_p + \left\|\prod_{i=0}^{k}\theta_i\right\|_p |y_0|.$$

By Lemmas 3.1 and 3.2,

$$\left\|\prod_{i=j}^{k}\theta_i\right\|_p^p = E\left(\prod_{i=j}^{k}|\theta_i|^p\right) \leqslant \gamma^{(k-j+1)p}E\left\{\exp\left(\frac{p|\alpha|}{1-|\alpha|}|\theta_{j-1}|\right)\right\}$$

$$\leqslant \gamma^{(k-j+1)p}\exp\left(\frac{p}{1-|\alpha|}\right)\exp\left(\frac{p|\alpha|^{j-1}}{1-|\alpha|}|\theta_0|\right)$$

from which the result follows. $\blacksquare$

4. PROOFS OF THE MAIN RESULTS USING A STOCHASTIC LYAPUNOV FUNCTION

The proof of Theorem 1.1 and Theorem 1.2 depends on the existence of a certain test function on the state process (θ, y), which we construct in this section.

The following three lemmas allow us to construct the appropriate Lyapunov–Foster test function for the process $\mathbf{\Phi}$. We first bound the process y.

By (10) and Lemma 3.1 we have for all $k \in \mathbb{Z}_+$

$$y_{k+1}^2 \leqslant 3\left\{\sum_{j=1}^{k}|v_j|\exp\left(\frac{|\alpha|}{1-|\alpha|}|\theta_{j-1}|\right)\prod_{i=j}^{k}\exp\left(\frac{1}{1-|\alpha|}|e_i|-1\right)\right\}^2$$

$$+ 3|v_{k+1}|^2 + 3\theta_0^2 y_0^2\exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_0|\right)\prod_{i=1}^{k}\exp\left(\frac{2}{1-|\alpha|}|e_i - 2\right). \qquad (14)$$

Let U_{k+1} denote the first squared term in (14):

$$U_{k+1} \triangleq \sum_{i=1}^{k}|v_j|\exp\left(\frac{|\alpha|}{1-|\alpha|}|\theta_{j-1}|\right)\prod_{i=j}^{k}\exp\left(\frac{1}{1-|\alpha|}|e_i|-1\right).$$

By rearranging terms we obtain

$$U_{k+1} =$$

$$\left\{\sum_{i=1}^{k-1}|v_j|\exp\left(\frac{|\alpha|}{1-|\alpha|}|\theta_{j-1}|\right)\prod_{i=j}^{k-1}\exp\left(\frac{|e_i|}{1-|\alpha|}-1\right)\right\}\exp\left(\frac{|e_k|}{1-|\alpha|}-1\right)$$

$$+\left\{|v_k|\exp\left(\frac{|\alpha|}{1-|\alpha|}|\theta_{k-1}|\right)\right\}\exp\left(\frac{1}{1-|\alpha|}|e_k|-1\right)$$

which implies that for all $k \geq 0$

$$U_{k+1} = \exp\left(\frac{1}{1-|\alpha|}\,|e_k| - 1\right)U_k + \exp\left(\frac{1}{1-|\alpha|}\,|e_k| - 1\right)\exp\left(\frac{|\alpha|}{1-|\alpha|}\,|\theta_{k-1}|\right)|v_k|. \tag{15}$$

Squaring both sides of (15) gives

$$U_{k+1}^2 = \exp\left(\frac{2}{1-|\alpha|}\,|e_k| - 2\right) \times \left\{U_k^2 + 2U_k|v_k|\exp\left(\frac{|\alpha|}{1-|\alpha|}\,|\theta_{k-1}|\right) + \exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|\theta_{k-1}|\right)|v_k|^2\right\}$$

Hence, using the estimate $2xy \leq \varepsilon x^2 + \varepsilon^{-1}y^2$, we have, for any $\varepsilon > 0$,

$$U_{k+1}^2 \leq \exp\left(\frac{2}{1-|\alpha|}\,|e_k| - 2\right)\left\{(1+\varepsilon)U_k^2 + (1+\varepsilon^{-1})\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|\theta_{k-1}|\right)|v_k|^2\right\}$$

This implies the following result. For $k \in \mathbb{Z}_+$, denote by $\mathcal{F}_k$ the σ-algebra

$$\mathcal{F}_k = \sigma\{\Phi_0, \ldots, \Phi_k\} \qquad k \in \mathbb{Z}_+.$$

LEMMA 4.1 *Under conditions (A1)–(A3) we have, for any $\varepsilon > 0$,*

$$E(U_{k+1}^2|\mathcal{F}_{k-1}) \leq \gamma^2\{(1+\varepsilon)E(U_k^2|\mathcal{F}_{k-1}) + (1+\varepsilon^{-1})\|v_1\|_2^2\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|\theta_{k-1}|\right)\} \qquad k \geq 0.$$

Given this result we now obtain a similar result for $\exp[\{2|\alpha|/(1-|\alpha|)\}|\theta_k|]$.

LEMMA 4.2. *Under conditions (A1)–(A3) we have*

$$E\left\{\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|\theta_{k+1}|\right)|\mathcal{F}_k\right\} \leq |\alpha|\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|\theta_k|\right) + (1-|\alpha|)\left[E\left\{\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|e_1|\right)\right\}\right]^{(1-|\alpha|)^{-1}}.$$

PROOF. For all $k \geq 0$ we have, by (2),

$$\begin{aligned} E\left\{\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|\theta_{k+1}|\right)|\mathcal{F}_k\right\} &\leq E\left\{\exp\left(|\alpha|\frac{2|\alpha|}{1-|\alpha|}\,|\theta_k|\right)\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|e_{k+1}|\right)|\mathcal{F}_k\right\} \\ &= \exp\left(|\alpha|\frac{2|\alpha|}{1-|\alpha|}\,|\theta_k|\right)E\left\{\exp\left(\frac{2|\alpha|}{1-|\alpha|}\,|e_{k+1}|\right)\right\}. \end{aligned}$$

To complete the proof, use the inequality $xy \leq |\alpha| x^{|\alpha|^{-1}} + (1 - |\alpha|) y^{(1-|\alpha|)^{-1}}$. ■

Consideration of the third term in the sum (14) motivates the following result.

LEMMA 4.3. *For all $k \geq 1$ we have*

$$E\left\{\prod_{i=1}^{k} \exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right) \Big| \mathcal{F}_{k-1}\right\} \leq \gamma^2 \prod_{i=1}^{k-1} \exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right).$$

The proof follows directly from conditions (A1) and (A3) and Jensen's inequality. ■

In the following result we combine the preceding bounds to construct an appropriate test function.

PROPOSITION 4.1. *There exists an adapted process $\{(V_k, \mathcal{F}_k) : k \geq 0\}$ such that*

(i) $E(V_{k+1}|\mathcal{F}_k) \leq \rho V_k + R$ *a.s.*

(ii) $\delta y_k^2 + \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_k|\right) \leq V_k$

for all $k \in \mathbb{Z}_+$ and all initial conditions, where $\rho < 1$, $R < \infty$ and $\delta > 0$.

PROOF. Let ε be a small positive constant to be specified below, and for $k \geq 0$ make the definition

$$\begin{aligned} V_k = {} & \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_k|\right) \\ & + y_0^2\theta_0^2 \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_0|\right) \prod_{i=1}^{k} \exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right) \\ & + \varepsilon^2 E(U_{k+1}^2|\mathcal{F}_k) \\ & + \varepsilon^3 y_k^2. \end{aligned}$$

It is easily seen that the conditional expectation may be defined so that, for each k, V_k is a continuous function of $(\Phi_0, \ldots, \Phi_k)$. When $k = 0$, we define the product

$$\prod_{i=1}^{k} \exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right)$$

to be equal to 1.

For $k \geq 0$ we have by Lemmas 4.1–4.3 and equation (14)

$$E(V_{k+1}|\mathcal{F}_k) \leq |\alpha| \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_k|\right)$$

$$
\begin{aligned}
&+ \gamma^2 y_0^2\theta_0^2 \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_0|\right)\prod_{i=1}^{k}\exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right)\\
&+ \gamma^2\varepsilon^2 E(U_{k+1}^2|\mathcal{F}_k)\\
&+ \gamma^2\varepsilon^2(1+\varepsilon^{-1})\|v_1\|_2^2 \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_k|\right)\\
&+ 3\varepsilon^3\left\{E(U_{k+1}^2|\mathcal{F}_k) + \theta_0^2 y_0^2 \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_0|\right)\prod_{i=1}^{k}\exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right)\right\}\\
&+ C_1
\end{aligned}
$$

where $C_1 \leqslant \infty$ is a constant independent of (θ_0, y_0). So, after rearranging terms,

$$
\begin{aligned}
E(V_{k+1}|\mathcal{F}_k) &\leqslant \{|\alpha| + \gamma^2\varepsilon^2(1+\varepsilon^{-1})\|v_1\|_2^2\} \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_k|\right)\\
&\quad + (\gamma^2 + 3\varepsilon^3)\theta_0^2 y_0^2 \exp\left(\frac{2|\alpha|}{1-|\alpha|}|\theta_0|\right)\prod_{i=1}^{k}\exp\left(\frac{2}{1-|\alpha|}|e_i| - 2\right)\\
&\quad + \{\gamma^2(1+\varepsilon) + 3\varepsilon\}\varepsilon^2 E(U_{k+1}^2|\mathcal{F}_k)\\
&\quad + C_1\\
&\leqslant \beta V_k + C_1
\end{aligned}
$$

where $\beta = \max\{|\alpha| + \gamma^2\varepsilon^2(1+\varepsilon^{-1})\|v_1\|_2^2,\ \ \gamma^2 + 3\varepsilon^3,\ \ \gamma^2(1+\varepsilon) + 3\varepsilon\}$, and $\beta < 1$ for ε sufficiently small. ■

PROOF OF THEOREM 1.1 AND THEOREM 1.2. It is a consequence of Lemma 2.2, Proposition 4.1 and Theorem 6.1 of Meyn and Tweedie (1992) that $\mathbf{\Phi}$ is geometrically ergodic. In particular, $\mathbf{\Phi}$ is positive Harris recurrent and hence the first limit follows from the law of large numbers of Athreya and Ney (1980).

The second limit follows from Theorem 6.1 of Meyn and Tweedie (1992).

The central limit theorem and law of the interated logorithm follow from Proposition 4.1 and Theorem 9.1 of Meyn and Tweedie (1992). ■

5. STATIONARITY WITHOUT THE EXISTENCE OF MOMENTS

Suppose that condition (A3) is replaced by the following.

(A3′) For some $1 > \varepsilon > 0$,

$$
E(|v_1|^{\varepsilon}) < \infty \qquad E\{\exp(\varepsilon|e_i|)\} < \infty \qquad E\left(\frac{1}{1-|\alpha|}|e_i|\right) < 1.
$$

In this case the function

$$f(t) \triangleq E\left\{\exp\left(\frac{t}{1-|\alpha|}|e_i| - t\right)\right\}$$

is smooth in a neighborhood of the origin with

$$f'(0) = E\left(\frac{1}{1-|\alpha|}|e_i| - 1\right) < 0.$$

Hence, for all $t > 0$ suitably small,

$$\gamma = E\left\{\exp\left(\frac{t}{1-|\alpha|}|e_i| - t\right)\right\} < 1. \tag{16}$$

With $0 < t \leq \varepsilon < 1$, satisfying (16) fixed, observe that by (10) and Lemma 3.1 we have

$$\begin{aligned}|y_{k+1}|^t \leq \sum_{j=1}^{k} |v_j|^t \exp\left(\frac{t|\alpha|}{1-|\alpha|}|\theta_{j-1}|\right)\prod_{i=j}^{k}\exp\left(\frac{t}{1-|\alpha|}|e_i| - t\right)\\ + |v_{k+1}|^t + |y_0|^t \exp\left(\frac{t|\alpha|}{1-|\alpha|}|\theta_0|\right)\prod_{i=1}^{k}\exp\left(\frac{t}{1-|\alpha|}|e_i| - t\right)\\ k \in \mathbb{Z}_+.\end{aligned} \tag{17}$$

Using (17) and repeating the argument used in the proof of Proposition 4.1 gives Theorem 5.1.

THEOREM 5.1. *If conditions (A1), (A2) and (A3') hold, then (θ, y) is geometrically ergodic.*

6. MULTIDIMENSIONAL MODELS

Consider the multidimensional case (1) in which

$$y_{k+1} = a_1(k)y_k + \ldots + a_m(k)y_{k-m+1} + v_{k+1} \qquad k \geq 0. \tag{18}$$

Assume that the parameter process $\{\theta_k = (a_1(k), \ldots, a_m(k))^{\mathrm{T}}\}$ is generated by the stable linear model

$$\theta_{k+1} = F\theta_k + Ge_{k+1}. \tag{19}$$

We define for $k \in \mathbb{Z}_+$,

$$\varphi_k = (y_k, \ldots, y_{k-m+1})^{\mathrm{T}}$$

$$A_k = \begin{bmatrix} \theta_k^{\mathrm{T}} & \\ 0 & \\ I & \vdots \\ 0 & \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

so that (18) may be rewritten as

$$\varphi_{k+1} = A_k \varphi_k + b v_{k+1}. \tag{20}$$

This model will be shown to be stable by a reduction to the scalar case.

6.1. *Reduction to the scalar case*

For any positive definite matrix Q, let us denote the norm of a vector x as

$$|x|_Q = (x^{\mathrm{T}} Q x)^{1/2}$$

and the norm of a matrix A as

$$|A|_Q = \sup_{x \neq 0} \frac{|Ax|_Q}{|x|_Q}$$

We have the following simple result.

LEMMA 6.1. *If $Q > 0$ satisfies $Q = A^{\mathrm{T}} Q A + I$, then*

$$|A|_Q = \left\{1 - \frac{1}{\lambda_{\max}(Q)}\right\},$$

where $\lambda_{\max}$ denotes the largest eigenvalue of Q.

Throughout this section we assume that $Q > 0$ and $R > 0$ are solutions of the following Lyapunov equations:

$$Q = S_1^{\mathrm{T}} Q S_1 + I$$
$$R = F^{\mathrm{T}} R F + I$$

where F is the matrix introduced in (19) and S_1 denotes the shift matrix defined as

$$S_1 = \begin{bmatrix} 0 & \cdots & & 0 \\ 1 & \ddots & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & 1 & 0 \end{bmatrix}.$$

It is easy to verify that Q is a diagonal matrix with diagonal entries $(m, m-1, \ldots, 1)$.

We may now formulate a key lemma.

LEMMA 6.2. *The following inequality holds:*

$$|\varphi_k|_Q \leqslant x_k \qquad k \geqslant 0$$

where the stochastic process x is defined as $x_0 = |\varphi_0|_Q$, $\delta_0 = \alpha + \beta|\theta_0|_R$ and

$$x_{k+1} = \delta_k x_k + \bar{v}_{k+1}$$
$$\delta_{k+1} = \varsigma \delta_k + \bar{e}_{k+1}$$

and where

$$\varsigma = \left\{1 - \frac{1}{\lambda_{\max}(R)}\right\}^{1/2}, \qquad \alpha = \left\{1 - \frac{1}{\lambda_{\max}(Q)}\right\}^{1/2} = \left(\frac{m-1}{m}\right)^{1/2}$$

$$\bar{v}_k = |bv_k|_Q \qquad \bar{e}_k = \alpha(1-\varsigma) + \beta|e_k|_R$$

and $\beta = \{m/\lambda_{\min}(R)\}^{1/2}$.

PROOF. By (20) we have

$$|\varphi_{k+1}|_Q \leqslant \left|S_1 + \begin{bmatrix} \theta_k^{\mathrm{T}} \\ \vdots \\ 0 \end{bmatrix}\right|_Q |\varphi_k|_Q + |bv_{k+1}|_Q$$

Hence, using the notation introduced in the lemma, we have by Lemma 6.1

$$|\varphi_{k+1}|_Q \leqslant \left(\alpha + \left|\begin{bmatrix} \theta_k^{\mathrm{T}} \\ 0 \end{bmatrix}\right|_Q\right)|\varphi_k|_Q + \bar{v}_{k+1} \leqslant \left(\alpha + \beta\left|\begin{bmatrix} \theta_k^{\mathrm{T}} \\ 0 \end{bmatrix}\right|_R\right)|\varphi_k|_Q + \bar{v}_{k+1} \tag{21}$$

Letting $\alpha_k = \alpha + \beta|\theta_k|_R$, we then see that

$$\begin{aligned} \alpha_{k+1} &= \alpha + \beta|F\theta_k + e_{k+1}|_R \\ &\leqslant \alpha + \beta(\varsigma|\theta_k|_R + |e_{k+1}|_R) \\ &= \varsigma(\alpha + \beta|\theta_k|_R) + \alpha(1-\varsigma) + \beta|e_{k+1}|_R \\ &= \varsigma\alpha_k + \bar{e}_{k+1} \end{aligned}$$

Hence, since by assumption $\alpha_0 = \delta_0$, we have $\alpha_k \leqslant \delta_k$ for all $k \in \mathbb{Z}_+$. Also by assumptions we have $|\varphi_0|_Q = x_0$, which by (21) implies that

$$|\varphi_k|_Q \leqslant x_k \qquad k \in \mathbb{Z}_+$$ ■

6.2. *Geometric ergodicity*

With Lemma 6.2 at hand we are in a position to generalize Theorem 1.1. The multidimensional case presents some technicalities with respect to irreducibility. Under conditions (A1) and (A2), the process $\boldsymbol{\theta}$ evolving on $\mathbb{R}^m$ will be geometrically ergodic if the matrix $[F^{m-1}G|\cdots|FG|G]$ has full row rank. That is, the pair of matrices (F, G) is controllable. If this is not the case, let S denote the range space of this matrix, which is known as the controllable subspace in the linear systems literature. The set S is absorbing: if $\theta_0 \in S$, then $\theta_k \in S$ for all $k \in \mathbb{Z}_+$.

We take $X = S \times \mathbb{R}$ as the state space for the process $\boldsymbol{\Phi}$, which under condition (A1) is a Markov chain with stationary transition probabilities.

The following result follows from the same argument used in Lemma 2.2, by considering the m-step transition function P^m.

LEMMA 6.3. *Suppose that conditions (A1) and (A2) hold for the model (18),* *(19). Then the Markov chain (θ, y) is irreducible and every compact subset of X is small, and hence $\mathbf{\Phi}$ is aperiodic.*

We may now state the main result of this section, which uses the following moment conditions on the 'disturbance process' $(\boldsymbol{e}, \boldsymbol{v})$.

(A4) For some $p > 2$,

$$E(|v_1|^p) < \infty \qquad \gamma^p \triangleq E\left[\exp\left\{\frac{p\beta}{1-\mu}|Ge_1|_R - p(1-\alpha)\right\}\right] < 1.$$

THEOREM 6.1. *Suppose that conditions (A1), (A2) and (A4) hold for the Markov chain $\mathbf{\Phi}$ described above. Then the conclusions of Theorem 1.1 follow.*

PROOF. Theorem 6.1 is a consequence of Lemma 6.3, Proposition 3.1 and Proposition 4.1 applied to the bilinear model defined in Lemma 6.2. These results allow us to use the same argument that was used to prove Theorem 1.1. ∎

ACKNOWLEDGEMENT

Work supported by NSF grant ECS 8910088.

REFERENCES

ATHREYA, K. B. and NEY, P. (1980) Some aspects of ergodic theory and laws of large numbers for Harris recurrent Markov chains. *Colloquia Mathematica Societatis János Bolyai* 32, 41–56. (Nonparametric statistical inference, Budapest, Hungary.)

CHAN, K. S. (1986) Topics in nonlinear time series analysis. Ph.D. Thesis Princeton University, Princeton, NJ.

COGBURN, R. (1975) A uniform theory for sums of Markov chain transition probabilities. *Ann. Probab.* 3, 191–214.

DIEBOLT, J. and GUÉGAN, D. (1990) Probabilistic properties of the general nonlinear Markovian process of order one and applications to time series modelling. Rapport 125, Universitè Paris, November.

KARLSEN, H. A. (1990) Existence of moments in a stationary stochastic difference equation. *Adv. Appl. Probab.* 22, 129–46.

KLEIN, A., LANDAU, L. J. and SCHUCKER, D. S. Decoupling inequalities for stationary Gaussian processes. *Ann. Probab.* 10, 702–8.

LASALLE, J. and LEFSCHETZ, S. (1961) *Stability by Liapunov's Direct Method*, New York: Academic Press.

MEYN, S. P. (1991) A stability theory for Feller Markov chains with applications to adaptive control and time series analysis. Technical report, University of Illinois, Urbana-Champaign, IL.

—— and GUO, L. (1990) Adaptive control of time varying stochastic systems. In *Proceedings of the 11th IFAC World Congress*, Vol. 3, pp. 198–202 (eds V. Utkin and O. Jaaksoo). Tallinn, Estonia, USSR.

—— and TWEEDIE, R. L. (1992) Criteria for stability of Markovian processes I: discrete time chains. To appear, *Adv. Appl. Probab.*, September, 1992.

—— and —— (1991) Markov chains and stochastic stability. To appear in *Control and Communication in Engineering*, New York: Springer-Verlag.

MOKKADEM, A. (1987) Sur un modèle autorégressif non linéaire, ergodicité et ergodicité géométrique. *J. Time Series Anal.* 8, 195–204.

NUMMELIN, E. (1984) *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge, MA: Cambridge University Press.

POURAHMAD, M. (1986) On stationarity of the solution of a doubly stochastic model. *J. Time Series Anal.* 7, 123–31.

RUDIN, W. (1974) *Real and Complex Analysis*, 2nd edn. New York: McGraw-Hill.

TJØSTHEIM, D. (1986) Some doubly stochastic time series models. *J. Time Series Anal.* 7, 51–72.

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

TUOMINEN P. and TWEEDIE, R. L. (1979) Markov chains with continuous components. *Proc. Lond. Math. Soc.* 3.

TWEEDIE, R. L. (1976) Criteria for classifying general Markov chains. *Adv. Appl. Probab.* 8, 737–71.